J. theor. Biol. (2003) **220,** 261–269
doi:10.1006/jtbi.2003.3164, available online at http://www.idealibrary.com on IDEAL®

The Role of Dimerization in Noise Reduction of Simple Genetic Networks

R. BUNDSCHUH*†, F. HAYOT† AND C. JAYAPRAKASH†

†*Department of Physics, The Ohio State University, Columbus, OH* 43210-1106, *U.S.A.*

(*Received on* 3 *June* 2002, *Accepted in revised form on* 11 *September* 2002)

Fluctuations are an intrinsic property of genetic networks due to the small number of interacting molecules. We study the role of dimerization reactions in controlling these fluctuations in a simple genetic circuit with negative feedback. We compare two different pathways. In the dimeric pathway the proteins to be regulated form dimers in solution that afterward bind to an operator site and inhibit transcription. In the monomeric pathway monomers bind to the operator site and then recruit another monomer to form a dimer directly on the DNA. We find that while both pathways implement the same negative feedback mechanism, the protein number fluctuations in the dimeric pathway are drastically reduced compared to the monomeric pathway. This difference in the ability to reduce fluctuations may be of importance in the design of genetic networks.

1. Introduction

Information processing in biological cells is often implemented by a genetic network. The state of such a network is represented by the concentrations and locations of the different species of molecules. A cell typically contains only small numbers of molecules of each kind. The interactions between these molecules occur in a random fashion. Thus, there is an intrinsic stochastic nature to biological information processing (McAdams & Shapiro, 1995; McAdams & Arkin, 1997,1999). In order to prevent detrimentally large fluctuations in the number of molecules of some species, the genetic network itself can contain negative feedback mechanisms that suppress these fluctuations (Savageau, 1974; Becskei & Serrano, 2000).

*Corresponding author. Tel.: +1-614-688-3978; fax: +1-614-292-7557.
E-mail address: bundschuh@mps.ohio-state.edu (R. Bundschuh).

The most basic design for a noise suppressing feedback mechanism consists of a gene that codes for a protein that in turn inhibits transcription of the gene itself (Hasty *et al.*, 2000; Thattai & van Oudenaarden, 2001). The inhibition of transcription typically involves binding of the protein to an operator region in a way that competes with the binding of RNA polymerase. Often, the proteins bind as dimers or higher-order multimers since this increases the stability of the transcription factor DNA complex and allows for tighter regulation. Binding DNA in the form of dimers allows for two possible scenarios: either the dimers assemble in solution and then bind to DNA as a preformed dimer (Ptashne *et al.*, 1980) or a monomer binds to DNA and afterward recruits a second monomer (Berger *et al.*, 1998) to form the DNA–dimer complex.

In this paper, we therefore explore two different regulatory networks. Our aim is to understand their effectiveness not only in

0022-5193/03/$35.00

controlling the mean number of proteins but in reducing intrinsic stochastic fluctuations. Both models implement the same basic negative feedback mechanism described above. They are based on the well-studied λ-phage in *E. coli*. In the first model, the protein dimerizes in the intracellular region first and then binds to an operator site on the DNA as realized in λ-phage (Ptashne *et al.*, 1980). In the second network we consider a case where the monomers dimerize only on the DNA as exemplified by the leucine zippers (Berger *et al.*, 1998). We study the reaction equations for the two models with biologically reasonable reaction rates chosen to make the comparison of the fluctuations in the protein concentration meaningful. The main result is that although the average number of produced proteins is the same in both networks the intracellular dimerization reaction reduces the fluctuations dramatically beyond what the feedback mechanism can do by itself. We elucidate the mechanism of this reduction. It is based on the buffering by the fast dimerization reaction of fluctuations arising from translation and transcription. While there are many reasons why evolution may choose one regulatory mechanism over another our work points to an important added advantage which results from choosing certain mechanisms: in cases where the smallness of protein fluctuations in a negative feedback scheme is an important design requirement, our study points out one general mechanism by which this can be accomplished.

2. Reactions and Methods

REACTIONS

We will start by specifying the chemical species and reactions that we will study. These reactions represent a generic network with negative feedback, such as the one found in the control circuit for the λ repressor protein cI of phage λ in *E. coli*. Our aim is to point out a generic feature of genetic networks of this type and we are not interested in describing specifics of the λ repressor system. Nevertheless, we will take advantage of the variety of studies of the specific λ repressor system (Arkin *et al.*, 1998; Hasty *et al.*, 2000; Thattai & van Oudenaarden, 2001) in guiding our choice of reasonable model parameters.

We first describe the reactions common to both our networks representing the production and decay of the proteins. In order to describe transcription and translation, we consider the free DNA D coding for the protein P to be regulated, the RNA polymerase R, the complex D^* of RNA polymerase bound to the promoter site on the DNA molecule, and the mRNA M. Then, transcription and translation are modeled as the single effective irreversible reactions

$$D^* \xrightarrow{k_1} D + M + R \tag{1}$$

and

$$M \xrightarrow{k_2} M + P \tag{2}$$

with reaction rates k_1 and k_2, respectively. The mRNA as well as the protein are degraded through the reactions

$$M \xrightarrow{k_3} \emptyset, \tag{3}$$

$$P \xrightarrow{k_4} \emptyset \tag{4}$$

with rate constants k_3 and k_4, respectively. We assume that even though the protein forms dimers and complexes with the DNA the degradation of the protein occurs predominantly in the monomeric form (Gottesman & Maurizi, 1992). The complex between DNA and RNA polymerase is formed in the equilibrium reaction

$$D + R \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} D^* \tag{5}$$

with forward and backward rates k_5 and k_{-5}, respectively.

The negative feedback mechanism is introduced through a single repressive operator site. A protein bound to this operator site prevents the binding of the RNA polymerase to its promoter site and thus inhibits transcription. In the systems we are interested in the protein binds to the operator site as a dimer and we

will denote the DNA–dimer complex by Q. We investigate two different pathways for the formation of the DNA–dimer complex. In the first network, as is realized in λ phage of *E. coli*, the monomers P can form dimers P_2 in solution and these dimers can then bind to the operator site. This dimeric pathway is described by the reactions

$$2P \underset{k_{-6}}{\overset{k_6}{\rightleftharpoons}} P_2, \tag{6}$$

$$D + P_2 \underset{k_{-7}}{\overset{k_7}{\rightleftharpoons}} Q, \tag{7}$$

where the forward and backward rates of these two equations are denoted by k_6, k_{-6}, k_7, and k_{-7}, respectively.

The alternative pathway is motivated by the behavior of leucine zipper proteins (Park *et al.*, 1996; Berger *et al.*, 1998). In this pathway, a protein monomer first forms a complex with the DNA by binding to the operator site as a monomer. We denote this complex by T. The monomer–DNA complex can then recruit another monomer to form the DNA–dimer complex Q. This monomeric pathway is described by reactions (1)–(5) and

$$D + P \underset{k_{-8}}{\overset{k_8}{\rightleftharpoons}} T \tag{8}$$

and

$$T + P \underset{k_{-9}}{\overset{k_9}{\rightleftharpoons}} Q, \tag{9}$$

where k_8, k_{-8}, k_9, and k_{-9} are the forward and backward rates, respectively.

In principle, both pathways exist in every genetic network with negative feedback of the type studied. The relations between the different rate constants $k_6, \ldots, k_9$ and $k_{-6}, \ldots, k_{-9}$ determine if dimerization in solution (Ptashne *et al.*, 1980) or dimerization on the DNA (Berger *et al.*, 1998) is the predominant pathway. Since we want to study the implications of choosing one pathway over the other, we will consider both pathways one at a time. To this end, we will first consider the reactions eqns (1)–(7) by themselves. We will call them model D. Then, we will compare the behavior of model D with model M defined by eqns (1)–(5), (8) and (9).

RATE CONSTANTS

In order to choose biologically reasonable values for the rate constants $k_1, \ldots, k_{-9}$ we use the specific implementation of the genetic circuit by the λ repressor protein in *E. coli*. The transcription reaction eqn (1) is limited in rate by the isomerization of the RNA polymerase from the closed into the open form (McClure, 1980). While the inhibitory operator site of interest in this work is the operator site O_{R3} the RNA polymerase isomerization rate depends also on the status of operator O_{R2}. If O_{R2} is not occupied, the rate is rather small while the rate is much larger if O_{R2} is occupied. We want to understand the generic features of a genetic circuit with negative feedback and are not interested in describing the specifics of λ repressor. Thus, we omit the complication of the operator site O_{R2} and choose as a biologically reasonable number the higher rate corresponding to an occupied operator O_{R2}. This yields $k_1 = 0.0078\,s^{-1}$ (Hawley & McClure, 1982).

For the rate k_3 of mRNA degradation we use a typical half-life of 3 min (Alberts *et al.*, 1994). This yields $k_3 = \ln 2/3\,\text{min} = 0.0039\,\text{s}^{-1}$. In order to achieve the average number of 11 proteins per open complex (Shea & Ackers, 1985) this also implies $k_2 = 11k_3 = 0.043\,\text{s}^{-1}$. Finally, we extract the protein monomer decay rate $k_4 = 0.0007\,s^{-1}$ from Arkin *et al.* (1998) chosen to match the degradation rate of λ repressor due to cell-growth induced dilution.

For the other reactions it is more difficult to obtain precise rate constants. The equilibrium constants of eqns (5)–(7) are $K_5 = k_{-5}/k_5 = 8nM$ (Shea & Ackers, 1985) $K_6 = k_{-6}/k_6 = 20\,\text{nM}$ (Sauer, 1979), and $K_7 = k_{-7}/k_7 = 75\,\text{nM}$ (Ackers *et al.*, 1982), respectively. As long as these reactions are much faster than eqns (1)–(4) the actual values of the rate constants are only of minor importance as long as they respect the experimentally determined equilibrium constants (Bundschuh, Hayot & Jayaprakash, in prep.); our choices of the individual rate constants will, however, be guided by experimental

observations where available. For reaction eqn (5) rate constants of $0.056\,s^{-1}(nM)^{-1}$ and $0.2\,s^{-1}$ are known for the RNA polymerase of phage T7 (Ujvari & Martin, 1996). Assuming that *E. coli* RNA polymerase follows similar kinetics but with a slightly different rate constant we choose $k_5 = 0.038\,s^{-1}(nM)^{-1}$ and $k_{-5} = 0.3\,s^{-1}$. The binding kinetics of λ repressor dimer to DNA should be of a similar order of magnitude but with an equilibrium constant of $K_7 = 75\,nM$. Thus, we choose $k_7 = 0.012\,s^{-1}(nM)^{-1}$ and $k_{-7} = 0.9\,s^{-1}$. While the individual rate constants for the dimerization reaction eqn (6) are not known, to the best of our knowledge, we make the reasonable choice $k_6 = 0.025\,s^{-1}(nM)^{-1}$ and $k_{-6} = 0.5\,s^{-1}$ of Arkin *et al.* (1998).*

The rate constants k_8, k_{-8}, k_9, and k_{-9} have to be chosen such that a sensible comparison of model D and model M is possible. If $K_8 = k_{-8}/k_8$ and $K_9 = k_{-9}/k_9$ are the respective equilibrium constants of reactions eqns (8) and (9) this means that the total equilibrium constant between the dimer–DNA complex and the monomer has to be identical in the two models, i.e. $K_8K_9 = K_6K_7 = 1500\,nM^2$. Since apart from this constraint we do not know which value of the individual equilibrium constants should be chosen, we vary K_9 in the range from 2 to 400 nM and adapt K_8 appropriately. It will turn out that our results do not change significantly over the whole range of values for the equilibrium constant K_9. Again, the individual rates k_8, k_{-8}, k_9, and k_{-9} should not matter as long as they are faster than transcription and translation and respect the equilibrium constants. To be specific, we choose $k_{-9} = 0.2\sqrt{K_9}s^{-1}(nM)^{-1/2}$, $k_9 = 0.2/\sqrt{K_9}s^{-1}(nM)^{-1/2}$, $k_{-8} = 7.75/\sqrt{K_9}\,s^{-1}(nM)^{1/2}$, and $k_8 = 0.00526\sqrt{K_9}\,s^{-1}(nM)^{-3/2}$.

SIMULATIONS

The reactions outlined above were numerically studied with the Gillespie algorithm (Gillespie, 1977). The intrinsic quantities the Gillespie algorithm acts on are not the concentrations but the actual numbers of molecules of each species. Thus, the Gillespie algorithm can correctly take all effects of the discrete number of molecules present in a cell at a given time into account. Instead of rate constants k_i, all reactions are characterized in the Gillespie framework by reaction probabilities c_i per second and per molecule. These are related to the rate constants k_i by powers of the volume V of the system where the exponent of V depends on the reaction (Gillespie, 1977). Since the volume V of an *E. coli* cell can be conveniently written as $V = 1/1\,nM$, the actual numerical values of the reaction probabilities c_i in the Gillespie framework are identical to the numerical values of the reaction rates k_i as long as the latter are expressed in nM. At the same time concentrations measured in nM also directly correspond to the number of molecules of the respective species in an *E. coli* cell.

Within the Gillespie algorithm the state of the system at a given time is described by the numbers of molecules of each species. In each step of the simulation, the reaction probabilities c_i per second and per molecule for each reaction equation are multiplied by the numbers of molecules of the species on the left-hand side of the corresponding reaction equation in order to obtain the total probability per second for each reaction to occur. Following these probabilities the next reaction to occur in the system is chosen randomly and it is determined how many seconds (or fractions thereof) pass until this reaction takes place. Then, the chosen reaction equation is used to update the numbers of molecules of each species, i.e. the numbers of molecules on the left-hand side of the reaction equation are reduced and the numbers on the right-hand side are increased. If the preset simulation time has not been reached the reaction probabilities for the new numbers of molecules are calculated and the next reaction is chosen.

For every choice of the rate constants, the Gillespie algorithm was used to simulate 10 000 independent time courses of the genetic network starting with a single DNA molecule and 30 molecules of RNA polymerase (McClure, 1983). Each of the 10 000 instances was run until

*In Arkin *et al.* (1998) the rate constant k_6 is actually by a factor of 2 higher which is probably due to the combinatorial factor of two associated with a reaction involving two identical molecules on the left-hand side of eqn (6) (Gillespie, 1977).

50 000 s of reaction time were simulated. By observing the number of self-regulating proteins as a function of reaction time, it was ensured that the system is well in the stationary state after this time (as also suggested by the slowest characteristic time $1/k_4 \approx 1500$ s of the system.) Figure 1(a) shows time courses of the number of monomers p during the last three seconds of three different instances. From each of the 10 000 instances the final number of monomers p was extracted as indicated by the circles for the three instances shown in Fig. 1(a). The histogram of these 10 000 independent numbers shown in Fig. 1(b) approximates the distribution $P(p)$ of the number p of monomers in the stationary state. From this distribution the average number $\langle p \rangle$ of monomers and the Fano factor

$$f \equiv \frac{\langle \Delta p^2 \rangle}{\langle p \rangle} \equiv \frac{\langle p^2 \rangle - \langle p \rangle^2}{\langle p \rangle} \tag{10}$$

characterizing its fluctuations are calculated.

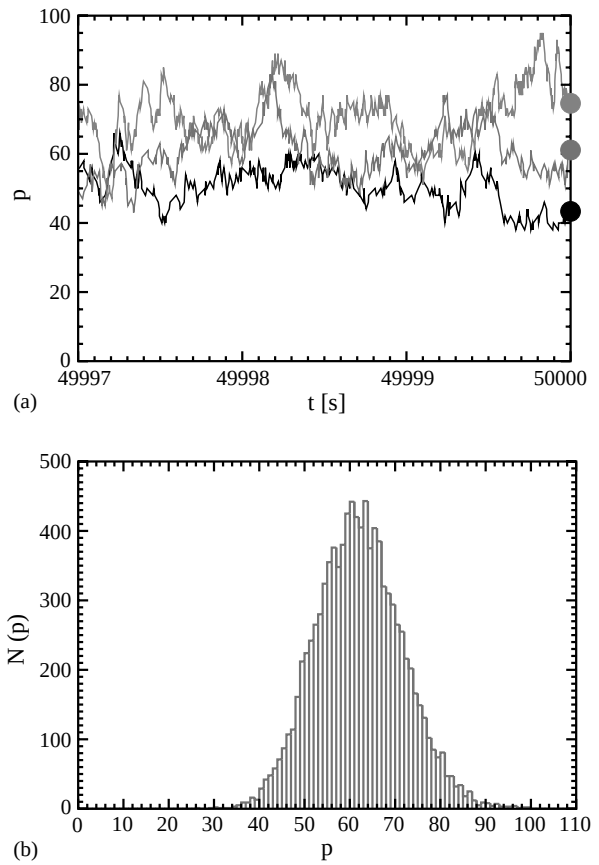

FIG. 1. Simulated number of monomers for model D. (a) shows the number of monomers p as a function of time t for the last three seconds of three independent simulations of model D with a total length of 50 000 s each. The circles on the right-hand side denote the final values of the monomer number p for each instance. In (b) the histogram of 10 000 of these final protein numbers is shown. From this histogram the average value $\langle p \rangle$ of the number of monomers and the Fano factor f are calculated.

Results

We determine the stationary values of the average number $\langle p \rangle$ of monomers and its fluctuations for two different gene network models with negative feedback through computer simulations with the Gillespie algorithm (Gillespie, 1977). The two models studied correspond to the two possible pathways of binding a protein dimer to the operator site on the DNA molecule. Model D assumes that the protein dimerizes in solution and that these dimers then bind to the operator site. Model M on the contrary assumes that a monomer binds to the operator site and then recruits another monomer from the solution to form a dimer. We choose biologically reasonable reaction rates for both models guided by the measured rate constants of the λ repressor gene cI in *E. coli*. One parameter remains undetermined, namely the equilibrium constant K_9 between the complexes of the protein monomer and DNA and the protein dimer and DNA, respectively.

Figure 2 shows the average number of protein monomers. The result for model D is $\langle p \rangle \approx 62.6$. It is indicated as the dashed line. The circles show the dependence of the average number of monomers on the equilibrium constant K_9 in model M. Small values of the equilibrium constant K_9 correspond to a situation where the equilibrium between monomeric and dimeric molecules bound to the operator site is far on the side of the dimers. In this case, the genetic network regulates the average number of monomers in model M to the same value as in model D. This confirms, that purely from the point of view of the average number of monomers, the dimeric pathway in model D and the monomeric pathway in model M are equivalent.

The difference in the feedback mechanism between models D and M is that in model M not only the DNA–dimer complex Q but also the

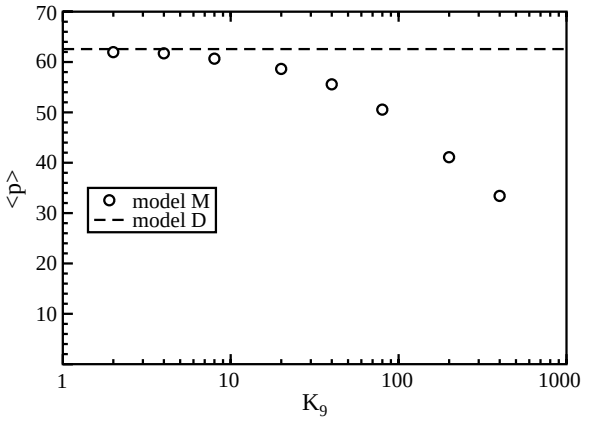

FIG. 2. Average number $\langle p \rangle$ of protein monomers. The dashed line corresponds to the value $\langle p \rangle \approx 62.6$ for model D with dimerization in solution. The circles show the dependence of the average number $\langle p \rangle$ of proteins as a function of the equilibrium constant K_9 in model M with dimerization only on the DNA. For small values of the equilibrium constant K_9 there is no difference between the average number of proteins in models M and D.

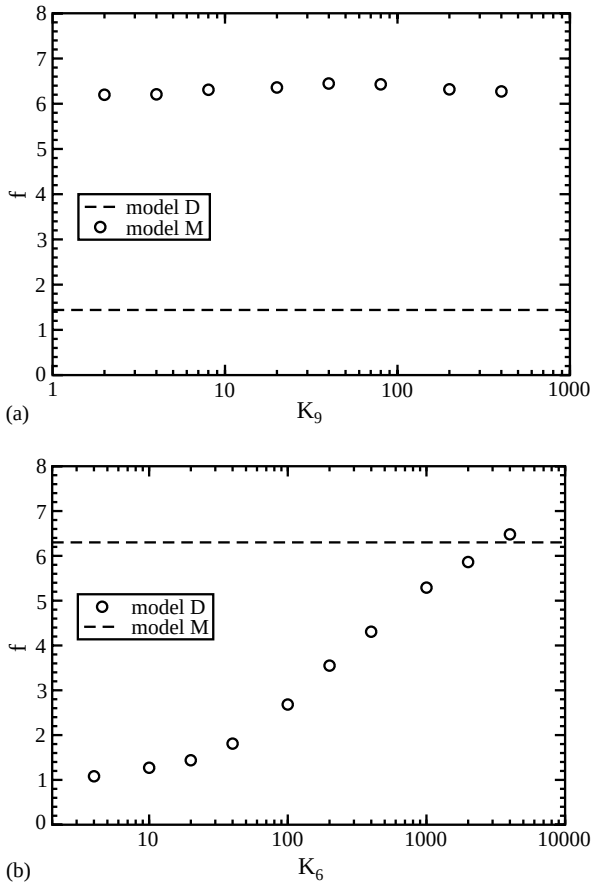

FIG. 3. Fano factor $f = \langle \Delta p^2 \rangle / \langle p \rangle$ characterizing the fluctuations of the number of protein monomers. In (a) circles show the Fano factor $f \approx 6.3$ as a function of the equilibrium constant K_9 in model M with dimerization only on the DNA. The dashed line corresponds to the value $f \approx 1.4$ for model D with dimerization in solution at the biological value $K_6 = 20$ nM of the equilibrium constant of the dimerization reaction. In (b) the equilibrium constant K_6 of the dimerization reaction in model *D* is varied. The circles show the Fano factor of model D while the dashed line indicates the Fano factor $f \approx 6.3$ of model M.

DNA–monomer complex T competes with the RNA polymerase for the DNA. If the number of protein monomers in solution is fixed the ratio between the DNA–dimer complex and free DNA remains unchanged as K_9 is varied since we hold the product $K_8 K_9$ constant. However, the ratio between the DNA–monomer complex and free DNA increases as K_9 becomes larger. This leads to an increased inhibition of transcription that in turn results in the observed reduction in the average number of protein monomers in solution.

We characterize the magnitude of the fluctuations in terms of the Fano factor $f = \langle \Delta p^2 \rangle / \langle p \rangle$. The Fano factor equals one for a Poisson distribution and deviations from one indicate how much larger or smaller the fluctuations are in comparison with the fluctuations in a Poisson process.

Figure 3(a) shows the Fano factors resulting from our simulations. Model D exhibits a Fano factor of $f \approx 1.4$ indicated by the dashed line. In model M, we find that the Fano factor is approximately 6.3 relatively insensitive to the value of the equilibrium constant K_9 over a wide range as seen in Fig. 3(a). Thus, we conclude that the intrinsic stochastic fluctuations in the number of protein molecules in the monomeric form, as measured by the Fano factor, are four times larger in model M than in model D.

This difference in Fano factor depends on the equilibrium constant K_6 of the dimerization reaction. If K_6 becomes very large the equilibrium between monomers and dimers is completely on the side of the monomers and only very rarely a dimer is in solution that can then bind to the DNA. This effectively restores the situation of model M and thus a larger Fano factor is expected. We test this hypothesis by varying the equilibrium constant K_6 in model D while keeping the product $K_6 K_7$ constant in

order to not change the underlying regulatory network. We verify that the average number $\langle p \rangle$ of monomers does not depend on K_6 unless K_6 becomes very small. However, the fluctuations as quantified by the Fano factor shown in Fig. 3(b) increase with increasing K_6 as expected. They approach the Fano factor $f \approx 6.3$ of model M for very large K_6 while they approach $f \approx 1$ for small K_6. We conclude that within model D the efficiency of the reduction of fluctuations can be adjusted by changing the equilibrium constant K_6 of the dimerization. The value $K_6 = 20\,\mathrm{nM}$ for λ repressor already achieves nearly maximal suppression of the fluctuations.

Discussion

Our numerical studies indicate that there is a significant difference between the fluctuations in the number of protein monomers between models with monomeric and dimeric DNA binding. The negative feedback mechanism itself of the genetic networks studied behaves identically whether the protein dimers bound to the operator site are bound via dimerization in solution and successive binding to the DNA or via binding of a monomer to the DNA with successive recruitment of a second monomer to form the dimer directly on the operator site. This is demonstrated by showing that the stationary average number of protein monomers does not depend on the pathway used for dimer formation on the operator site (at least for certain compatible choices of the reaction rates). However, the fluctuations in the number of protein monomers are much stronger in the monomeric pathway than in the dimeric pathway. Thus, if controlling fluctuations in the monomer concentration is a critical design criterion for a genetic network, it is much more effective to employ the negative feedback through the dimeric pathway than through the monomeric pathway. We note in passing that we can also study the fluctuations of the number p_2 of dimers (in the dimeric pathway.) We find for the Fano factor $f_2 = \langle \Delta p_2^2 \rangle / \langle p_2 \rangle$ of the dimer population $f_2 \approx 6.5$. This is comparable to the Fano factor for fluctuations of the monomer concentration in the monomeric pathway. A different measure of the fluctuations such as the ratio of the standard deviation to the mean would indicate that the fluctuations in the dimeric pathway are somewhat lower. Therefore, if controlling the fluctuations in the binding species were the only criterion the dimeric pathway only provides a slight advantage.

We now provide a qualitative explanation for the effectiveness of the dimeric pathway in controlling the fluctuations in the number of proteins. The difference between the two pathways is that in the dimeric pathway the equilibrium between protein monomers and protein dimers provides an additional buffer that reduces the fluctuations of the number of monomers. A fluctuation in the number of translated proteins is largely absorbed by the dimerization. Although the number of monomers is the same in both pathways, in the dimeric pathway there are actually $\langle n \rangle \equiv \langle p \rangle + 2\langle p_2 \rangle$ proteins in the solution where p_2 denotes the number of protein dimers. In our simulation of model D we estimate $\langle n \rangle \approx 455$. The two relations $\langle n \rangle \equiv \langle p \rangle + 2\langle p_2 \rangle$ and $\langle p \rangle^2 = 20\langle p_2 \rangle$ yield that every protein translated (i.e. every increase of $\langle n \rangle$ by one) increases the number $\langle p \rangle$ of monomers only by 0.074. The rest of the proteins are absorbed by the dimer population in solution. This dimer population in solution is absent in the case of the monomeric pathway. This buffering should reduce the fluctuations in the case of the dimeric pathway by a factor of $1/0.074 \approx 13.5$ in comparison with the monomeric pathway. This effect becomes weaker as the equilibrium between monomers and dimers shifts more toward the monomer side as is observed in Fig. 3(b). It only depends on the existence of the dimerization reaction. Thus, if both pathways are present simultaneously and K_9 is small such that inhibition by monomers bound to the DNA is negligible the reduction in fluctuations is the same as in model D as we verified by a simulation of a model with both pathways. If K_9 becomes large, the inhibition by monomers bound to the DNA reduces the number of monomers $\langle p \rangle$ exactly like in model M. This reduction makes the buffering mechanism of the dimerization reaction less efficient. Nevertheless, the Fano factor for the combined model remains close to its value for model D: for $K_9 = 400$

where $\langle p \rangle \approx 32.6$ the Fano factor is $f \approx 1.8$ and thus much smaller than in model M.

The actual observed reduction of the fluctuations is only by a factor of 4 which is less than the expected 13.5 according to the argument above. The reason is that we have not taken into account the additional fluctuations associated with the monomer–dimer equilibrium. The dimerization reaction eqn (6) by itself with a fixed total number of $n = 455$ exhibits monomer number fluctuations characterized by a Fano factor of $f_{\mathrm{dim}} \approx 0.9$ at $K_6 = 20$ nM. This value can be obtained analytically or numerically by simulating the dimerization reaction eqn (6) by itself with the Gillespie algorithm. Together with the fluctuations $f_{trans} \approx 6.3/13.5 \approx 0.5$ introduced by the translation reaction discussed above this explains the measured Fano factor $f \approx f_{dim} + f_{trans} \approx 1.4$ of the complete dimeric system. We conclude that the fluctuations of the number of protein monomers due to the dimerization reaction are the larger contribution to the Fano factor in the dimeric pathway.

Conclusions

We have studied a model system for gene autoregulation, which relies on an inhibitory feedback effected by the formation of a dimer–DNA complex. It derives from phage λ and appears typical for a large class of self-regulatory biological systems. The values we have used for the rate constants of the reactions considered are derived from experiment wherever possible. We do not expect our general conclusions to be too sensitive to variations in the individual rate constants as long as the DNA binding and dimerization reactions can be considered fast compared to the translation and transcription reactions.

There are two possible ways to implement negative feedback by protein dimers, where dimerization occurs either in solution or on the DNA itself through successive binding of monomers. Although the negative feedback pathway is common, and the average number of protein monomers nearly the same, we show that fluctuations in their number differ dramatically from one way to the other. We have verified, that the same scenario holds if dimers are replaced by tetramers and we believe our result to hold for any order of multimers.

We have highlighted the crucial role played by intracellular dimerization reactions in absorbing fluctuations which occur in the number of produced proteins, and show that fluctuations of the latter are actually dominated by equilibrium fluctuations of the dimerization reaction itself.

In general, many different and potentially contradictory aspects have to be taken into account in designing a specific genetic network. If the control of fluctuations of the monomer population is a major design goal our study shows that this strongly favors the implementation of the dimeric over the monomeric pathway.

We gratefully acknowledge insightful discussions with T. Dijkstra and T. Hwa.

REFERENCES

Ackers, G. K., Johnson, A. D. & Shea, M. A. (1982). Quantitative model for gene regulation by lambda phage repressor. *Proc. Natl. Acad. Sci. U.S.A.* **79,** 1129–1133.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. (1994). *Molecular Biology of the Cell*. New York: Garland Publishing.

Arkin, A., Ross, J. & McAdams, H. H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* **149,** 1633–1648.

Becskei, A. & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature* **405,** 590–593.

Berger, C., Piubelli, L., Haditsch, U. & Bosshard, H. R. (1998). Diffusion-controlled DNA recognition by an unfolded, monomeric bZIP transcription factor. *FEBS Lett.* **425,** 14–18.

Gillespie, D. T. (1977). Stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81,** 2340–2361.

Gottesman, S. & Maurizi, M. R. (1992). Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Rev.* **56,** 592–621.

Hasty, J., Pradines, J., Dolnik, M. & Collins, J. J. (2000). Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 2075–2080.

Hawley, D. K. & McClure, W. R. (1982). Mechanism of activation of transcription initiation from the lambda PRM promoter. *J. Mol. Biol.* **157,** 493–525.

McAdams, H. H. & Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 814–819.

McAdams, H. H. & Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.* **15,** 65–69.

McAdams, H. H. & Shapiro, L. (1995). Circuit simulation of genetic networks. *Science* **269,** 650–656.

McClure, W. R. (1980). Rate-limiting steps in RNA chain initiation. *Proc. Natl. Acad. Sci. U.S.A.* **77,** 5634–5638.

McClure, W. R. (1983). A biochemical analysis of the effect of RNA polymerase concentration on the *in vivo* control of RNA chain initiation frequency. In: *Biochemistry of Metabolic Processes* (Lennon, D. L. F., Stratman, F. W. & Zahltne, R. N., eds), pp. 207–217. New York: Elsevier Science Publishing Co.

Park, C., Campbell, J. L. & Goddard III, W. A. (1996). Protein stitchery: design of a protein for selective binding to a specific DNA sequence. *J. Am. Chem. Soc.* **118,** 4235–4239.

Ptashne, M., Jeffrey, A., Johnson, A. D., Maurer, R., Meyer, B. J., Pabo, C. O., Roberts, T. M. & Sauer, R. T. (1980). How the lambda repressor and cro work. *Cell* **19,** 1–11.

Savageau, M. A. (1974). Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* **252,** 546–549.

Sauer, R. T. (1979). *Molecular Characterization of the Lambda Repressor and Its Gene CI.* Cambridge: Harvard University Press.

Shea, M. A. & Ackers, G. K. (1985). The OR control system of bacteriophage lambda. A physical–chemical model for gene regulation. *J. Mol. Biol.* **181,** 211–230.

Thattai, M. & van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 8614–8619.

Ujvari, A. & Martin, C. T. (1996). Thermodynamic and kinetic measurements of promoter binding by T7 RNA polymerase. *Biochemistry* **35,** 14 574–14 582.