

SCIENTIFIC REPORTS



OPEN

Intrinsic phenotypic stability of a bi-stable auto regulatory gene

Azim-Berdy Besya & Andreas Grönlund

Received: 15 October 2015

Accepted: 24 February 2016

Published: 10 March 2016

Even under homogenous conditions clonal cells can assume different distinct states for generations to follow, also known as epigenetic inheritance. Such long periods of different phenotypic states can be formed due to the existence of more than one stable state in the molecule concentration, where the different states are explored through molecular fluctuations. By formulating a single reaction variable representing the birth and death of molecules, including transcription, translation and decay, we calculate the escape time from the phenotypic states attained from autocatalytic synthesis through a Fokker-Planck formulation and integration of an effective pseudo-potential. We calculate the stability of the phenotypic states both for cooperative binding feedback and dimer binding feedback, resulting in non-linear decay.

The cellular machinery often work with few molecules at hand. A gene may exist as one single copy on the DNA and the copy number of the messenger RNA and following protein are typically very low¹. Chemical reactions are probabilistic – before a reaction can take place the participating molecules first need to find each other through random motion, chemical bonds form at random. Molecular decay is also random. Combining low copy number and random birth and death of molecules make biochemical reaction networks noisy, even if we eliminate all external variations in the environment of a given biochemical process. Two clonal cells placed in the same environment may therefore at times exhibit very different behaviour. Fluctuations generated from birth and death of molecules in a reaction network are commonly referred to as *intrinsic fluctuations* whereas *extrinsic fluctuations* are fluctuations of the environment that the studied reaction network are embedded in, i.e. variations in the reaction rates of the reaction network^{2,3}. The magnitude of the molecular fluctuations can be altered by a change in the reaction rates and by regulatory processes. Negative autoregulation is shown both to reduce noise⁴ and to amplify fluctuations if the feedback mechanism is delayed^{5–7} or if the feedback is strong⁸. The fidelity of regulatory control is also limited by the molecular noise associated with the process of transmitting biochemical signals⁹ and by finding the regulatory sites^{10–12}. However, there are no a priori reasons to assume molecular fluctuations only being detrimental for an organism. In theory noise could be exploited for enhancing the sensitivity of regulatory control processes¹³. Moreover, there are many situations where noise is an essential component of a process. Overcoming potential barriers by thermal fluctuations are what drives chemical reactions in the first place and in are this respect of course essential, but potential wells and multi-stable states can also be formed by the dynamics of molecule synthesis, e.g. if the synthesis is enhanced by its own presence. The consequence of such feedback is that there are two or more stable concentrations that the different molecule species of the reaction network may attain. Autocatalytic protein synthesis is shown both theoretically and experimentally to generate meta stable states and the consequence of such seemingly simple positive feedback is that genetically identical individuals may even at identical homogenous conditions attain different phenotypic states^{14–18}. The states are referred to as phenotypic states, since the protein concentration may increase (or decrease) by orders of magnitude and last for generations. E.g. the spontaneous exit of the lysogenic state for the lambda phage is only about 10^{-7} – 10^{-5} per cell and generation^{18,19} showing that dynamic stability can be on par with the stability of DNA^{20,21}. External perturbations, e.g. by sudden changes in the reaction rates, can push the biochemical reaction network from one meta-stable state to another. Random switching from one stable state to another may occur from extrinsic fluctuation or perturbations but also as a consequence of the intrinsic noise generated from the birth and death of molecules, without any external signal preceding the escape of the state. Random switching is exploited in bacteria as a stochastic survival strategy to blindly anticipate variations in the environment, where different rates of switching is resulting in different fitness²². Note that the random switching mechanism display an intrinsic time scale which is not triggered by changes in the environment. Random switching is also observed to induce antibiotic resistance in otherwise identical cells²³.

Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden. Correspondence and requests for materials should be addressed to A.G. (email: andreas.gronlund@umu.se)

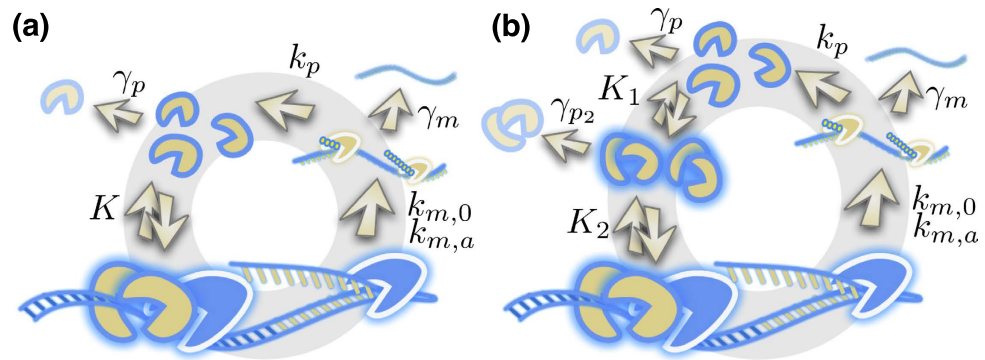


Figure 1. Illustration of the regulatory loop of the two studied motifs of positive autoregulation. The illustration display synthesis of messenger RNA from DNA, synthesis of proteins from messenger RNA and binding of the protein to the promoter to its own gene. The non-linearity of the feedback regulation is the same for the two studied motifs but is in (a) achieved by a cooperative binding of two proteins to the promoter and in (b) by forming dimers before binding to the promoter. An important difference, which is illustrated in (b), is that feedback mediated by dimers also imply that proteins decay not only as monomers but also as dimers. Whenever the promoter is occupied by proteins, and active, the synthesis rate of messenger RNA is boosted to $k_{m,0} + k_{m,a}$ from the basal rate $k_{m,0}$. Protein translation is proceeding with the rate k_p . The dissociation binding constants between protein monomers and dimers K_1 and the promoter K_2 (dimers) and K (monomers). Decay rate constants of messenger RNA γ_m , proteins γ_p and protein dimers γ_{p_2} .

Traditionally, calculations of the stability of metastable states are done for Brownian particles performing random walks in a potential well, or similar systems, where the noise is not a function of the position in space. For autocatalytic transcription factors, where metastable states in the concentration can be generated as a result of the protein exhibiting positive feedback on its own synthesis, the birth (and death) of molecules also generates fluctuations. Thus, not only is the potential a function of the protein concentration but also the noise, which makes analytical progress challenging. The analogue of such noise for a brownian particle escaping a potential well would be to have noise that is different at different locations in space. In addition, cellular molecules typically are synthesised in a series of steps where each intermediate product or state leading to the final product will also leave its contribution to the overall fluctuations. To account for the problem just described, the WKB method can be exploited to give approximate solutions in leading order up to a constant pre-exponential factor of the escape time^{24–27}. We will derive another approximate method to account for the molecule dependent noise by taking the classic approach of integrating the Fokker-Planck equation²⁸ where the pre-exponential factor directly will be given from the kinetic parameters. We will collapse the dynamics of all chemical components into a single variable and describe the noise level of the collapsed variable for concentrations outside of stable states such that the escape time can be integrated through a pseudo-potential. The calculations improves the accuracy of what can be obtained from a constant noise approximation and in addition accounts for intermediate product noise and non-linear degradation as a result of dimer formation.

Results

The necessary criteria for generating two stable phenotypic states from positive transcriptional feedback is that the feedback is sufficiently non-linear, which can be quantified by the Hill coefficient^{29–31}. We will restrict the calculations to feedback generating a Hill coefficient of two, the smallest coefficient that generates bi-stability. This can be achieved when monomers bind cooperative on the promoter to two adjacent binding sites or when proteins form dimers that bind on a single binding site.

The biochemical reaction schemes are illustrated in Fig. 1. In Fig. 1a,b cooperative binding feedback and dimer binding feedback is illustrated, respectively. Protein synthesis is modeled as a two-step process with mRNA as intermediate species. The gene can be either in its basal or activated state, active whenever the promoter is bound with protein, with mRNA synthesis rates of $k_{m,0}$ in the un-bound state and $k_{m,0} + k_{m,a}$ in the active state. The concentrations of; mRNA m , proteins p and protein dimers p_2 decay with rates γ_m , γ_p and γ_{p_2} , respectively. The autocatalytic positive feedback loop is then a result by the protein enhancing its own synthesis by binding to its promoter either in the form of dimers or by two proteins binding cooperatively.

We will derive a stochastic differential equation that describes the evolution of the feedback system through a single reaction variable x ,

$$dx(t) = \Phi(x)dt + B(x)dW(t) \quad (1)$$

The equation contains a concentration drift term $\Phi(x)$ that is given by the (ensemble) average rate difference of synthesis and decay and a noise term $B(x)$, originating from the small perturbations of birth and death of molecules. The drift term is calculated in the limit of large molecules whereas the noise term is given by the relative perturbation at finite molecule numbers. When we have expressed $\Phi(x)$ and $B(x)$ we can through a pseudo potential, $\Psi(x) = -\int \frac{\Phi(x)}{\theta(x)} dx$, where $\theta(x) = \frac{B^2(x)}{2}$, calculate the average escape time as

$$\Gamma_{LH,HL} = \frac{2\pi e^{\Delta\Psi_{LH,HL}}}{\theta(x_{L,H}) \sqrt{\left(\frac{d^2}{dx^2}(\Psi + \ln(\theta))\right)_{x_{L,H}} \left|\frac{d^2}{dx^2}(\Psi)\right|_{x_M}}}, \tag{2}$$

where $\Delta\Psi_{LH,HL} = \Psi(x_M) - \Psi(x_{L,H})$. The indices L and H denotes Low and High stable concentration which are separated by the potential peak at the Middle unstable point, M . The derivation is given in the Methods Section. Thus we need to choose our variable x , calculate $\Phi(x)$ and $\theta(x)$ and then integrate the pseudo potential $\Psi(x)$ and insert into equation 2.

Concentration drift and stability. The concentrations of the molecules illustrated in Fig. 1 will in the large molecule limit (for dimer binding feedback) evolve according to the following set of ordinary differential equations (ODE)

$$\begin{aligned} \frac{dm}{dt} &= k_{m,0} + k_{m,a} \left(\frac{p_2}{K_2 + p_2} \right) - \gamma_m m \\ \frac{dp}{dt} &= k_p m + 2k_{-1} p_2 - 2k_1 p^2 - \gamma_p p \\ \frac{dp_2}{dt} &= k_1 p^2 - k_{-1} p_2 - \gamma_{p_2} p_2 \end{aligned} \tag{3}$$

Typically, the protein life-time is large compared to the mRNA life-time and the equilibration of monomer and dimer concentrations. Under such conditions we can formulate an equation for the concentration drift $\Phi(\rho_T)$ in the total scaled protein concentration, $\rho_T = \rho + 2\rho_2$ where $\rho_i = \frac{p_i}{K}$ as

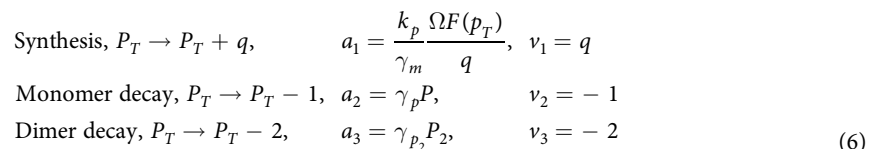
$$\begin{aligned} \Phi(\rho_T) &\equiv \frac{d\rho_T}{d\tau} = S_0 + S_a \left(\frac{\rho^2}{1 + \rho^2} \right) - (\rho + \alpha\rho^2), \\ \rho &= \left(\sqrt{\frac{K_1}{2K} \left(\rho_T + \frac{K_1}{8K} \right)} - \frac{K_1}{4K} \right) \end{aligned} \tag{4}$$

giving us the effective synthesis rate constants $S_0 = \frac{k_{m,0}}{K} \frac{k_p}{\gamma_m \gamma_p}$, $S_a = \frac{k_{m,a}}{K} \frac{k_p}{\gamma_m \gamma_p}$ and the scaled dimer decay rate constant $\alpha = 2 \frac{\gamma_{p_2} K}{\gamma_p K_1}$ and the dissociation constants $K^2 = K_1 K_2$. Time is normalized by the protein life-time as $\tau = \gamma_p t$. The equations describing cooperative binding feedback, illustrated in Fig. 1a, can be obtained by letting $K_1 \rightarrow \infty$ making $\alpha = 0$. Moreover, since $\rho_T = \rho + 2\rho_2$ and $\rho^2 = \rho_2 K_1 / K$ we can formulate equation 4 closed in terms of any of the concentrations ρ , ρ_2 or ρ_T . For the rate parameters where the system displays bi-stability there are always two stable stationary solutions separated by an unstable one and the region of bi-stability is given by the interval $(\rho_{*,1}, \rho_{*,2})$ where both endpoints are solutions of $\Phi(\rho_*) = 0$ and $\Phi'(\rho_*) = 0$. The region of bi-stability is given by

$$\begin{aligned} S_a(\rho_*) &= \frac{(1 + 2\alpha\rho_*)(1 + \rho_*^2)^2}{2\rho_*} \\ S_0(\rho_*) &= \rho_* + \alpha\rho_*^2 - \frac{\rho_*}{2}(1 + 2\alpha\rho_*)(1 + \rho_*^2), \end{aligned} \tag{5}$$

where ρ_* are the monomer concentrations separating the bi-stable region. The region of bi-stability is displayed in Fig. 2a. In the limit of $\alpha \rightarrow 0$ the dimer binding feedback give the same region of bi-stability as cooperative binding. In the limit of having large (scaled) activated synthesis rate, S_a , the limit for having bi-stability is given by $S_0 = 1/4S_a$. Thus a large feedback activation need to be accompanied by a low basal synthesis rate for bi-stability to occur. In Fig. 2b we display the potential function, given by the primitive function of the concentration drift $\Phi(\rho_T)$, at the three marked points in Fig. 2a.

Intrinsic noise. We will now derive an expression for the intrinsic noise of the total concentration of proteins, $B(p_T)$, assuming that mRNA turnover and protein dimerization equilibration are both fast compared to protein turnover. One species, total protein copy number, P_T , and three reactions,



The function $F(p_T)$ in the propensity of the synthesis is given by the two first terms in the first row of equation 3, i.e the synthesis rate of proteins. However, we have scaled the propensity of synthesis with the average protein synthesised per mRNA and the stoichiometric correction q . The idea is that the average synthesis rate, which is the propensity times the stoichiometry, is not dependent on q and that q gives a measure of the burst rate of the

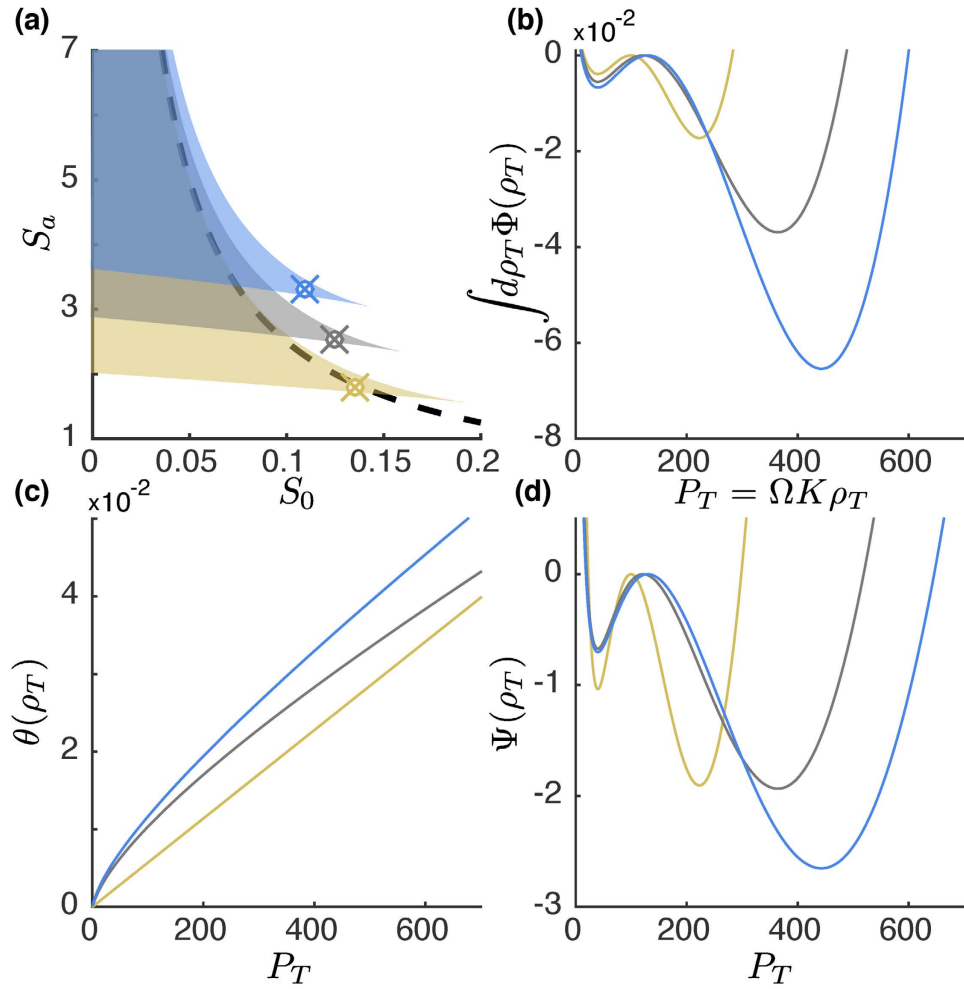


Figure 2. The potential and noise properties as a function of the total protein number $P_T = \Omega K \rho_T$. (a) The region of bi-stability for cooperative binding feedback (yellow) and dimer binding feedback with the decay rate parameter $\alpha = 0.5$ (grey) and $\alpha = 1$ (blue). The dashed line is the asymptote $S_0 = 1/4S_a$ gives the upper limit of the basal synthesis rate S_0 generating bi-stability in the limit of large activation synthesis S_a . (b) The potential, which is given by the integral of the concentration drift, of the three points marked in (a) as a function of the number of proteins. In (c) the noise $\theta(\rho_T)$ needed for calculating the effective quasi-potential $\Psi(x) = -\int \frac{\Phi(x)}{\theta(x)} dx$, de-pictured in (d).

protein synthesis in the full reaction system. Increasing q implies that more proteins are produced per synthesis event but that the synthesis events appear less frequently. We can calculate the variance C of the total protein concentration using the fluctuation dissipation relation (see Methods section), $C = -D/2A$, where

$$A = \left[\frac{k_p}{\gamma_m} \frac{dF(p_T)}{dp_T} - \gamma_p \left(1 + 2\beta \frac{\partial p_2}{\partial p_T} \right) \right]_{(p, p_2, p_T)_s} \tag{7}$$

letting $\beta = \frac{\gamma_{p_2}}{\gamma_p} - 1$ and where the diffusion component is

$$D = \left[\frac{\gamma_p}{\Omega} (2(p_T + (3\beta + 1)p_2) + (q - 1)(p_T + 2\beta p_2)) \right]_{(p, p_2, p_T)_s} \tag{8}$$

By matching the variance of the total protein concentration for the full system, calculated in the same way as above using fluctuation dissipation relation, we obtain the effective stoichiometry for the reduced system 6 to be $q = 1 + 2b$ when mRNA decay is assumed to be rapid compared to protein decay (calculations of q is presented in the supplementary information). The parameter $b = \frac{k_p}{\gamma_m}$ is the number of proteins translated per mRNA. We then obtain the noise component for the dimer binding feedback by inserting the calculated q in equation 8,

$$\theta(\rho_T) = \frac{D}{2K^2\gamma_p} = \frac{1}{K\Omega} \left((1+b)\rho_T + (1+3\beta+2\beta b) \frac{K}{K_1} \left(\sqrt{\frac{K_1}{2K} \left(\rho_T + \frac{K_1}{8K} \right)} - \frac{K_1}{4K} \right)^2 \right) \quad (9)$$

where we have re-written ρ_2 in terms of ρ_T . For cooperative binding feedback we get $\theta(\rho_T) = \frac{1}{K\Omega}(1+b)\rho_T$. Note that $\theta(\rho_T)$ is only valid at the two stable fixed points $\rho_T = \rho_L$ and $\rho_T = \rho_H$, but we now assume that the dependence on ρ_T can be approximated to hold for all concentrations.

Escaping the stable states. Now we know both $\Phi(\rho_T)$ and $\theta(\rho_T)$ are ready to calculate the escape time of the phenotypic states by integrating the quasi-potential $\Psi(\rho_T)$. The integration of Ψ is presented in the Supplementary Information and for brevity we just state the results here. For dimer binding feedback we obtain

$$\begin{aligned} \Psi(\tilde{\rho}_T) = & \frac{K\Omega}{1+b} \left(\frac{1}{1 + \frac{1}{2(1+b)}} \tilde{\rho}_T + \ln \left(\frac{(\sqrt{\tilde{\rho}_T} + \delta_1)^{n_1} (\sqrt{\tilde{\rho}_T} + \delta_2)^{n_2}}{\left(\left(\sqrt{\tilde{\rho}_T} - \sqrt{\frac{K_1}{8K}} \right)^2 + \frac{1}{4} \left(\frac{K_1}{K} + \frac{8K}{K_1} \right) \right)^{n_3}} \right) \right) \\ & + n_4 \tan^{-1} \left(\frac{\sqrt{\tilde{\rho}_T} - \sqrt{\frac{K_1}{8K}}}{\frac{1}{2} \sqrt{\frac{K_1}{K} + \frac{8K}{K_1}}} \right) + n_5 \sqrt{\tilde{\rho}_T} \end{aligned} \quad (10)$$

where $\tilde{\rho}_T = \rho_T + \frac{K_1}{8K}$. The parameters δ_i and n_i are presented in the Supplementary Information. For cooperative binding feedback

$$\Psi(\rho_T) = \frac{K\Omega}{1+b} \left(\rho_T - S_0 \ln(\rho_T) - \frac{S_a}{2} \ln(1 + \rho_T^2) \right) \quad (11)$$

We note that the pseudo potential difference between the stable states and the unstable peak, $\Delta\Psi_{LH,HL} = \Psi(x_M) - \Psi(x_{L,H})$, scale to first order as $\Delta\Psi_{LH,HL} = \frac{P_{TM} - P_{TH,L}}{\frac{3}{2} + b}$ for dimer binding feedback $\Delta\Psi_{LH,HL} = \frac{P_{TM} - P_{TH,L}}{1+b}$ for cooperative binding feedback.

We have plotted both the noise component θ and the quasi potential Ψ in Fig. 2b,c. The escape time is obtained by insertion of $\Psi(\rho_T)$ and $\theta(\rho_T)$ in equation 2 with $x = \rho_T$. In Fig. 3a–c we display escape properties for cooperative binding feedback and in panels 3(d)–(f) for dimer binding feedback. The accuracy of the escape time is displayed in Fig. 3a,b,d,e where we as a reference also display the escape time calculated for the constant noise approximation (dashed line). The constant noise approximation is performed by approximating to the noise level at the stable concentration from which the escape is taking place, $\theta(\rho_T) = \theta(\rho_L)$ or $\theta(\rho_H)$ ³². The benefit of accurately describing the noise is seen when comparing with the constant noise approximation, which is systematically underestimating the noise levels whenever going from the low to the high stable concentration, resulting in an escape time that is longer than what is obtained from exact stochastic simulations of the reaction network. For escaping from the high stable concentrations we see the opposite behaviour, which is natural since the noise is overestimated when approximated with the level at the high stable concentration. The stochastic simulations of the escape rate is done using Gillespie's stochastic simulation algorithm³³.

Now, since the fluctuations at the stationary points occur with a time-scale much shorter than the time scale of switching from one phenotypic state to another, implying that the actual jump from one phenotypic state to the another it fast compared to the time spent in either of the phenotypic states, it is reasonable to assume that the switching process is memoryless and can be approximated as a one-step process (similar to a regular chemical reaction). Under such assumptions the time between switching events can be described by the exponential distribution, with the average rate m_i^{-1} of switching as the only parameter describing the distribution. We test this assumption by measuring the probability of escaping from the low state to the high state during the time interval $[0, t]$ from stochastic simulations (Figure 3c and 3f, diamonds and circles), that is the cumulative distribution $P(\leq t)$. The cumulative probability of the exponential distribution is given by $P(\leq t) = 1 - e^{-t/\Gamma_{LH}}$, where we have inserted our calculated mean time of switching $m_i = \Gamma_{LH}$. The small insets display the probability density function $P(t) = \Gamma_{LH}^{-1} e^{-t/\Gamma_{LH}}$.

Discussion

We have presented a method for calculating the switching kinetics between phenotypic states, formed by auto-catalytic synthesis, where the switching is driven by intrinsic fluctuations. The method relies on deriving a single reaction parameter by a separation of timescales and formulating an effective stoichiometric parameter such that both the concentration drift and fluctuations represent what is observed in the full system. Here we have applied the results on transcription factors binding to its own promoter and enhancing its own synthesis. Such feedback is the most common transcriptional feedback motif in bacteria³⁴ with the possibility of generating meta-stable states. The feedback need to be sufficiently sensitive to generate metastable states and this can be achieved either by forming multimeric complexes before binding to the promoter or/and having multiple and cooperative binding sites on the promoter. Typically, transcription factors form multimeric complexes, as dimers or tetramers,

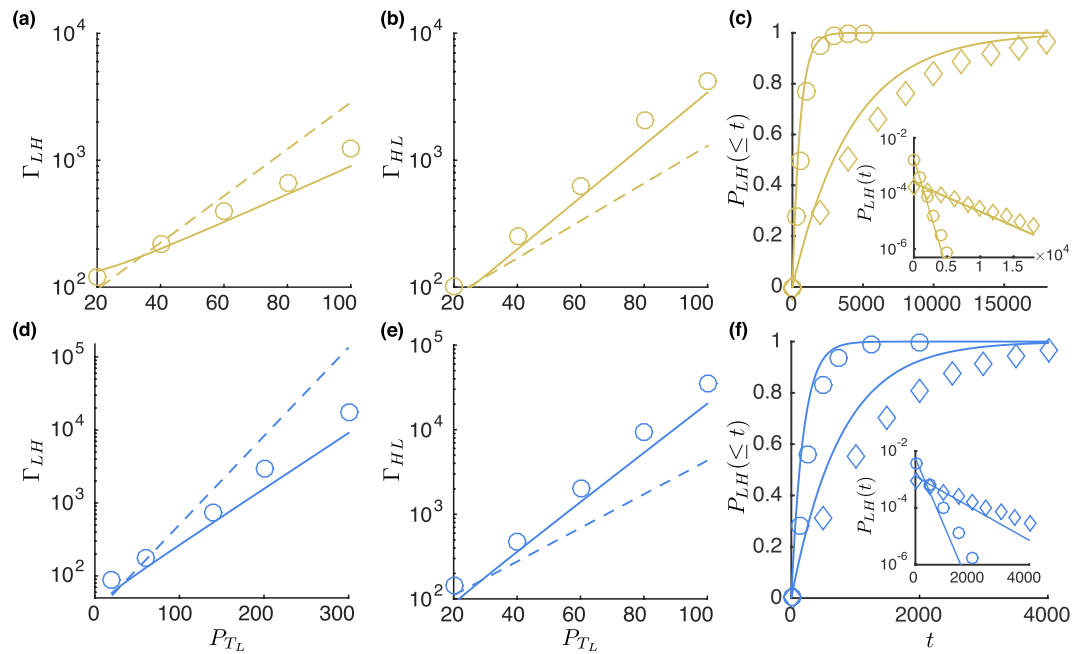


Figure 3. The accuracy of the escape time from the phenotypic low states L and high states H with increasing number of molecules. Upper panels show escape properties for cooperative feedback binding and lower panels for dimer binding feedback. The rate parameters are given from the yellow and blue points in Fig. 2a for cooperative and dimer feedback, respectively. In addition we set the monomer-to-dimer decay rate ratio to 10 for dimer binding feedback, making the lifetime of dimers ten times longer than monomers. Errors of the stochastic simulations are smaller than the symbols. The panels (a,b,d,e) show the mean escape rate, where Γ_{LH} denotes escape from Low to High state and vice versa. In the rightmost panels we display the probability of escaping the low state to the high state as a function of increasing time where the lines are the theoretical expected probability assuming an cumulative probability of the exponential distribution, $P_{LH}(\leq t) = (1 - e^{-t\Gamma_{LH}})$. The small insets show the probability density function of escaping assuming an exponential distribution, $P_{LH} = \Gamma_{LH}^{-1}e^{-t\Gamma_{LH}}$. Errors are smaller than symbols and circles are for the low state having 80 proteins and diamonds with the low state having 160 molecules.

and most likely it is not only due to improving the feedback sensitivity but also due to symmetry in the molecule's binding to DNA. The formation of a multimeric complex of the feedback molecule give that the molecule can bind in several orientations which makes the search time to the specific binding site shorter, since the binding (and re-binding after dissociation) becomes more probable³⁵. Assuming all other reactions parameters fixed, faster search kinetics of the transcription factor will increase the equilibrium binding constant. From the method in this paper, one may directly investigate such effects since both the pre-exponential factor and the exponential rate dependence is directly expressed from the kinetic parameters of the reaction networks. We have also incorporated another effect of forming multimeric complexes in the method, the effect of having non-linear decay since more than one molecule may be removed at one instant, e.g. by spontaneous decay of a multimer, or by a multimer being consumed in some biochemical pathway. Since the time scale of the stability may change orders of magnitude between feedback from dimer binding and cooperative binding with otherwise similar reaction parameters we conclude that to understand phenotypic switching we need to be careful when formulating a model of an existing reaction network.

Methods

Escape rate calculation. Assume that we have a chemical species whose concentration is described by the variable x . Assume furthermore that the species are displaying two stable concentrations x_L and x_H , a low and high stable concentration, respectively. The two stable concentrations are separated by the middle unstable concentration x_M . We want to calculate the intrinsic stability of the two stable phenotypic states x_L and x_H , that is, for how long will the system stay (on average) in a stable state before exit to the other stable state due to the intrinsic fluctuations generated from synthesis and decay of molecules. Individual trajectories $x(t)$ evolve according to the stochastic differential equation

$$dx(t) = \Phi(x)dt + B(x)dW(t) \quad (12)$$

The term $\Phi(x)$ gives the average drift in concentration and $B(x)$ the fluctuations, where $W(t)$ is a Wiener Process. Notably, both the drift and the fluctuations depend on the concentration since both the drift and fluctuations are generated from the biochemical reactions of the system (which typically depends on the concentrations). Since the system is time-homogenous, i.e. the conditional probabilities satisfying $\Pi(x', t | x, 0) = \Pi(x', 0 | x, -t)$, the

backward Kolmogorov (or Fokker-Planck) equation corresponding to the stochastic process described in equation 12 can be written

$$\frac{\partial \Pi(x', t|x, 0)}{\partial t} = \Phi(x) \frac{\partial \Pi(x', t|x, 0)}{\partial x} + \theta(x) \frac{\partial^2 \Pi(x', t|x, 0)}{\partial x^2}, \quad (13)$$

where $\theta(x) = \frac{B^2(x)}{2}$. The probability that the concentration remains in the region \mathbf{R} at time t starting from some concentration x inside \mathbf{R} is $G(x, t) = \int_{\mathbf{R}} \Pi(x', t|x, 0) dx'$. By integration of the backward equation 13 over x' we get

$$\frac{\partial G(x, t)}{\partial t} = \Phi(x) \frac{\partial G(x, t)}{\partial x} + \theta(x) \frac{\partial^2 G(x, t)}{\partial x^2} \quad (14)$$

Now, the average time of escaping \mathbf{R} starting in x is given by $\Gamma(x) = -\int_0^\infty t \frac{\partial G(x, t)}{\partial t} dt = \int_0^\infty G(x, t) dt$, is therefore obeying the equation

$$\Phi(x) \frac{d\Gamma(x)}{dx} + \theta(x) \frac{d^2\Gamma(x)}{dx^2} = G(x, \infty) - G(x, 0) = -1 \quad (15)$$

Introducing the quasi-potential $\Psi(x)$ where $\frac{d\Psi(x)}{dx} = -\frac{\Phi(x)}{\theta(x)}$, gives the solution by the integrating factor $e^{-\Psi}$. The escape time from the low state to the high state, Γ_{LH} , is obtained by integrating twice, from 0 to x' with boundary condition $\left. \frac{d\Gamma(x)}{dx} \right|_{x=0} = 0$, and from the low stable concentration, x_L , to the high stable concentration, x_H , with boundary condition $\Gamma(x_H) = 0$ gives $\Gamma_{LH} = \int_{x_L}^{x_H} e^{\Psi} dx' \int_0^{x'} e^{-\Psi - \ln(\theta)} dx''$. In the same manner we obtain the escape time from the high state to the low state, Γ_{HL} . Integrating from ∞ to x' with boundary condition $\left. \frac{d\Gamma(x)}{dx} \right|_{x \rightarrow \infty} = 0$, and from x_H to x_L with boundary condition $\Gamma(x_L) = 0$ gives $\Gamma_{HL} = \int_{x_L}^{x_H} e^{\Psi} dx' \int_{x'}^{\infty} e^{-\Psi - \ln(\theta)} dx''$. In these equations for the escape time, the external integrands get their main contribution near the middle unstable stationary point, x_M , and the internal integrands get their main contribution near the low stable stationary point x_L for Γ_{LH} , and near high stable stationary point x_H for Γ_{HL} . Therefore, a parabolic approximation of the integrand functions around the dominant points gives the average escape time as

$$\Gamma_{LH,HL} = \frac{2\pi e^{\Delta\Psi_{LH,HL}}}{\theta(x_{L,H}) \sqrt{\left(\frac{d^2}{dx^2} (\Psi + \ln(\theta)) \right)_{x_{L,H}} \left| \frac{d^2}{dx^2} (\Psi)_{x_M} \right|}}, \quad (16)$$

where $\Delta\Psi_{LH,HL} = \Psi(x_M) - \Psi(x_{L,H})$.

System size scaling. Since we are interested in analysing the intrinsic, finite-molecule, stochastic properties of the bistable system it is convenient if the reaction network can be scaled from a system of few molecules to a large number of molecules without affecting the bi-stability. This can be achieved by introducing a system size parameter Ω ³⁶. In this setting we can view the scaled variables as concentrations, e.g. the protein monomer concentration $p = P/\Omega$, such that all fixed point concentrations $x = X/\Omega$ are invariant under a change in Ω . Even though changing Ω can be viewed as a change in volume it should primarily be viewed as a scaling parameter that gives us additional freedom when choosing the number of molecules in the stable states.

Fluctuation dissipation relation. For a multivariate Ornstein-Uhlenbeck process defined by the stochastic differential equation $dx(t) = -Ax(t)dt + BdW(t)$, where \mathbf{A} and \mathbf{B} are constant matrices, the covariance matrix \mathbf{C} at stationary conditions satisfies a fluctuation-dissipation relation

$$\mathbf{AC} + \mathbf{CA}^T - \mathbf{BB}^T = 0 \quad (17)$$

If \mathbf{x} is sufficiently close to a stationary point \mathbf{x}_s we can for a more general system $d\mathbf{x}(t) = -\mathbf{f}(\mathbf{x}(t))dt + \mathbf{B}dW(t)$ linearise the drift $\mathbf{f}(\mathbf{x}(t))$. The matrix \mathbf{A} is then the Jacobian matrix evaluated at the stationary point \mathbf{x}_s , $\mathbf{A}_{ij} = \frac{\partial f_i}{\partial x_{sj}}$.

For a reaction network, the diffusion matrix is given by $\mathbf{D}_{ij} = (\mathbf{BB}^T)_{ij} = \frac{1}{\Omega} \sum_k a_k v_{ki} v_{kj}$, where a_k is the propensity of reaction k and v_{ki} the stoichiometry of species i in reaction k , evaluated at the stationary concentration \mathbf{x}_s ³⁶. Once \mathbf{A} and \mathbf{D} are calculated we can compute the covariance matrix \mathbf{C} through the fluctuation dissipation theorem 17. For a single variable process the variance is $C = -\frac{A}{2D}$.

References

1. Taniguchi, Y. *et al.* Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329** (2010).
2. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science (New York, N.Y.)* **297**, 1183–6 (2002).
3. Paulsson, J. Summing up the noise in gene networks. *Nature* **427**, 415–8 (2004).
4. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
5. Grönlund, A., Lötstedt, P. & Elf, J. Costs and constraints from time-delayed feedback in small gene regulatory motifs. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 8171–8176 (2010).
6. Grönlund, A., Lötstedt, P. & Elf, J. Delay-induced anomalous fluctuations in intracellular regulation. *Nature communications* **2**, 419 (2011).

7. Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516–9 (2008).
8. Grönlund, A., Lötstedt, P. & Elf, J. Transcription factor binding kinetics constrain noise suppression via negative feedback. *Nature communications* **4**, 1864 (2013).
9. Bialek, W. & Setayeshgar, S. Physical limits to biochemical signaling. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10040–10045 (2005).
10. Pulkkinen, O. & Metzler, R. Distance Matters: The Impact of Gene Proximity in Bacterial Gene Regulation. *Physical Review Letters* **110**, 198101 (2013).
11. Hammar, P. *et al.* The lac repressor displays facilitated diffusion in living cells. *Science (New York, N.Y.)* **336**, 1595–8 (2012).
12. Hammar, P. *et al.* Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. *Nature genetics* **46**, 405–8 (2014).
13. Paulsson, J., Berg, O. G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7148–53 (2000).
14. Becskei, A., Séraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO journal* **20**, 2528–2535 (2001).
15. Smits, W. K., Kuipers, O. P. & Veening, J.-W. Phenotypic variation in bacteria: the role of feedback regulation. *Nature reviews. Microbiology* **4**, 259–71 (2006).
16. Xiong, W. & Ferrell, J. E. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460–465 (2003).
17. Mora, T. & Walczak, A. M. Effect of phenotypic selection on stochastic gene expression. *The Journal of Physical chemistry B* **117**, 13194–13205 (2013).
18. Aurell, E. & Sneppen, K. Epigenetics as a first exit problem. *Physical review letters* **88**, 048101 (2002).
19. Ptashne, M. A genetic switch: Phage lambda and higher organisms. *Cell and Blackwell Scientific*, Cambridge, MA (1992).
20. Kessler, D. A. & Levine, H. Large population solution of the stochastic Luria Delbruck evolution model. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11682–7 (2013).
21. Luria, S. E. & Delbrück, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **28**, 491–511 (1943).
22. Acar, M., Mettetal, J. T. & van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. *Nature genetics* **40**, 471–5 (2008).
23. Deris, J. B. *et al.* The innate growth bistability and fitness landscapes of antibiotic-resistant bacteria. *Science (New York, N.Y.)* **342**, 1237435 (2013).
24. Ge, H., Qian, H. & Xie, X. S. Stochastic phenotype transition of a single cell in an intermediate region of gene state switching. *Physical review letters* **114**, 078101 (2015).
25. Assaf, M., Roberts, E., Luthey-Schulten, Z. & Goldenfeld, N. Extrinsic noise driven phenotype switching in a self-regulating gene. *Physical review letters* **111**, 058102 (2013).
26. Assaf, M., Roberts, E. & Luthey-Schulten, Z. Determining the stability of genetic switches: explicitly accounting for mrna noise. *Physical review letters* **106**, 248102 (2011).
27. Escudero, C. & Kamenov, A. Switching rates of multi-step reactions. *arXiv preprint arXiv:0811.0090* (2008).
28. Gardiner, C. *Stochastic methods* (Springer, Berlin, 2009).
29. Hill, a. V. The Combinations of Haemoglobin with Oxygen and with Carbon Monoxide. I. *The Biochemical journal* **7**, 471–480 (1913).
30. Weiss, J. N. The Hill equation revisited: uses and misuses. *The FASEB Journal* **11**, 835–841 (1997).
31. Bintu, L. *et al.* Transcriptional regulation by the numbers: models. *Current opinion in genetics & development* **15**, 116–24 (2005).
32. Lestas, I., Paulsson, J., Ross, N. E. & Vinnicombe, G. On the analysis of noise in gene regulatory networks. *2007 46th IEEE Conference on Decision and Control* 621–626 (2007).
33. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics* **22**, 403–434 (1976).
34. Salgado, H. *et al.* RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic acids research* **41**, D203–13 (2013).
35. Hammar, P. *et al.* The lac repressor displays facilitated diffusion in living cells. *Science* **336**, 1595–1598 (2012).
36. Van Kampen, N. G. *Stochastic processes in physics and chemistry*, vol. 1 (Elsevier, 1992).

Author Contributions

A.-B.B. and A.G. jointly conceived the study and wrote the paper.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Besya, A.-B. and Grönlund, A. Intrinsic phenotypic stability of a bi-stable auto regulatory gene. *Sci. Rep.* **6**, 22951; doi: 10.1038/srep22951 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>