

Chance and memory

Theodore J. Perkins

Ottawa Hospital Research Institute
Ottawa, Ontario, Canada

Andrea Y. Weiße

Centre for Systems Biology at Edinburgh
University of Edinburgh, Edinburgh, Scotland

Peter S. Swain

Centre for Systems Biology at Edinburgh
University of Edinburgh, Edinburgh, Scotland

Introduction

Memory and chance are two properties or, more evocatively, two strategies that cells have evolved. Cellular memory is beneficial for broadly the same reasons that memory is beneficial to humans. How well could we function if we forgot the contents of the refrigerator every time the door was closed, or forgot where the children were, or who our spouse was? The inability to form new memories, called anterograde amnesia in humans, is tremendously debilitating, as seen in the famous psychiatric patient H.M. [1] and the popular movie “Memento” (Christopher Nolan, 2000).

Cells have limited views of their surroundings. They are able to sense various chemical and physical properties of their immediate environment, but distal information can come only through stochastic processes, such as diffusion of chemicals. Cells lack the ability to directly sense distal parts of their environment, as humans can through vision and hearing. Thus, we might expect that cells have adopted memory of the past as a way of accumulating information about their environment and so gain an advantage in their behaviour. In a simple application, memory can be used to track how environmental signals change over time. For example, in the chemotaxis network of *E. coli*, cells use memory of previous measurements of chemical concentrations to estimate the time derivative of the chemical concentration and so to infer whether the organism is swimming up or down a chemical gradient [2]. Conversely, memory can be useful in enabling persistent behaviour despite changing environments. For example, differentiated cells in a multicellular organism must remain differentiated and not redifferentiate despite existing under a variety of conditions. If such cells become damaged or infected, they may undergo apoptosis, but we do not observe neurons becoming liver cells, or liver cells becoming muscle cells, and so on.

Different mechanisms for achieving cellular memory operate at different time scales. For example, the chemotaxis machinery of *E. coli* can change response in seconds. Gene regulatory systems controlling sugar catabolism, such as the *lac* operon in *E. coli* [3] or the galactose network in *Saccharomyces cerevisiae* [4], take minutes or hours to switch state. Terminally differentiated cells can retain their state for years – essentially permanently, although scientists are learning how to reset their memories and return cells to a pluripotent status [5]. Of course, that cells

have complex internal chemical and structural states implies a certain amount of “inertia”. A cell cannot change its behaviour instantly, so in a weak sense, all cellular systems display some degree of memory. We, however, will focus primarily on forms of cellular memory that operate on longer time scales than the physical limits imposed by changing cellular state. We will describe forms of memory that require active maintenance, as embodied by positive feedback. These memories are more similar to human memories in that their state may represent some specific feature of the cell’s environment, such as “high lactose”.

While the potential for memory to be advantageous is clear, the benefit of chance behaviour is perhaps less obvious. Some human activities provide clear examples, especially in competitive situations. For example, if a tennis player always served to the same spot on the court, the opponent would have a much easier time of returning the ball. Predictable behaviours give competitors an opportunity to adapt and gain an advantage, if they can detect the pattern of the behaviour. Thus, intentional randomization of one’s own behaviour can be beneficial. Unicellular organisms are virtually always in a competitive situation – if not from other species, then from others of their own kind seeking out limited resources such as nutrients or space. One well-known example is the phenomenon of bacterial persistence, where some members of a genetically identical population of bacteria spontaneously and stochastically adopt a phenotype that grows slowly but is more resistant to environmental insults such as antibiotics [6]. The cells spend some time in the persister state and then spontaneously and stochastically switch back to normal, more rapid growth. This strategy allows most of the population to grow rapidly in ideal conditions, but prevents the extinction of the population if an insult should occur. Its stochastic aspect makes it challenging to predict which cells will be persisters at any particular time. Thus the population as a whole, in its competition with the doctor prescribing the antibiotics, increases its chances of success.

Like memory, chance behaviour can have different time scales and specificities. For example, fluctuations in numbers of ribosomes can affect the expression of virtually all genes and persist for hours or longer in microbes. Fluctuations in the number of transcripts of a particular gene may last for only seconds or minutes and may directly influence only the expression of that gene.

The phenomena of memory and chance can be at odds. Memory implies that decision or behaviours are contingent on the past. Chance implies that decisions are randomized, implying a lesser dependence on the past or present. If chance plays a role in the memory mechanism itself, then it tends to degrade the memory, resulting in a neutral or random state that is independent of the past. Yet, both memory and chance are important cellular strategies. Intriguingly, they are often found in the same systems. Many of the model systems used as examples of cellular memory, such as gene regulatory networks involved in catabolism, determination of cell type, and chemotaxis, have also been used in studies of cellular stochasticity. In this chapter, we will discuss the dual and often related phenomena of cellular memory and chance. We begin by reviewing stochastic effects in gene expression and some of the basic mathematical and computational tools used for modelling and analyzing stochastic biochemical networks. We then describe various model cellular systems that have been studied from these perspectives.

Chance, or stochasticity

Life at the cellular level is inevitably stochastic [7, 8]. Biochemistry drives all of life’s processes, and reactants come together through diffusion with their motion being driven by rapid and

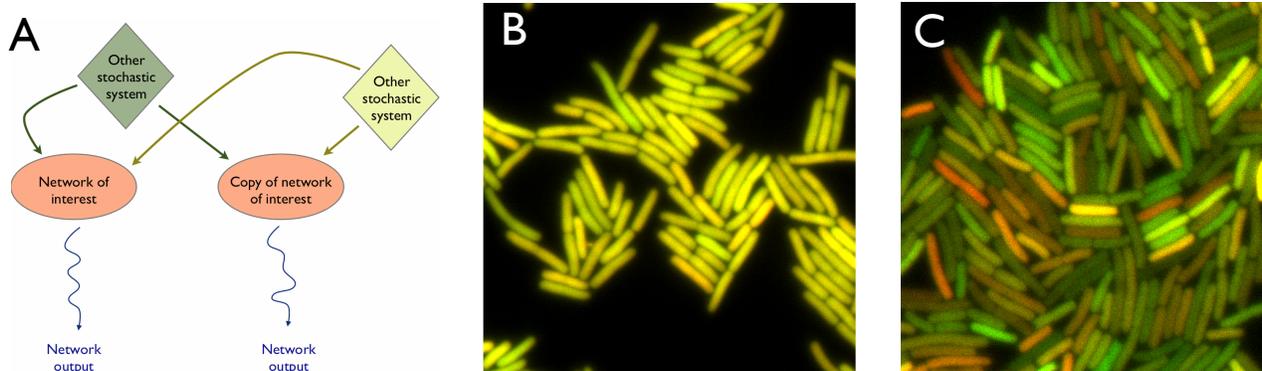


Figure 1: Intrinsic and extrinsic fluctuations. **A** Extrinsic fluctuations are generated by interactions of the system of interest with other stochastic systems. They affect the system and a copy of the system identically. Intrinsic fluctuations are inherent to each copy of the system and cause the behaviour of each copy to be different. **B** Strains of *Escherichia coli* expressing two distinguishable alleles of Green Fluorescent Protein controlled by identical promoters. Intrinsic noise is given by the variation in colour (differential expression of each allele) over the population and extrinsic noise by the correlation between colours. **C** The same strains under conditions of low expression. Intrinsic noise increases. From Elowitz *et al.* (2002).

frequent collisions with other molecules. Once together, these same collisions alter the reactants' internal energies and so their propensities to react. Both effects cause individual reaction events to occur randomly and the overall process to be stochastic. Such stochasticity need not always be significant. Intuitively, we expect the random timings of individual reactions to matter only when typical numbers of molecules are low, because then each individual event, which at most changes the numbers of molecules by one or two, generates a substantial change. Low numbers are not uncommon intracellularly: gene copy numbers are usually one or two, and transcription factors, at least for microbes, frequently number in the tens [9, 10].

Intrinsic and extrinsic stochasticity

Two types of stochasticity affect any biochemical system [11, 12]. Intuitively, intrinsic stochasticity is stochasticity that arises inherently in the dynamics of the system from the occurrence of biochemical reactions. Extrinsic stochasticity is additional variation that is generated because the system interacts with other stochastic systems in the cell or in the extracellular environment (Fig. 1A).

In principle, intrinsic and extrinsic stochasticity can be measured by creating a copy of the system of interest in the same cellular environment as the original network (Fig. 1A) [12]. We can define intrinsic and extrinsic variables with fluctuations in these variables together generating intrinsic and extrinsic stochasticity [11]. The intrinsic variables will typically specify the copy numbers of the molecular components of the system. For the expression of a particular gene, the level of occupancy of the promoter by transcription factors, the numbers of mRNA molecules, and the number of proteins are all intrinsic variables. Imagining a second copy of the system – an identical gene and promoter elsewhere in the genome – then the instantaneous values of the intrinsic variables of this copy of the system will frequently differ from those of the original

system. At any point in time, for example, the number of mRNAs transcribed from the first copy of the gene will usually be different from the number of mRNAs transcribed from the second copy. Extrinsic variables, however, describe processes that equally affect each copy. Their values are therefore the same for each copy. For example, the number of cytosolic RNA polymerases is an extrinsic variable because the rate of gene expression from both copies of the gene will increase if the number of cytosolic RNA polymerases increases and decrease if the number of cytosolic RNA polymerases decreases.

Stochasticity is quantified by measuring an intrinsic variable for both copies of the system. For our example of gene expression, the number of proteins is typically measured by using fluorescent proteins as markers [13, 12, 14, 15] (other techniques are described in the chapter by B. Munsky). Imaging a population of cells then allows estimation of the distribution of protein levels, typically at steady-state. Fluctuations of the intrinsic variable will have both intrinsic and extrinsic sources *in vivo*. We will use the term “noise” to mean an empirical measure of stochasticity – for example, the coefficient of variation (the standard deviation divided by the mean) of a stochastic variable.

Intrinsic noise is defined as a measure of the difference between the value of an intrinsic variable for one copy of the system and its counterpart in the second copy. For gene expression, the intrinsic noise is typically the mean absolute difference (suitably normalized) between the number of proteins expressed from one copy of the gene, which we will denote I_1 , and the number of proteins expressed from the other copy [12, 16], denoted I_2 :

$$\eta_{\text{int}}^2 = \frac{\langle (I_1 - I_2)^2 \rangle}{2\langle I \rangle^2} \quad (1)$$

where we use angled brackets to denote an expectation over all stochastic variables, and $\langle I_1 \rangle = \langle I_2 \rangle = \langle I \rangle$. Such a definition supports the intuition that intrinsic fluctuations cause variation in one copy of the system to be uncorrelated with variation in the other copy. Extrinsic noise is defined as the correlation coefficient of the intrinsic variable between the two copies:

$$\eta_{\text{ext}}^2 = \frac{\langle I_1 I_2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}. \quad (2)$$

Extrinsic fluctuations equally affect both copies of the system and consequently cause correlations between variation in one copy and variation in the other. Intuitively, the intrinsic and extrinsic noise should be related to the coefficient of variation of the intrinsic variable of the original system of interest. This so-called total noise is given by the square root of the sum of the squares of the intrinsic and the extrinsic noise [11]:

$$\eta_{\text{tot}}^2 = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2} = \eta_{\text{int}}^2 + \eta_{\text{ext}}^2 \quad (3)$$

which can be verified from adding Eqs. 1 and 2.

Such two-colour measurements of stochasticity have been applied to bacteria and yeast where gene expression has been characterized by using two copies of a promoter integrated in the genome with each copy driving a distinguishable allele of Green Fluorescent Protein [12, 15] (Fig. 1B and C). Both intrinsic and extrinsic noise can be substantial giving, for example, a total noise of around 0.4: the standard deviation of protein numbers is 40% of the mean. Extrinsic noise is usually higher than intrinsic noise. There are some experimental caveats: both copies of

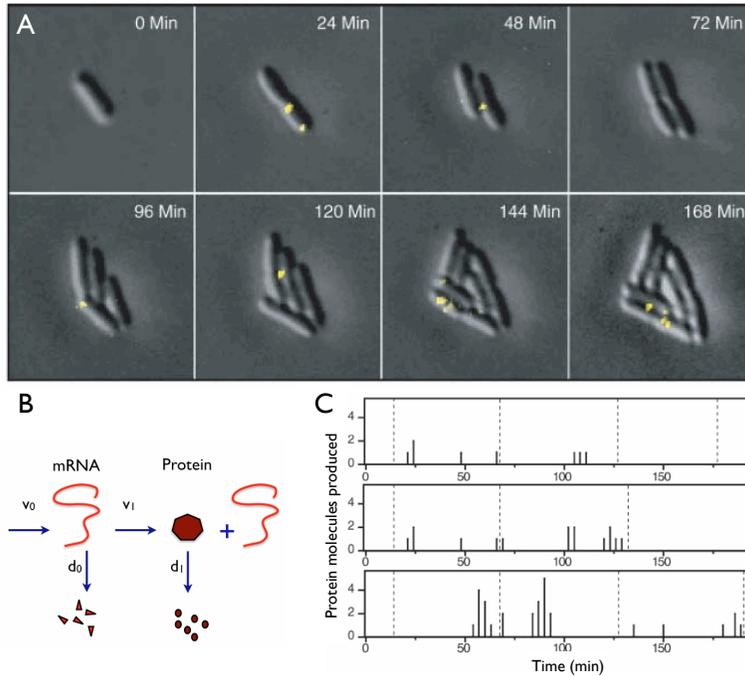


Figure 2: Stochasticity is generated by bursts of translation. **A** Time-series showing occasional expression from a repressed gene in *E. coli*, which results in the synthesis of a single mRNA, but several proteins. Proteins are membrane proteins tagged with Yellow Fluorescent Protein (yellow dots). From Yu *et al.* (2006). **B** The two-stage model of constitutive gene expression. Transcription and translation are included. **C** Quantitation of the data from **A** showing exponentially distributed bursts of translated protein. From Yu *et al.* (2006).

the system should be placed “equally” in the genome so that the probabilities of transcription and replication are equal. This equality is perhaps best met by placing the two genes adjacent to each other [12]. Although conceptually there are no difficulties, practically problems arise with feedback. If the protein synthesized in one system can influence its own expression, the same protein will also influence expression in a second copy of the system. The two copies of the system have then lost the (conditional) independence they require to be two simultaneous measurements of the same stochastic process.

Understanding stochasticity

Although measurements of fluorescence in populations of cells allow us to quantify the magnitude of noise, time-series measurements, where protein or mRNA levels are followed in single cells over time, best determine the mechanisms that generate noise. With the adoption of microfluidics by biologists [17, 18], such experiments are becoming more common, but, at least for bacteria, simpler approaches have also been insightful.

Translational bursts

Following over time the expression of a repressed gene in *E. coli*, where occasionally and stochastically repression was spontaneously lifted, revealed that translation of mRNA can occur in bursts

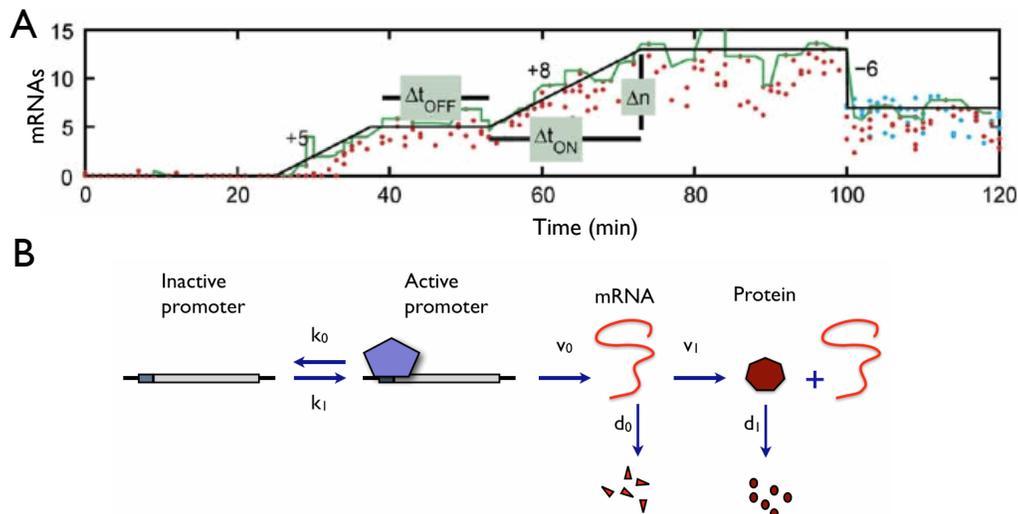


Figure 3: Stochasticity arises from transcriptional bursting. **A** Time-lapse data of the number of mRNAs from a phage lambda promoter in single cells of *E. coli*. Red dots are the number of mRNAs in one cell; the green line is a smoothed version of the red dots; the blue dots are the number of mRNAs in the sister cell after division; and the black line is a piece-wise linear fit. From Golding *et al.* (2005). **B** The three-stage model of gene expression with activation of the promoter, transcription, and translation and the potential for both transcriptional and translational bursting. The rate k_0 is inversely related to the mean Δt_{ON} in **A** and the rate k_1 is inversely related to the mean Δt_{OFF} .

[19]. Using a *lac* promoter and a fusion of Yellow Fluorescent Protein to a membrane protein so that slow diffusion in the membrane allowed single-molecule sensitivity, Yu *et al.* showed that an individual mRNA is translated in a series of exponentially distributed bursts (Fig. 2A) [19]. Such bursting was expected because of the competition between ribosomes and the RNA degradation machinery for the 5' end of the mRNA [20]. If a ribosome binds first to the mRNA, then translation occurs, and the mRNA is protected from degradation because the ribosome physically obstructs binding of the degradation machinery; if the degradation machinery binds first, then further ribosomes can no longer bind and the mRNA is degraded from its 5' end. Let v_1 be the probability of translation per unit time of a particular mRNA and d_0 be its probability of degradation (Fig. 2B). Then the probability that a ribosome binds rather than the degradation machinery is $\frac{v_1}{v_1+d_0}$ or $\frac{b}{1+b}$, where $b = v_1/d_0$ is the typical number of proteins translated from one mRNA during the mRNA's lifetime – the so-called burst size [21]. The probability of r proteins being translated from a particular mRNA, P_r , requires that ribosomes win the competition with the mRNA degradation machinery r times and lose once [20]:

$$P_r = \left(\frac{b}{1+b} \right)^r \left(1 - \frac{b}{1+b} \right). \quad (4)$$

The number of proteins translated from one mRNA molecule thus has a geometric distribution, which if $b < 1$ can be approximated by an exponential distribution with parameter b . Such exponential distributions of bursts of translation (Fig. 2C) were measured by Yu *et al.* [19].

If the mRNA lifetime is substantially shorter than the protein lifetime, then all the proteins

translated from any particular mRNA appear “at once” if viewed on the timescale associated with typical changes in the number of proteins: translational bursting occurs [22]. This separation of timescales is common, being present, for example, in around 50% of genes in budding yeast [22]. It allows for an approximate formula for the distribution of proteins [23], which if transcription occurs with a constant probability per unit time is a negative binomial distribution (Fig. 2B):

$$P_n = \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)} \left(\frac{b}{1+b}\right)^n \left(1 - \frac{b}{1+b}\right)^a \quad (5)$$

with a being the number of mRNAs synthesized per protein lifetime (v_0/d_1). This distribution is generated if translation from each mRNA is independent and if translation occurs in geometrically distributed translational bursts [22]. Such bursting allows the master equation for the two-dimensional system of mRNA and protein to be reduced to a one-dimensional equation just describing proteins. This equation is, however, more complex than a typical master equation for a one-step process because it contains a term that is effectively a convolution between the geometric distribution of bursts and the distribution for the number of proteins.

Transcriptional bursts

Transcription can also occur in bursts. By using the MS2 RNA-binding protein tagged with a fluorescent protein and by inserting tens of binding sites for MS2 in the mRNA of a gene of interest, Golding *et al.* were able to follow individual mRNAs in *E. coli* over time [24]. They observed that these tagged mRNAs are not synthesized with a constant probability per unit time even for a constitutive promoter, but that synthesis is consistent with a random telegraph process where the promoter stochastically transitions between an “on” state, capable of transcription, and an “off” state, incapable of transcription (Fig. 3A). These “on” and “off” states of the promoter could reflect binding of RNA polymerase or a transcription factor, changes in the structure of chromatin in higher organisms, or even transcriptional “traffic jams” caused by multiple RNA polymerases transcribing simultaneously [24, 25]. Whatever its cause, transcriptional bursting has been observed widely, in budding yeast [14, 15], slime molds [26], and human cells [27], as well as in bacteria. This broad applicability has led to what might be called the standard model of gene expression, which includes stochastic transitions of the promoter and translational bursting if the mRNA lifetime is substantially smaller than the protein lifetime (Fig. 3B). An exact expression for the distribution of mRNAs [27] and an approximate expression for the distribution of proteins are known [22], although both include only intrinsic but not extrinsic fluctuations.

Simulation methods

We may gain intuition about the behaviour of more complex networks using stochastic simulation. For intrinsic fluctuations or for networks where all the potential chemical reactions are known, Gillespie’s algorithm [28] is most commonly used. At its simplest, this algorithm involves rolling the equivalent of two dice for each reaction that occurs (further details are given in the chapter by B. Munsky). One die determines which reaction out of the potential reactions happens, with the propensity for each reaction being determined by the probability per unit time of occurrence of a single reaction times the number of ways that reaction can occur given the numbers of molecules of reactants present. The second die determines the time when this reaction occurs by generating a sample from an exponential distribution. Usually extrinsic

fluctuations enter into simulations because kinetic rate constants are often implicitly functions of the concentration of a particular type of protein. For example, degradation is commonly described as a first-order process with each molecule having a constant probability of degradation per time. Yet the rate of degradation is actually a function of the concentration of proteasomes, and if the number of proteasomes fluctuates the rate of degradation fluctuates. Such fluctuations are extrinsic and can also occur, for example, in translation if the number of ribosomes or tRNAs fluctuates or in transcription if the number of RNA polymerases or nucleotides fluctuates. Extrinsic fluctuations can be simulated by including, for example, a model for the availability of proteasomes, but such models can be complicated containing unknown parameters and, in principle, other extrinsic fluctuations. It therefore can be preferable to have an effective model of extrinsic fluctuations by letting the rate of a reaction, or potentially many rates, fluctuate directly in the Gillespie algorithm. These fluctuations can be simulated with a minor modification of the algorithm to include piece-wise linear changes in reaction rates and then by simulating the extrinsic fluctuations of a particular reaction rate ahead of time and feeding these fluctuations, or more correctly a piece-wise linear approximation of the fluctuations, into the Gillespie algorithm as it runs [16]. Such an approach also has the advantage of potentially generating extrinsic fluctuations with any desired properties. Where measured, extrinsic fluctuations have a lifetime comparable to the cell-cycle time [29] – they are “coloured” – and it is important to include this long lifetime because such fluctuations are not typically averaged away by the system and can generate counter-intuitive phenomena [16].

Bistability and stochasticity

Given that life is fundamentally biochemistry, how do biochemical systems generate memory? One mechanism is positive feedback that is sufficiently strong to create bistability [31]. At long times, a bistable system tends towards one of two stable steady-states. If the system starts with an initial condition in one particular collection of initial conditions, it will tend to one steady-state; if it starts with any other initial condition, it will tend to the other steady-state. The space of initial conditions – the space of all possible concentrations of the molecular components of the system – has two, so-called, domains of attraction, one for each stable steady-state (Fig. 4A). Sufficiently strong positive feedback can create such bistable behaviour. If the output of a system can feedback on the system and encourage more output, then intuitively we can see that an output sufficiently low only weakly activates such positive feedback, and the system remains at a low level of output. If, however, the output becomes high then the positive feedback is strongly activated and the output becomes higher still generating more feedback and more output. This “runaway” feedback causes the system to tend to a high level of output. No intermediate output level is possible: either output is high enough to start “runaway” positive feedback or it is not.

Bistable systems can have memory

Biochemical systems that are bistable have the potential for history-dependent behaviour and so for memory. Imagine a typical dose-response curve: as the input to the system increases so too does the output. A linear change in input often gives a non-linear change in output and so a sigmoidal input-output relationship. For bistable systems, the output increases with respect to the input for low levels of input (Fig. 4B). Such levels are not strong enough to

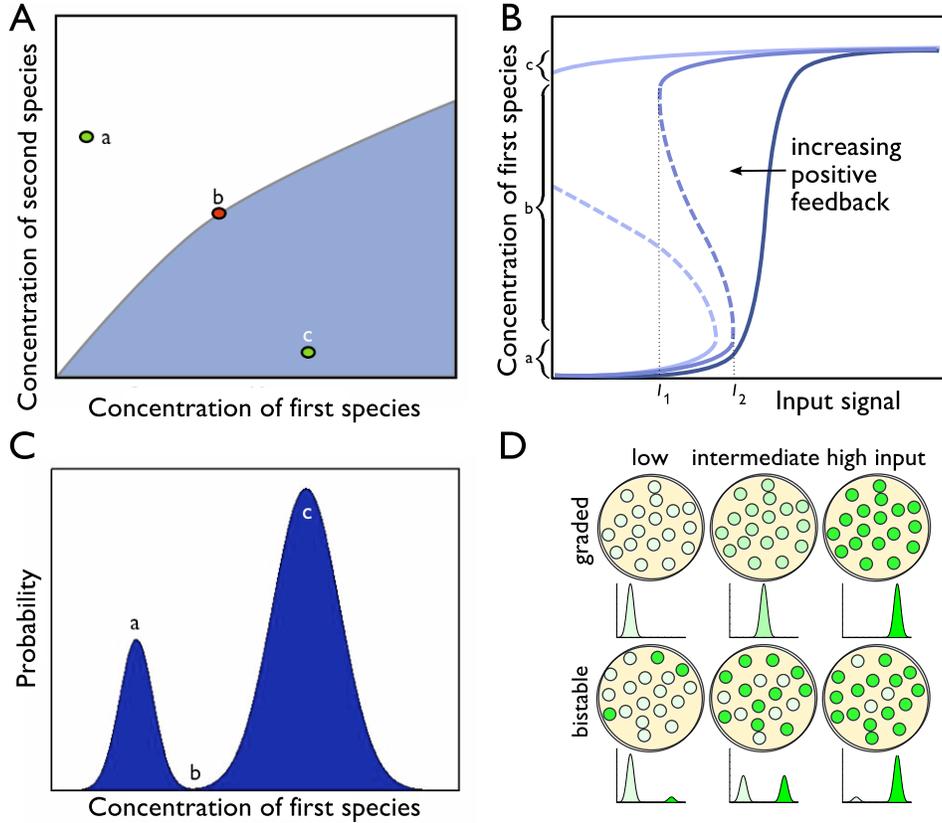


Figure 4: Bistability. **A** A steady-state has its own domain of attraction. The unshaded region is the domain of attraction of stable steady-state a ; the shaded region is the domain of attraction of stable steady-state c . The steady-state b is unstable. **B** Bifurcation diagram as a function of the level of input signal. With positive feedback an otherwise monostable system (dark line) can become bistable (medium shaded line). Two saddle-node bifurcations occur when the input is at I_1 and I_2 . The system is bistable for inputs between I_1 and I_2 and will tend to steady-state at either a or c depending on initial concentrations. It is monostable otherwise. For sufficiently strong positive feedback, the system becomes irreversible (light line); it becomes “locked” at steady-state c once it jumps from a to c . **C** Stochasticity undermines bistability. Although deterministically the system is either at steady-state a or c , stochastic effects generate non-zero probabilities to be at either steady-state in the bistable regime. **D** Single-cell measurements can reveal differences between graded and bistable responses that are lost in population-averaged measurements [30].

generate output that can initiate runaway positive feedback, and all initial conditions tend to one “inactivated” steady-state. As the input increases, however, the system can undergo what is called a saddle-node bifurcation that creates a second steady-state (at I_1 in Fig. 4B): the input has become sufficiently high to generate positive feedback strong enough to create bistability. This second, “activated” steady-state is not apparent at this level of input though, and only on further increases is there a dramatic, discontinuous change in output at a critical threshold of input when the system “jumps” from one steady-state to the other (at I_2 in Fig. 4B). This jump occurs at another saddle-node bifurcation where the inactivated steady-state is lost because the positive feedback is so strong that all systems have sufficient output to generate runaway feedback. Regardless of the initial condition, there is just one steady-state, and the system is no longer bistable. At this new steady-state, the output again increases monotonically with the input. If the input is then lowered, the strength of the feedback falls, a saddle-node bifurcation occurs (at I_2), and the low, inactivated steady-state is again created at the same threshold of input where it was lost. This bifurcation does not typically perturb the dynamics of the system until the input is lowered sufficiently that a second saddle-node bifurcation removes the activated state and the bistability of the system (at I_1). At this threshold of input, the steady-state behaviour of the system changes discontinuously to the inactivated steady-state, and there its output decreases continuously with decreasing input. The system has memory, and the past affects whether the system jumps or not. For example, if the system has already experienced levels of input higher than I_2 , there will be a jump at I_1 as the input decreases; if the system has not previously experienced input higher than I_2 , there will be no jump at I_1 . Such history-dependent behaviour is called hysteresis. It allows information to be stored despite potentially rapid turnover of the system’s molecules [32].

Permanent memory can be created by sufficiently strong feedback. If the feedback is so strong that the system is bistable even with no input then once the input is enough to drive the system from the inactivated to the activated state then no transition back to the inactivated state is possible. Although biochemical bistability has been proposed to describe neural memories [33, 34], permanent memory is perhaps more appropriate for differentiation where cells should remain differentiated once committed to differentiation [35].

Stochasticity undermines memory

Biochemical systems are stochastic, and stochasticity undermines bistability by generating fluctuations that drive the system from one stable steady-state to the other. Typically, we can think of two sets of timescales in the system: fast timescales that control the return of the system to steady-state in response to small, frequent perturbations and determine the lifetime of typical fluctuations around a steady-state; and slow timescales generated by rare fluctuations between steady-states that determine the time spent by the system at each steady-state. Stochasticity in a bistable system can cause the distribution for the numbers of molecules to become bimodal (Fig. 4C). Frequent (fast) fluctuations determine the width of this distribution close to each steady-state; rare fluctuations determine the relative height of each peak. If both fluctuations are common, they can drive a bimodal distribution to be unimodal. With measurements of populations of cells, however, averaging can obscure bistable behaviour [30]: the average response of a population of bistable cells, particularly if stochastic effects cause each cell to respond at a different threshold of input concentration, and the average response of cells with a monostable, graded behaviour can be indistinguishable (Fig. 4D).

Nevertheless, a bimodal distribution of, for example, the numbers of protein does not mean that the system's underlying dynamics are necessarily bistable. Even expression from a single gene can generate a bimodal distribution if the promoter transitions slowly between an active and an inactive state [23, 22, 36, 37, 38]. This bimodality is lost if fluctuations at the promoter are sufficiently fast because they are then averaged by other processes of gene expression [39].

Estimating the time spent at each steady-state

A challenge for theory is to calculate the fraction of time spent near each steady-state in a bistable system. Such a calculation requires a description of the rare fluctuations that drive the system from one steady-state to the other. It therefore cannot be tackled with Van Kampen's linear noise approximation [40], the workhorse of stochastic systems biology, because this approximation requires that fluctuations generate only small changes in numbers of molecules compared to the deterministic behaviour and therefore can only be applied near one of the system's steady-states. Formally, the master equation for a vector of molecules \mathbf{n} obeys

$$\frac{\partial}{\partial t}P(\mathbf{n}, t) = AP(\mathbf{n}, t) \quad (6)$$

where the operator A is determined by the chemical reactions (see also the chapter by B. Munsky). This equation has a solution in terms of an expansion of the eigenfunctions of A :

$$P(\mathbf{n}, t) = a_0\psi_0(\mathbf{n}) + \sum_{i=1} a_i e^{-\lambda_i t} \psi_i(\mathbf{n}) \quad (7)$$

where $\psi_i(t)$ is the eigenfunction of A with eigenvalue λ_i . The steady-state distribution $P^{(s)}(\mathbf{n})$ is determined by the zero'th eigenfunction, $P^{(s)}(\mathbf{n}) = a_0\psi_0(\mathbf{n})$, which has a zero eigenvalue. If the system is bistable, this distribution will typically be bimodal with maxima at $\mathbf{n} = \mathbf{a}$ and $\mathbf{n} = \mathbf{c}$, say, and a minimum at $\mathbf{n} = \mathbf{b}$, corresponding deterministically to stable steady-states at a and c and an unstable steady-state at b (Fig. 4C). The next smallest eigenvalue λ_1 determines the slow timescales. As the system approaches steady-state, peaks at a and c will develop, and we expect metastable behaviour with molecules mostly being near the steady-state at either a or c but undergoing occasional transitions between a and c . The probability distribution at this long-lived metastable stage will be approximately

$$P(\mathbf{n}, t) \simeq P^{(s)}(\mathbf{n}) + a_1 e^{-\lambda_1 t} \psi_1(\mathbf{n}). \quad (8)$$

Defining $\Pi_{\mathbf{a}}$ as the probability of the system being near the steady-state at \mathbf{a} and $\Pi_{\mathbf{c}}$ as the probability of the system being near the steady-state at \mathbf{c} :

$$\Pi_{\mathbf{a}} = \int_{\mathbf{0}}^{\mathbf{b}} d\mathbf{n}P(\mathbf{n}, t) \quad ; \quad \Pi_{\mathbf{c}} = \int_{\mathbf{b}}^{\infty} d\mathbf{n}P(\mathbf{n}, t) \quad (9)$$

then we expect

$$\frac{\partial \Pi_{\mathbf{a}}}{\partial t} = -\frac{\partial \Pi_{\mathbf{c}}}{\partial t} = \kappa_{\mathbf{c} \rightarrow \mathbf{a}} \Pi_{\mathbf{c}} - \kappa_{\mathbf{a} \rightarrow \mathbf{c}} \Pi_{\mathbf{a}} \quad (10)$$

because $\Pi_{\mathbf{a}} + \Pi_{\mathbf{c}} = 1$, and if $\kappa_{\mathbf{a} \rightarrow \mathbf{c}}$ is the rate of fluctuating from the neighbourhood of \mathbf{a} to the neighbourhood of \mathbf{c} and $\kappa_{\mathbf{c} \rightarrow \mathbf{a}}$ is the rate of fluctuating back. At steady-state, we have $\Pi_{\mathbf{a}}^{(s)} \kappa_{\mathbf{a} \rightarrow \mathbf{c}} = \Pi_{\mathbf{c}}^{(s)} \kappa_{\mathbf{c} \rightarrow \mathbf{a}}$. If we integrate Eq. 8 with respect to \mathbf{n} following Eq. 9, insert the resulting expressions into Eq. 10, and simplify with the steady-state equation, we find that $\kappa_{\mathbf{a} \rightarrow \mathbf{c}} + \kappa_{\mathbf{c} \rightarrow \mathbf{a}} =$

λ_1 [41]: the rate of fluctuating from one steady-state to another is, indeed, determined by the slow timescale, $1/\lambda_1$.

For a one-dimensional system where chemical reactions, which may occur with non-linear rates, only increase or decrease the number of molecules by one molecule (bursting, for example, cannot occur), an analytical expression for the average time to fluctuate from one steady-state to another can be calculated [40]. The exact initial and final numbers of molecules are not important because the time to jump between steady-states is determined by the slow timescales and the additional time, determined by the fast timescales, to move the molecules between different initial conditions or between different final conditions is usually negligible. If n molecules are present, the probability of creating a molecule is $\Omega w_+(n)$ for a system of volume Ω and the probability of degrading a molecule is $\Omega w_-(n)$, then the steady-state probability of the concentration $x = n/\Omega$ is approximately [42]

$$P^{(s)}(x) \sim e^{-\Omega\phi(x)} \quad (11)$$

for large Ω and where $\phi(x)$ is an effective potential:

$$\phi(x) = - \int_0^x dx' \log \frac{w_+(x')}{w_-(x')}. \quad (12)$$

To calculate the average time to fluctuate between steady-states, typically the average time, $t_{a \rightarrow b}$, to fluctuate from $x = a$ to b , where b is the deterministically unstable steady-state, is calculated because this time is much longer than the time to fluctuate from b to c , which is determined by the fast timescales. Then $\kappa_{a \rightarrow c} = \frac{1}{2} t_{a \rightarrow b}^{-1}$ because at $x = b$ a fluctuation to either a or c is equally likely. For large Ω , the rates are [42]

$$\kappa_{a \rightarrow c} \simeq \frac{1}{2\pi} w_+(a) \sqrt{\phi''(a) |\phi''(b)|} e^{-\Omega[\phi(b) - \phi(a)]} \quad (13)$$

and

$$\kappa_{c \rightarrow a} \simeq \frac{1}{2\pi} w_+(c) \sqrt{\phi''(c) |\phi''(b)|} e^{-\Omega[\phi(b) - \phi(c)]} \quad (14)$$

and depend exponentially on the difference in the effective potential at the unstable and a stable steady-state. They are exponentially small in the size of the system, Ω , and therefore vanish as expected in the deterministic limit. In this deterministic limit, the dynamics of x obey

$$\dot{x} = w_+(x) - w_-(x) \quad (15)$$

and the potential of motion, $-\int dx' [w_+(x') - w_-(x')]$, is not the same as the effective potential of Eq. 12 [42]. The rates are exponentially sensitive to changes in the parameters of the system because the parameters determine the effective potential. Such exponential sensitivity implies a large variation in fluctuation-driven switching rates between cells, perhaps suggesting that other, more robust switching mechanisms also exist [43].

To estimate transition rates between steady-states for systems that cannot be reduced to one variable, numerical methods are most often used, although some analytical progress is possible [44]. Perhaps the most efficient numerical algorithm developed to estimate transition rates is forward flux sampling [45]. The two steady-states, a and c , are defined through a parameter λ , which may be a function of all the chemical species in the system. The system is at steady-state a if $\lambda \leq \lambda_a$ for some λ_a and at steady-state c if $\lambda \geq \lambda_c$ for some λ_c . A series of interfaces is then defined between λ_a and λ_c . If a trajectory starts from steady-state a and reaches one

of the interfaces, then we can consider the probability that this trajectory will reach the next interface closest to λ_c on its first crossing of the current interface. The transition rate $\kappa_{a \rightarrow c}$ is approximated by a suitably normalized product of these probabilities for all interfaces between λ_a and λ_c [45]. The probabilities can be estimated sequentially using the Gillespie algorithm. Initially simulations are started from a to estimate the first probability and generate a collection of initial conditions for the estimation of the second probability. These initial conditions are given by the state of the system each time the simulation trajectory comes from a and crosses the first interface. Further simulation from these initial conditions then allows estimation of the second probability and generation of the initial conditions for estimation of the next probability, and so on. The efficiency of the method is determined by the choice of the parameter λ and the accuracy by the number of interfaces between λ_a and λ_c and the number of initial conditions used to estimate each probability. In general, forward flux sampling can be highly efficient [45].

Examples

The *lac* operon in *E. coli*

The *lac* operon of *E. coli* was first studied by Jacob and Monod (see [31] for an early summary of some of their and others' work), and indeed was the first system for which the basic mechanism of transcriptional regulation was understood (see Wilson *et al.* [46] and references therein). The operon includes genes important in the intake and metabolism of lactose: *lacZ* and *lacY* (Fig. 5A). In particular, LacZ is an enzyme that breaks lactose down into glucose and galactose, and LacY encodes a permease that transports lactose into the cell. Expression of the *lac* operon is repressed by the transcription factor LacI, appropriately called the lac repressor. LacI binds to the promoter of the *lac* operon, where it blocks transcription. If intracellular allolactose, a metabolite of lactose, binds to LacI, however, then LacI loses its ability to bind to the promoter, and the *lac* operon becomes de-repressed. These qualitative influences are summarized in Figure 5A, with some abridgement in intracellular lactose metabolism [46].

The system has the potential for strong positive feedback. In a low lactose environment, intracellular lactose, and thus allolactose, is negligible. LacI remains unbound by allolactose, and acts to repress expression of the *lac* operon. Conversely, in a high lactose environment, intracellular lactose, and thus allolactose, becomes substantial, and binds up all the free molecules of LacI. Without LacI binding the promoter, the *lac* operon becomes expressed and leads to the synthesis of more LacY permeases. These permeases increase the cell's potential to import lactose and therefore the potential for greater expression of the *lac* operon: the system has positive feedback through the double negative feedback of lactose on LacI and LacI on expression.

Nevertheless, bistable behaviour has only been observed with thio-methylgalactoside (TMG), a non-metabolizable analogue of lactose [47, 48]. TMG also binds LacI and de-represses the *lac* operon. When the amount of TMG in the extracellular environment is at an intermediate level and if there is only a small number of LacY permeases in the cell membrane, perhaps because the *lac* operon has not been expressed for some time, then only a small amount of TMG will enter the cell – not enough to de-repress the *lac* operon. On the other hand, if the membrane has many LacY permeases, perhaps because of recent *lac* operon expression, then enough TMG can enter the cell to enhance expression of the operon. The exact “shape” of this bistable behaviour has been mapped in some detail, along with the influence of other factors (Fig. 5C & D)[49, 47, 50].

How does stochasticity influence the function of the *lac* regulatory machinery? The early

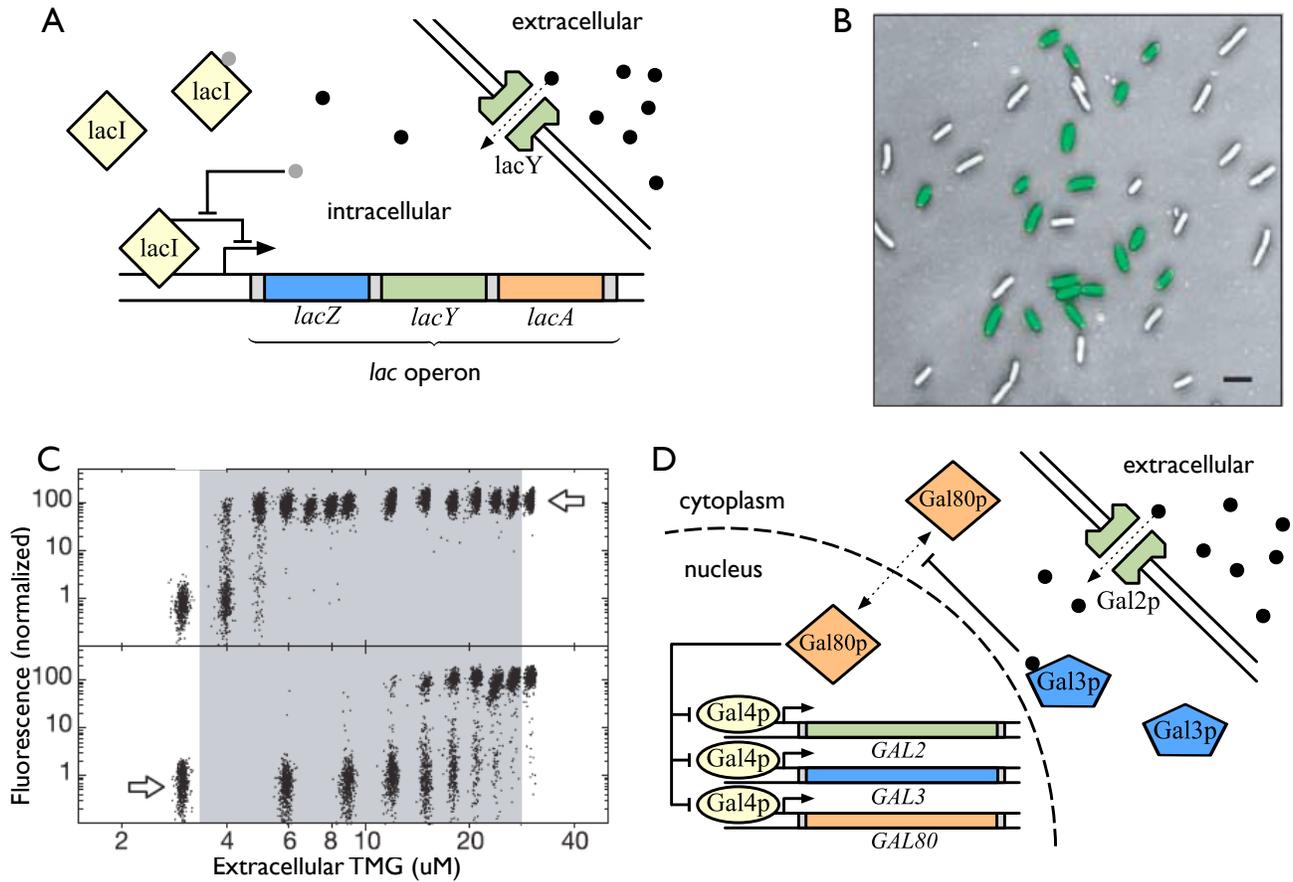


Figure 5: Sugar utilization networks in *E. coli* and *S. cerevisiae*. **A** The *lac* operon of *E. coli*. The *lac* operon comprises three genes, one of which, *lacY*, encodes a lactose permease. The permease allows lactose (black circles) to enter the cell. Intracellular allolactose (gray circles), a metabolite of lactose, binds the transcription factor LacI, preventing it from repressing the operon [46]. **B** Bistable responses observed in uninduced *E. coli* cells upon exposure to 18 μM of TMG, a nonmetabolizable analog of lactose. Levels of induction in each cell is monitored by a Green Fluorescent Protein (GFP) reporter controlled by a chromosomally integrated *lac* promoter. From Ozbudak *et al.* (2004). **C** A hysteresis curve is an unequivocal proof of bistability. Scatterplots of the GFP reporter and another estimating catabolite repression are shown. In the top panel, cells are initially fully induced; in the lower panels, cells are initially uninduced. Bistable behaviour occurs in the gray regime, and there are two separate thresholds of TMG concentration for the “on”-to-“off” and “off”-to-“on” threshold transitions. From Ozbudak *et al.* (2004). **D** The galactose utilization network in *S. cerevisiae*. Galactose enters the cell via Gal2p and binds the sensor Gal3p, which then inhibits transcriptional repression of the GAL genes by the repressor Gal80p. This inhibition of repression increases synthesis of both Gal2p and Gal3p giving positive feedback and of Gal80p giving negative feedback. Gal4p is here a constitutively expressed activator.

work of Novick and Weiner was remarkably prescient [51]. They were the first to recognize the bimodal expression of the *lac* operon at intermediate extracellular TMG concentrations, as well as the hysteretic nature of its regulation. Moreover, by detailed study of the kinetics of the induction process, they concluded that a single, random chemical event leads to a “runaway” positive feedback that fully induces the operon – a hypothesis that is essentially correct. They hypothesized that this single event was a single lactose permease entering the cell membrane, whereupon the TMG “floodgate” was opened. This explanation is incorrect, though it took decades of biochemical research as well as advances in biotechnology to establish the true mechanism [52]. A typical uninduced *E. coli* harbours a handful of lactose permeases, perhaps as many as ten, in its membrane. These are the result of small bursts of expression from the operon. Induction at an intermediate level of extracellular TMG occurs when a random much larger burst of expression produces on the order of hundreds of permeases, which then sustain expression. The mechanism supporting these large and small bursts is intimately connecting with the DNA binding behaviour of LacI. There are binding sites on either side of the operon, and, through DNA looping, a single LacI molecule is able to bind on both sides. Short-lived dissociations from one side or the other allow small bursts of expression. Complete dissociation from the DNA, a much rarer event, leads to a much longer period of de-repression and a large burst of expression. Thus, as Novick and Weiner hypothesized, it is a single, rare, random chemical event – complete dissociation of LacI from the DNA – which flips the switch from uninduced to induced [52].

There are, of course, other stochastic aspects to the system. LacI levels fluctuate, as do permease levels. These fluctuations are much smaller than the difference between “off” and “on” states [47]. Nevertheless, Robert *et al.* [53] have shown that various measurable properties of individual cells can be used to predict whether the cells will become induced or not at intermediate levels of inducer. Moreover, they show that these proclivities can be inherited epigenetically. Thus, the stochastic flipping of the *lac* operon switch is driven by both intrinsic and extrinsic fluctuations, with the latter just beginning to be identified.

The GAL system in budding yeast

The theme of positive feedback via permeases resulting in bistability and hysteresis also occurs in the arabinose utilization system in *E. coli* [54], but a better-known example is the galactose utilization network of *S. cerevisiae*, which we treat here briefly. In this system, the master activator is the transcription factor Gal4p, which turns on a number of regulatory and metabolism-related genes, including the galactose permease *GAL2* and the galactose sensor *GAL3* [4]. In the absence of intracellular galactose, the repressor Gal80p binds to Gal4p and inhibits its activation of expression. When galactose binds to Gal3p, however, Gal3p becomes activated and can bind to Gal80p and so remove its ability to repress Gal4p (Fig. 5B). The system has two positive feedbacks through expression of both the permease, *GAL2*, and the sensor, *GAL3*, once Gal4p is no longer repressed by Gal80p, and one negative feedback through expression of *GAL80*.

As with the *lac* operon, yeast cells in an intermediate extracellular galactose concentration may be either stably “on” or “off”. However, that stability is relative to the stochasticity of the biochemistry, so that at sufficiently long time scales any cell may turn on or turn off. Indeed, Acar *et al.* and Song *et al.* observed bimodal and history-dependent distributions for the galactose response of *S. cerevisiae* (analogous to those of Fig. 5D) [55, 56]. Acar *et al.*, by sorting a bimodal population to select only the “off” or “on” cells and subsequently assaying the population distribution over time, furthermore estimated the duration of memory in the system

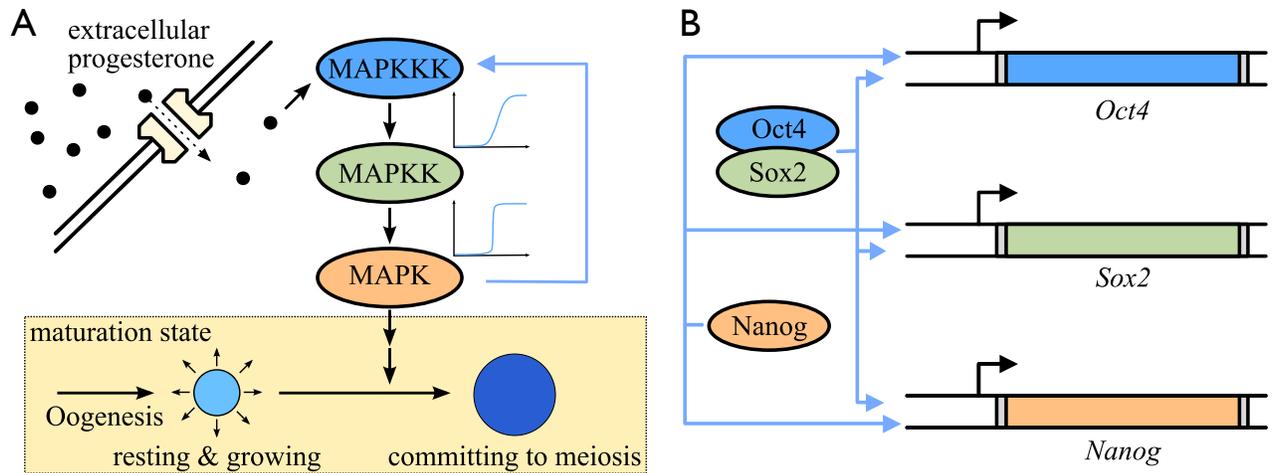


Figure 6: Positive feedback in higher organisms. **A** Progesterone-induced oocyte maturation in *X. laevis*. Through both the activation of a MAP kinase cascade that converts a graded signal into a switch-like response and positive feedback from MAPK to MAPKKK (blue arrow), progesterone induces all-or-none, irreversible maturation of oocytes. Inset: the degree of ultrasensitivity in the response increases down the cascade [57]. **B** The core genetic network driving pluripotency in human embryonic stem cell. Blue arrows represent positive transcriptional regulation. Diagram adapted from Boyer *et al.* [58].

could be at least on the order of tens if not hundreds of hours, depending on experimental parameters [55]. Perhaps surprisingly, the *GAL2* gene is not necessary for bistability. Upon its deletion, bistability is still observed, albeit with an “on” state with significantly less expression than in the wild-type [55]. Galactose is still able to enter the cell via other means, and the positive feedback loop via *GAL3* mediates bistability. Removing the feedbacks to *GAL3* or *GAL80*, by putting the genes under the control of different promoters, eliminates bistability. When *GAL3* is not activated by Gal4p, but instead by an alternate activator, cell populations show a unimodal distribution of induction, and one that loses any history dependence. Interestingly, when the gene encoding the repressor Gal80p is not activated by Gal4p, cell populations can show a bimodal induction distribution – some cells “off” and some “on” – but one with little history dependence. That is, individual cells switch comparatively rapidly between the two states, though there is still some tendency to be in one state or the other, and not in between. Acar *et al.* dubbed this phenomenon “destabilized memory” [55].

Bistability in higher organisms

Perhaps the most understood example of bistable behaviour in higher organisms is oocyte maturation in the frog *Xenopus laevis*. After S-phase, immature oocytes carry out several early steps of meiotic prophase, but only proceed with the first meiotic division in response to the hormone progesterone. The commitment to meiosis is all-or-none and irreversible. It is generated biochemically by two mechanisms: a mitogen-activated protein (MAP) kinase cascade of three kinases and a positive feedback loop from the last kinase, MAP kinase, to the first kinase, MAP kinase kinase kinase (Fig. 6A). The MAP kinase cascade converts the input signal of progesterone

terone into an ultrasensitive or switch-like response with each level of the cascade increasing the degree of ultrasensitivity [59]. This ultrasensitivity is necessary to generate bistable behaviour, but not sufficient, and the additional positive feedback is required [30]. When this feedback is blocked, the cells exhibit intermediate responses that are not irreversible (Fig. 4B) [35].

Such bistability with strong positive feedback might be expected in differentiated cells in a multicellular organism. Differentiation represents the locking-in of an extremely stable state – one that lasts until the death of the cell, potentially many years. How then do stem cells fit into this framework?

Accumulating evidence suggests that stem cells are “poised” for differentiation and must be actively maintained in their “stem” state [60]. Stem cells in culture must be treated with various chemical factors to prevent differentiation (e.g., [61, 62]). In this sense, the stem state is not naturally stable. *In vivo*, a variety of signals affect stem cell function. For example, many stem cell systems become less active, and thus less able to repair or maintain tissue, as an animal ages [63, 64, 65]. However, this loss of activity does not appear to be from changes in the stem cells themselves. Subjecting the cells to signals present in younger animals re-enlivens the aged cells [65, 66]. Given that the right signals are provided, are stem cells at a kind of conditional steady state? Stem cell identity is usually defined by one or a few “master regulator” genes. For example, human embryonic stem cells require the coordinated expression of three genes, Sox2, Oct4 and Nanog, which are connected in positive feedback loops (Fig. 6B) [58]. Several works have found bimodal expression distributions for these genes in populations of cultured cells, suggesting some form of stochastic bistability [67, 68], and various models have been proposed [69, 70, 71]. Although thinking of the pluripotency properties of stem cells as driven by stochastic, multistable genetic networks is perhaps the best approach [72], there are many challenges to link these ideas to mechanisms and quantitative predictions.

Conclusions

Bistability is one mechanism that cells can use to retain memory. If the positive feedback driving the bistability is sufficiently strong, then once a biochemical network “switches”, it can never switch back (Fig. 4B). Such permanent memories were once thought to be maintained as long as the cell has sufficient resources to generate the positive feedback. We now recognize that stochasticity is pervasive both intra- and extracellularly and that fluctuations can undermine memory with rare events causing networks that are “permanently” on to switch off.

Stochasticity also challenges our study of bistable systems. We should study single cells because extrinsic fluctuations can “wash-out” bistable responses when averaged over measurements of populations of cells [30]. We should look for hysteresis because bimodal distributions can be generated by slow reactions in biochemical networks that have no positive feedback [23, 22, 36, 37, 38]. Such short-term memory, however, may be sufficient for some systems, and indeed positive feedback need not generate bistability but simply long-lived transient behaviour [73]. Single-cell studies, and particularly the time-lapse studies useful for investigating cellular memory, are challenging. Nevertheless, developments in microfluidic technology are easing the generation of dynamic changes in the cellular environment and the quantification and tracking of individual cellular responses over many generations [17, 18]. We expect that such experiments will reveal new insights into cellular memory and its role in cellular decision-making [74].

By having switches of greater complexity, cells can limit the effects of stochastic fluctuations,

but cells can also evolve biochemistry that exploits stochasticity. Bet-hedging by populations of microbes is one example [75], where some cells in the population temporarily and stochastically switch to a behaviour that is inefficient in the current environment but potentially efficient if a change in the environment occurs. Perhaps the most intriguing example is, however, in our own cells. Mammalian embryonic stem cells appear to use stochasticity to become temporarily “primed” for different cell fates and so to become susceptible over time to the different differentiation signals for all possible choices of cell fate [60].

References

- [1] Scoville WB, Milner B (2000) Loss of recent memory after bilateral hippocampal lesions. *J Neuropsychiatry Clin Neurosci* 12:103–13.
- [2] Berg HC, Brown DA (1972) Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature* 239:500–4.
- [3] Muller-Hill B (1996) *The lac operon: a short history of a genetic paradigm* (Walter de Gruyter, Berlin).
- [4] Lohr D, Venkov P, Zlatanova J (1995) Transcriptional regulation in the yeast GAL gene family: a complex genetic network. *FASEB J* 9:777–87.
- [5] Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–76.
- [6] Lewis K (2007) Persister cells, dormancy and infectious disease. *Nat Rev Micro* 5:48–56.
- [7] Shahrezaei V, Swain PS (2008) The stochastic nature of biochemical networks. *Curr Opin Biotechnol* 19:369–74.
- [8] Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135:216–26.
- [9] Ghaemmaghami S, *et al.* (2003) Global analysis of protein expression in yeast. *Nature* 425:737–41.
- [10] Taniguchi Y, *et al.* (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* 329:533–8.
- [11] Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* 99:12795–800.
- [12] Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297:1183–6.
- [13] Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. *Nat Genet* 31:69–73.
- [14] Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. *Nature* 422:633–7.

- [15] Raser JM, O’Shea EK (2004) Control of stochasticity in eukaryotic gene expression. *Science* 304:1811–4.
- [16] Shahrezaei V, Ollivier JF, Swain PS (2008) Colored extrinsic fluctuations and stochastic gene expression. *Mol Syst Biol* 4:196.
- [17] Bennett MR, Hasty J (2009) Microfluidic devices for measuring gene network dynamics in single cells. *Nat Rev Genet* 10:628–38.
- [18] Wang CJ, Levchenko A (2009) Microfluidics technology for systems biology research. *Methods Mol Biol* 500:203–19.
- [19] Yu J, Xiao J, Ren X, Lao K, Xie XS (2006) Probing gene expression in live cells, one protein molecule at a time. *Science* 311:1600–3.
- [20] McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA* 94:814–9.
- [21] Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci USA* 98:8614–9.
- [22] Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. *Proc Natl Acad Sci USA* 105:17256–61.
- [23] Friedman N, Cai L, Xie XS (2006) Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys Rev Lett* 97:168302.
- [24] Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. *Cell* 123:1025–36.
- [25] Voliotis M, Cohen N, Molina-París C, Liverpool TB (2008) Fluctuations, pauses, and backtracking in DNA transcription. *Biophys J* 94:334–48.
- [26] Chubb JR, Treck T, Shenoy SM, Singer RH (2006) Transcriptional pulsing of a developmental gene. *Curr Biol* 16:1018–25.
- [27] Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. *Plos Biol* 4:e309.
- [28] Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361.
- [29] Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB (2005) Gene regulation at the single-cell level. *Science* 307:1962–5.
- [30] Ferrell JE, Machleder EM (1998) The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* 280:895–8.
- [31] Monod J, Jacob F (1961) General conclusions: Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* 26:389.

- [32] Lisman JE (1985) A mechanism for memory storage insensitive to molecular turnover: a bistable autophosphorylating kinase. *Proc Natl Acad Sci USA* 82:3055–7.
- [33] Lisman JE, Goldring MA (1988) Feasibility of long-term storage of graded information by the Ca^{2+} /calmodulin-dependent protein kinase molecules of the postsynaptic density. *Proc Natl Acad Sci USA* 85:5320–4.
- [34] Bhalla US, Iyengar R (1999) Emergent properties of networks of biological signaling pathways. *Science* 283:381–7.
- [35] Xiong W, Ferrell JE (2003) A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* 426:460–5.
- [36] Pirone JR, Elston TC (2004) Fluctuations in transcription factor binding can explain the graded and binary responses observed in inducible gene expression. *J Theor Biol* 226:111–21.
- [37] Karmakar R, Bose I (2004) Graded and binary responses in stochastic gene expression. *Physical biology* 1:197–204.
- [38] Hornos JEM, *et al.* (2005) Self-regulating gene: an exact solution. *Phys Rev E* 72:051907.
- [39] Qian H, Shi PZ, Xing J (2009) Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity. *Phys Chem Chem Phys* 11:4861–70.
- [40] Van Kampen NG (1981) *Stochastic processes in physics and chemistry* (North-Holland, Amsterdam, The Netherlands).
- [41] Procaccia I, Ross J (1977) Stability and relative stability in reactive systems far from equilibrium. ii. Kinetic analysis of relative stability of multiple stationary states. *J Chem Phys* 67:5565–5571.
- [42] Hanggi P, Grabert H, Talkner P, Thomas H (1984) Bistable systems: Master equation versus Fokker-Planck modeling. *Phys Rev, A* 29:371–378.
- [43] Mehta P, Mukhopadhyay R, Wingreen NS (2008) Exponential sensitivity of noise-driven switching in genetic networks. *Physical biology* 5:26005.
- [44] Roma DM, O'Flanagan RA, Ruckenstein AE, Sengupta AM, Mukhopadhyay R (2005) Optimal path to epigenetic switching. *Phys Rev E* 71:011902.
- [45] Allen RJ, Warren PB, ten Wolde PR (2005) Sampling rare switching events in biochemical networks. *Phys Rev Lett* 94:018104.
- [46] Wilson CJ, Zhan H, Swint-Kruse L, Matthews KS (2007) The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell Mol Life Sci* 64:3–16.
- [47] 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.
- [48] Santillán M, Mackey MC, Zeron ES (2007) Origin of bistability in the lac operon. *Biophys J* 92:3830–42.

- [49] Setty Y, Mayo AE, Surette MG, Alon U (2003) Detailed map of a cis-regulatory input function. *Proc Natl Acad Sci USA* 100:7702–7.
- [50] Kuhlman T, Zhang Z, Saier MH, Hwa T (2007) Combinatorial transcriptional control of the lactose operon of Escherichia coli. *Proc Natl Acad Sci USA* 104:6043–8.
- [51] Novick A, Weiner M (1957) Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci USA* 43:553–66.
- [52] Choi PJ, Cai L, Frieda K, Xie XS (2008) A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* 322:442–6.
- [53] Robert L, *et al.* (2010) Pre-dispositions and epigenetic inheritance in the escherichia coli lactose operon bistable switch. *Mol Syst Biol* 6:357.
- [54] Megerle JA, Fritz G, Gerland U, Jung K, Rädler JO (2008) Timing and dynamics of single cell gene expression in the arabinose utilization system. *Biophys J* 95:2103–15.
- [55] Acar M, Becskei A, van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435:228–32.
- [56] Song C, *et al.* (2010) Estimating the stochastic bifurcation structure of cellular networks. *PLoS Comput Biol* 6:e1000699.
- [57] Huang CY, Ferrell JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci USA* 93:10078–83.
- [58] Boyer LA, *et al.* (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947–56.
- [59] Ferrell JE (1996) Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends in Biochemical Sciences* 21:460–6.
- [60] Silva J, Smith A (2008) Capturing pluripotency. *Cell* 132:532–6.
- [61] Smith AG, *et al.* (1988) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* 336:688–90.
- [62] Ying QL, Nichols J, Chambers I, Smith A (2003) BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with Stat3. *Cell* 115:281–92.
- [63] Kuhn HG, Dickinson-Anson H, Gage FH (1996) Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *J Neurosci* 16:2027–33.
- [64] Morrison SJ, Wandycz AM, Akashi K, Globerson A, Weissman IL (1996) The aging of hematopoietic stem cells. *Nat Med* 2:1011–6.
- [65] Conboy IM, Conboy MJ, Smythe GM, Rando TA (2003) Notch-mediated restoration of regenerative potential to aged muscle. *Science* 302:1575–7.

- [66] Conboy IM, *et al.* (2005) Rejuvenation of aged progenitor cells by exposure to a young systemic environment. *Nature* 433:760–4.
- [67] Chambers I, *et al.* (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* 450:1230–4.
- [68] Davey RE, Onishi K, Mahdavi A, Zandstra PW (2007) LIF-mediated control of embryonic stem cell self-renewal emerges due to an autoregulatory loop. *FASEB J* 21:2020–32.
- [69] Chickarmane V, Troein C, Nuber UA, Sauro HM, Peterson C (2006) Transcriptional dynamics of the embryonic stem cell switch. *PLoS Comput Biol* 2:e123.
- [70] Kalmar T, *et al.* (2009) Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol* 7:e1000149.
- [71] Glauche I, Herberg M, Roeder I (2010) Nanog variability and pluripotency regulation of embryonic stem cells – insights from a mathematical model analysis. *PLoS ONE* 5:e11238.
- [72] Huang S (2009) Reprogramming cell fates: reconciling rarity with robustness. *Bioessays* 31:546–60.
- [73] Weinberger LS, Dar RD, Simpson ML (2008) Transient-mediated fate determination in a transcriptional circuit of HIV. *Nat Genet* 40:466–70.
- [74] Perkins TJ, Swain PS (2009) Strategies for cellular decision-making. *Mol Syst Biol* 5:326.
- [75] Veening JW, Smits WK, Kuipers OP (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* 62:193–210.