Interface focus

*J. R. Soc. Interface* (2008) **5**, S29–S39
doi:10.1098/rsif.2008.0086.focus
*Published online* 15 April 2008

Review

# Quantitative approaches to the study of bistability in the *lac* operon of *Escherichia coli*

**Moisés Santillán[1,2,*] and Michael C. Mackey[2]**

[1]*Centro de Investigación y Estudios Avanzados del IPN, Unidad Monterrey, Parque de Investigación e Innovación Tecnológica, 66600 Apodaca, NL, Mexico*
[2]*Centre for Nonlinear Dynamics in Physiology and Medicine, McGill University, 3655 Promenade Sir William Osler, Montreal, Que., Canada H3G 1Y6*

In this paper, the history and importance of the *lac* operon in the development of molecular and systems biology are briefly reviewed. We start by presenting a description of the regulatory mechanisms in this operon, taking into account the most recent discoveries. Then we offer a survey of the history of the *lac* operon, including the discovery of its main elements and the subsequent influence on the development of molecular and systems biology. Next the bistable behaviour of the operon is discussed, both with respect to its discovery and its molecular origin. A review of the literature in which this bistable phenomenon has been studied from a mathematical modelling viewpoint is then given. We conclude with some brief remarks.

**Keywords: *lac* operon; multistability; gene regulatory networks; systems biology; computational biology; mathematical model**

## 1. INTRODUCTION

Glucose is the favourite carbon and energy source for *Escherichia coli*, as well as for many other organisms. Although this bacterium can also feed on other sugars, it only does so when glucose is absent. Thus, if a bacterial culture grows in a medium containing a mixture of glucose and another sugar (such as lactose), it will exclusively feed on the former until it is exhausted, before switching on to the second one. A consequence of this behaviour is that the bacterial growth curve shows two distinctive phases, as can be seen in figure 1. This phenomenon was originally studied by Monod (1941), who described it as diauxic growth. It is worth mentioning at this point that diauxic growth only occurs in batch cultures, and simultaneous usage of sugars is often observed in continuous cultures (Lendenmann *et al.* 1996).

Molecular level understanding of how an organism sequentially uses different metabolites has been attracting tremendous interest for the past fifty years. Jacob & Monod (1961*a*,*b*) tackled this problem and conceptually outlined how bacterial cultures could switch from one mode of growth to another so rapidly and completely. In the process of doing so, they introduced the operon concept (Jacob *et al.* 1960), which has become a paradigmatic example of gene regulation.

According to the construction of Jacob & Monod, an operon consists of a set of structural genes that are regulated together, depending on the cell metabolic requirements. These structural genes code for a group of enzymes or proteins that are responsible for a specific task or metabolic process, and their regulation is achieved via one or more common regulatory mechanisms. Repression was the first regulatory mechanism to be discovered by them. In it, a repressor molecule binds a specific DNA site (which they termed the operator) located upstream from the structural genes, and inhibits their transcription. The regulation of the structural genes' expression is achieved by varying the number of active repressor molecules. Although Jacob and Monod originally thought of the repressors as RNA molecules, they are now known to be proteins.

The lactose (or simply *lac*) operon is composed of three structural genes: *lacZ*, *lacY* and *lacA*. These genes, respectively, code for β-galactosidase, *lac* permease and a transacetylase. β-Galactosidase acts to cleave lactose into galactose and glucose, which is the first step in lactose metabolism; *lac* permease is a transmembrane protein, which is necessary for lactose uptake; transacetylase transfers an acetyl group from coenzyme A (CoA) to the hydroxyl group of the galactosides. Of these proteins,

*Author and address for correspondence: Centro de Investigación y Estudios Avanzados del IPN, Unidad Monterrey, Parque de Investigación e Innovación Tecnológica, 66600 Apodaca NL, Mexico (msantillan@cinvestav.mx).

One contribution of 10 to a Theme Supplement 'Biological switches and clocks'.

*Received* 27 February 2008
*Accepted* 27 March 2008

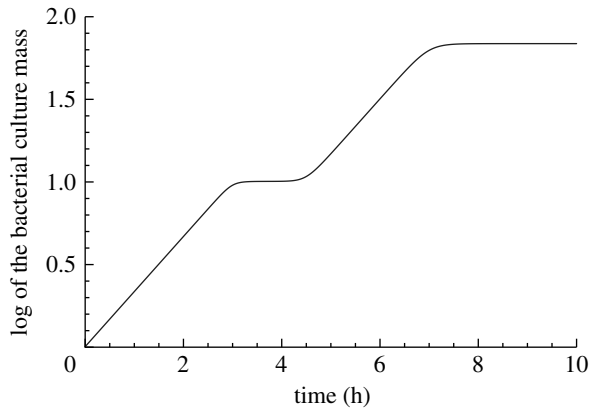


Figure 1. Typical diauxic growth curve. Note the existence of two different exponential growth phases, separated by a short interval in which the culture does not grow. The first (second) phase corresponds to the bacterial culture feeding on glucose (lactose), while the interval with no growth corresponds to the time the bacteria need to turn on the genes needed to metabolize lactose after glucose exhaustion.

only β-galactosidase and *lac* permease play an active role in the regulation of the *lac* operon. The regulatory gene *lacI* (in a different operon) codes for the *lac* repressor which, when active, is capable of inhibiting the transcription of the structural genes by binding an operator. The *lac* repressor is inactivated when it is bound by allolactose, a by-product of lactose metabolism. Finally, the *lac* operon genes are also upregulated by an activator that increases the affinity of the mRNA polymerase for the *lac* promoter, and whose production is controlled by the concentration of extracellular glucose.

The *lac* operon has been pivotal in the development of molecular biology and is currently having an important impact on the development of systems biology. In §2 we present a more detailed description of the *lac* operon regulatory mechanisms, in §3 we review the most significant aspects of the development of molecular biology influenced by the *lac* operon and in §4 we give a comprehensive review of the quantitative studies of the *lac* operon bistable behaviour.

## 2. THE *lac* OPERON REGULATORY MECHANISMS IN DETAIL

In §1, the *lac* operon control system was briefly reviewed. Although that brief review gives a good idea of how this system functions, in reality it is far more complex. In the following paragraphs, a more detailed description of all the regulatory mechanisms in the *lac* operon, including the most recent discoveries, will be given.

The *lac* operon regulatory elements (pictured in figure 2*a*) are distributed along the DNA chain as follows (Reznikoff 1992; Müller-Hill 1998): the *lac* promoter is located between bp −36 (bp stands for base pair, and positions are referred relative to the starting point of gene *lacZ*, bp +1) and bp −7. Operator O1 is 21 bp long and is centred around bp +11. There are two additional operators, denoted O2 and O3, which are, respectively, located at 401 bp downstream and 92 bp upstream from O1. Finally, the activator (CAP)-binding site spans from bp −72 to bp −50.

The *lac* repressor is a homotetramer (consisting of two functional homodimers) of *lacI* polypeptides (Lewis 2005; Wilson *et al.* 2007). Each functional dimer can bind operators O1, O2 and O3. Furthermore, DNA can also fold in such a way that a single repressor binds two operators simultaneously, one per dimer. Each monomer in the *lac* repressor can be bound by an allolactose molecule, inhibiting the capability of the corresponding dimer to bind an operator. This means that free repressors can bind one operator (figure 2*b*) or two of them simultaneously (figure 2*c*), repressors with three free monomers can bind one but not two operators (figure 2*d*), repressors with two free monomers can bind one operator, if the bound monomers belong to the same dimer (figure 2*e*), or none at all, and that repressors with only one free monomer are unable to bind any operator, as are repressors with all four monomers bound by allolactose (Narang 2007).

Deletion experiments have shown that a repressor bound to O1 inhibits transcription initiation, while a repressor bound to either O2 or O3 has almost no effect on the expression of the *lac* operon structural genes. Nevertheless, O2 and O3 do have an indirect effect because the complex formed by a single repressor simultaneously bound to O1 and either O2 or O3 is far more stable than that of a repressor bound only to O1. The consequence of this is that by interacting with the *lac* repressor operator O1 is only capable of decreasing the expression of the operon genes 18 times; when it cooperates with O2, the repression level can be as high as 700-fold; when O1 and O3 act together, they can reduce the operon activity up to 440 times; when all three operators are present, the repression intensity can be as high as 1300-fold (Oehler *et al.* 1990).

The intracellular production of cyclic AMP (cAMP) decreases as the concentration of extracellular glucose increases. cAMP further binds a specific receptor molecule (CRP) to form the so-called CAP complex. Finally, CAP binds a specific DNA site (denoted here as C) upstream from the *lac* promoter, and by doing so it increases the affinity of the mRNA polymerase for this promoter (Reznikoff 1992). This regulatory mechanism is known as catabolite repression.

A novel source of cooperativity has been recently discovered in the *lac* operon: when a CAP complex is bound to site C, it bends DNA locally and increases the probability of the complex in which a repressor simultaneously binds operators O1 and O3 (Kuhlman *et al.* 2007).

The last regulatory mechanism in the *lac* operon is the so-called inducer exclusion. In it, external glucose decreases the efficiency of *lac* permease to transport lactose (Reznikoff 1992), and by doing so negatively affects the induction of the operon genes.

These regulatory mechanisms reviewed above are summarized in figure 3. As we have seen, the activity of the *lac* operon is regulated by extracellular glucose and lactose. While extracellular glucose decreases the operon activity via catabolite repression and inducer exclusion, extracellular lactose increases the operon expression level by deactivating the repressor. Another fact worth noticing is the existence of a positive feedback loop: as more molecules of *lac* permease and

β-galactosidase are produced, there is an elevated lactose uptake flux and an increased lactose metabolism rate; this further increases the production of allolactose and, as a consequence, diminishes the amount of active repressor. This, in turn, increases the operon activity, and thus more *lac* permease and β-galactosidase molecules are produced.

The reader interested in the details of the *lac* operon regulatory mechanisms is referred to the excellent review by Beckwith (1987) and the references there. A good description of the operon regulatory elements and their location on the DNA chain can be found in Reznikoff (1992). The most recent discoveries regarding the cooperativity between CAP-binding site and operator O3 are reported in Kuhlman *et al.* (2007).

## 3. IMPORTANCE OF THE *lac* OPERON

The operon model, developed by Jacob and Monod in their attempt to explain diauxic growth, depicted how genetic mechanisms can control metabolic events in response to environmental stimuli via the coordinated transcription of a set of genes with related function. It literally became a paradigm for gene regulation in prokaryotes, where many more operons have been discovered, and has also influenced the understanding of gene regulation in eukaryotes. Furthermore, not only the *lac* operon as a whole, but also its individual components, such as the *lac* repressor, the three known *lac* operators, the enzyme β-galactosidase and the protein *lac* permease, have influenced the development of molecular biology themselves.

### 3.1. *The* lac *repressor*

A year after Jacob and Monod received the Nobel Prize for their contributions to gene regulation, Müller-Hill and Gilbert isolated the *lac* repressor. This is a protein of 360 amino acids which associates with a homotetramer with 154 520 Da molecular mass (Lewis 2005).

The molecular mechanism of repressing the *lac* operon requires the repressor to be capable of binding both operator DNA and allolactose (or similar inducers). The possibility of competitive binding by these ligands was eliminated by the demonstration that protease digestion selectively cleaves the repressor into two fragments: a tetrameric 'core' (residues 60–360 of the monomer) that retains inducer-binding properties, and a monomeric N-terminal headpiece (amino acids 1–59) capable of binding DNA. To explain the repressor inactivation by allolactose, Monod, Changeux and Jacob proposed that the repressor undergoes a conformational transition in response to bound ligands, and that this alters its ability to bind DNA. They named this phenomenon allostery (Monod *et al.* 1963).

In the early 1970s, several hundred milligrams of the repressor were purified and used for crystallization. Yet, its three-dimensional architecture remained elusive until the early 1990s (Lewis 2005). The three-dimensional structure of the *lac* repressor provided insight into how the repressor may function, as well as the three-dimensional framework for interpreting a huge amount of biochemical and genetic information. Most importantly, when the biochemical and genetic data were viewed in the context of the structure, a detailed molecular model could be constructed to provide a physical basis for the allosteric response, as well as a more detailed understanding of the genetic switch in the *lac* operon.

The allosteric response discovered in the lactose repressor opened a whole new area of research. Allostery has been found in many other proteins and has also been extended to a variety of cellular signalling pathways in all organisms. Notwithstanding, the transcendence of the lactose repressor does not end there. For instance, its monomer has recently been used as a model system for experimental and theoretical explorations of protein-folding mechanisms (Wilson *et al.* 2005).

Those interested in knowing more about the lactose repressor can refer to the review papers by Lewis (2005) and Wilson *et al.* (2007).

### 3.2. *The three* lac *operators*

The primary operator site (O1) for the *lac* operon was sequenced by Gilbert & Maxam (1973) nearly a decade after Jacob and Monod had published their model. In addition to O1, two auxiliary operators (O2 and O3) were identified with sequences similar to those of the primary operator (Reznikoff *et al.* 1974). We now know that tetrameric *lac* repressor is ideally suited to bind two operators simultaneously, creating the so-called 'repression loops' (Reznikoff 1992). DNA looping enhances the repressor affinity for multi-operator sequences, and supercoiling these DNAs yields complexes with remarkable stability. A number of synthetic operator variants have also been constructed and have proved very useful for understanding the molecular mechanisms of repression (Wilson *et al.* 2007).

It was thought for decades that all the signals that control the initiation of bacterial gene transcription are clustered at the 5′ ends of operons, as proposed originally in the models of Jacob, Monod and coworkers. This aspect of their pioneering work is now known to be an oversimplification, as initiation control signals have since been found within, downstream and upstream of the genes regulated by them. Although the phenomenon is not as widespread as in higher cells, its study in bacteria can be, in particular, illuminating. Together with phage lambda switch, the three-operator system of the *lac* operon has been extremely useful (Gralla 1989) in this respect. Moreover, not only have the *lac* operators been helpful to understand the molecular mechanisms of gene regulation, but they have also been employed for other practical purposes. For instance, a technique for *in vivo* visualizing chromosome dynamics using *lac* operator–repressor binding has been proposed (Belmont & Straight 1998).

### 3.3. *The β-galactosidase enzyme*

Few genes have a history of study as long and distinguished as *lacZ*. The *lacZ* gene encodes an open reading frame of 1024 amino acids and is one of the first

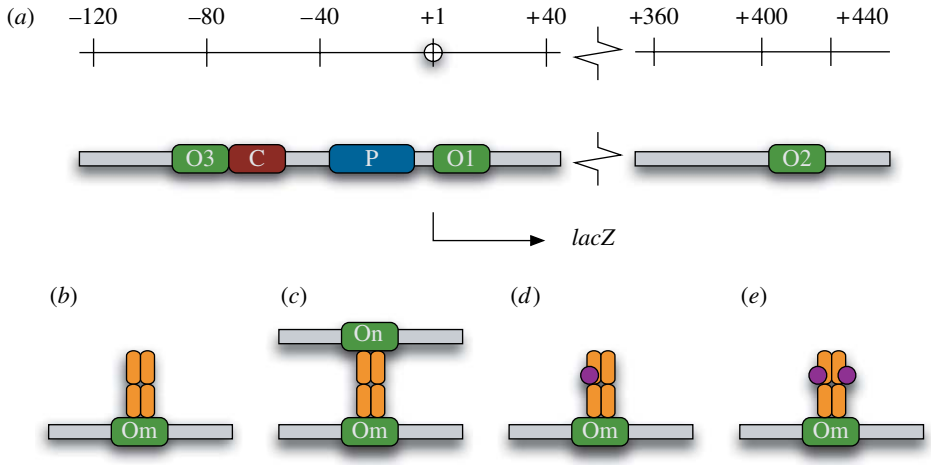


Figure 2. (*a*) Schematic of the regulatory elements located in *lac* operon DNA. P denotes the promoter, O1, O2 and O3 correspond to the three operators (repressor-binding sites), and C is the binding site for the cAMP–CRP complex. The different ways in which a repressor molecule can interact with the operator sites are represented in *b*, *c*, *d* and *e*. Namely, a free repressor molecule (*b*), one with a single subunit bound by allolactose (*d*) or one with the two subunits in the same side bound by allolactose (*e*) can bind a single operator. Moreover, a free repressor molecule can bind two different operators simultaneously (*c*). Figure adapted from Santillán (2008).

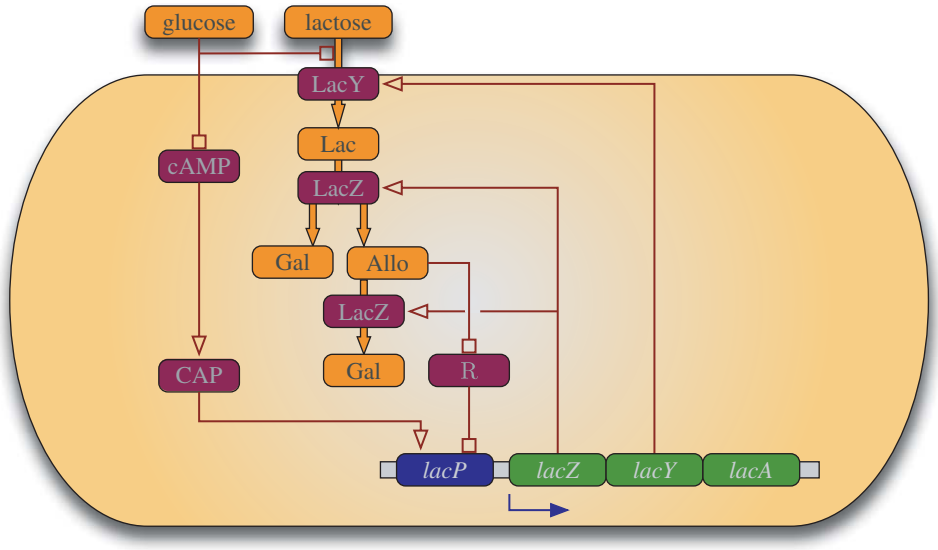


Figure 3. Schematic of the *lac* operon regulatory mechanisms. This operon consists of genes *lacZ*, *lacY* and *lacA*. Protein LacY is a permease that transports external lactose into the cell. Protein LacZ polymerizes into a homotetramer named β-galactosidase. This enzyme transforms internal lactose (Lac) to allolactose (Allo) or to glucose and galactose (Gal). It also converts allolactose to glucose and galactose. Allolactose can bind to the repressor (R) inhibiting it. When not bound by allolactose, R can bind to a specific site upstream of the operon structural genes and thus avoid transcription initiation. External glucose inhibits the production of cAMP that, when bound to protein CRP to form complex CAP, acts as an activator of the *lac* operon. External glucose also inhibits lactose uptake by permease proteins. Figure adapted from Santillán *et al.* (2007).

large genes to be completely sequenced. In *E. coli*, the biologically active β-galactosidase protein exists as a tetramer of four identical subunits and has a molecular weight of approximately 480–500 kDa. The primary enzymatic function of β-galactosidase relevant to its role as a biotechnological tool is to cleave the chemical bond between the anomeric carbon and glycosyl oxygen of appropriate substrates (Serebriiskii & Golemis 2000).

*lacZ* was chosen as the target of a very extensive early analysis, in part owing to specific experimental advantages accompanying work with β-galactosidase. These advantages continue to provide a rationale for using this protein in biotechnological applications today. They are given below (Serebriiskii & Golemis 2000).

— Induction of β-galactosidase synthesis occurs over a large dynamic range (up to 10 000-fold over baseline levels with some inducers). This large range is achievable, in part, because the β-galactosidase protein can be tolerated at extremely high levels in *E. coli*, as well as in many other organisms such as yeasts, *Caenorhabditis elegans*, *Drosophila melanogaster* and mammals. Further, the β-galactosidase protein is readily purified by a number of relatively simple techniques, facilitating *in vitro* analysis of its activity.

— A number of substrates (inducers) for β-galactosidase are either naturally available or very easily chemically synthesized. These enhance the development of

models for β-galactosidase enzymatic activity and also provide a practical tool to finely modulate the expression or dissect catalytic activity of the β-galactosidase protein product.

— β-Galactosidase activity is easily assayed, both *in vivo* and *in vitro*. Assays that have achieved prominence involve the use of colorimetric substrates in which the cleavage of specific β-D-galactopyranoside-coupled aglycone moieties releases coloured dyes. More recently, the panel of available colorimetric substrates has been augmented with fluorescent or chemiluminescent alternative substrates, which further expand sensitivity and applications.

— The β-galactosidase protein is structurally malleable. It consists of three separable functional domains: alpha (α, amino-terminal), beta (β, central) and omega (ω, carboxy-terminal). Independent coexpression of the separated domains of the β-galactosidase protein successfully reconstitutes the activity of the full enzyme. This ability, as well as the additional capacity of β-galactosidase to function enzymatically when expressed as a translational fusion to a varied group of protein or peptide moieties, enables further applications.

These characteristics have allowed the usage of *lacZ* and its product (β-galactosidase) in many scientific and technological applications. Reviewing them all is quite beyond the scope of this paper, but the excellent review by Silhavy & Beckwith (1985) can be consulted. Further, those interested in learning more about the applications of the gene *lacZ* and the protein β-galactosidase can consult Silhavy & Beckwith (1985), Josephy (1996), Serebriiskii & Golemis (2000) and Shuman & Silhavy (2003).

### *3.4. The* lac *permease protein*

Active transporters (pumps) require a cellular energy source (i.e. ATP hydrolysis) to catalyse the transport of charged components against an electrochemical gradient. Depending on their energy source, active transporters are classified as primary or secondary. Secondary transporters, in particular, use the free energy stored in a given electrochemical ion gradient (Abramson *et al.* 2004). LacY is a secondary transporter that couples free energy released from downhill translocation of protons to drive the uphill translocation of galactosides against a concentration gradient.

Lactose permease of *E. coli* (LacY ) is composed of 417 amino acid residues and has 12 helices that transverse the membrane in zigzag fashion, connected by relatively hydrophilic loops with both N and C termini on the cytoplasm side. This protein is encoded by *lacY*, the second structural gene in the *lac* operon. *lacY* was the first gene encoding a membrane transport protein to be cloned into a recombinant plasmid, overexpressed and sequenced (see Kaback 2005 and references therein). This success in the early days of molecular biology opened the study of secondary active transport at the molecular level. Thus, LacY was the first protein of its class to be solubilized and purified in a completely functional state, thereby demonstrating that this single gene product is solely responsible for all the translocation reactions catalysed by the galactoside transport system in *E. coli*. It has also been shown that LacY is both structurally and functionally a monomer in the membrane (Kaback 2005).

Since the discovery of lactose permease, a number of molecular biological, biochemical and biophysical approaches have been used to study this protein. Analysis of this extensive data, and in particular of recent discoveries regarding the LacY structure and the properties of mutants in the irreplaceable residues (Kaback 2005), has led to the formulation of a model for this protein transport mechanism (Abramson *et al.* 2004). Furthermore, comparison of the structures of LacY and other MFS transporters (such as the $P_i$/glycerol-3-phosphate antiporter (GlpT)) has yielded valuable information on the functioning of secondary active transporters in general.

Abramson *et al.* (2004) and Kaback (2005) review the state of knowledge on secondary active transporters, of which the lactose permease is a paradigm.

In summary, we can see from the above considerations that the *lac* system has been extremely important and continues to advance our molecular understanding of genetic control and the relationship between sequence, structure and function.

## 4. QUANTITATIVE EXPERIMENTAL AND THEORETICAL APPROACHES

Two different interpretations of the *lac* operon dynamic performance existed in the beginning. Monod argued that the inducer concentration in the growing environment completely determines the operon induction level. On the other hand, Cohn & Horibata (1959) proposed a more subtle interpretation of their experiments. They suggested that the *lac* system provides an 'experimental example of the Delbrück model' (Delbrück 1949). According to Delbrück, biological systems with identical genotypes may display different behaviours under particular external conditions, due to 'epigenetic' differences that can be transmitted in the cell lineage in the absence of genetic modification. This hypothesis corresponds to a very early formulation of the general principle of phenotypic inheritance.

Novick & Weiner (1957) and Cohn & Horibata (1959) discovered the so-called 'maintenance effect', according to which a single cell may have two alternative states: induced, in which it can metabolize lactose, or uninduced, in which the corresponding genes are switched off and lactose metabolism does not occur. Their experimental protocol was as follows. First, a large amount of inducer was added to the extracellular medium of a culture of uninduced *E. coli* cells; then, the culture was split into two parts: U and I. Part U was immediately diluted, and so the cells remained uninduced; part I was diluted after several minutes, allowing the cells in this subculture to become induced. They further observed that, when induced cells were transferred to a medium with an intermediate 'maintenance concentration' of inducer, they and their progeny remained induced. Similarly, when uninduced cells were transferred to a medium with a 'maintenance' concentration, they and their progeny remained uninduced.

The 'maintenance effect' was interpreted as the consequence of a high permease concentration in induced cells, which would also have high inducer pumping efficiency. This would enable these cells to maintain the induced state and to transmit it to their progeny, even if placed in a medium with a low concentration of inducer. This interpretation accounts for the existence of two distinct phenotypes and provides an explanation of why induced cells placed in media with low inducer concentrations remain indefinitely induced, whereas cells that have never been induced stay uninduced. However, it does not explain what makes the cells switch between alternative states. This switching remained a mystery for a long time and it had to wait for the introduction of the concept of multistability to be fully explained.

Griffith (1968) developed a mathematical model (using ordinary differential equations, ODEs) for a single gene controlled by a positive feedback loop. He found that, under certain conditions, two stable states may be accessible for the system simultaneously. However, Griffith did not use his model to explain the maintenance effect of the *lac* operon. The first models that took into account all the relevant processes to unravel the dynamics of the *lac* operon were by Babloyantz & Sanglier (1972) and Nicolis & Prigogine (1977). Using a mathematical modelling approach, they interpreted the maintenance effect as the biological facet of the physical process of multistability. This model, involving a nonlinear feedback loop, accounted for the main behavioural features of the lactose–operon bistable transition. However, even though the mathematical description of the model required five differential equations (plus one conservation equation), the model did not take into account the detailed information available concerning molecular interactions between the operon components.

### *4.1. The Ozbudak* et al. *experiments*

In the last few years, the interest in the bistable behaviour of the lactose operon has been renewed, and this is in part due to a paper recently published by Ozbudak *et al.* (2004). In it, the authors report the results of a series of ingenious experiments designed to study the bistable *lac* operon response when induced with lactose and the artificial non-metabolizable inducer thiomethylgalactoside (TMG). Ozbudak *et al.* incorporated a single copy of the green fluorescent protein gene (*gfp*) under the control of the *lac* promoter into the chromosome of *E. coli*. The cells also contained a plasmid encoding a red fluorescent reporter (HcRed) under the control of the galactitol (*gat*) promoter. This promoter includes a CRP-binding site, as well as a binding site for the galactitol repressor GatR. However, GatR is absent in *E. coli*. Therefore, transcription at the *gat* promoter, measured by red fluorescence, is a direct measure of CRP-cAMP levels. They further measured the response of single cells, initially in a given state of *lac* expression, to exposure to various combinations of glucose and TMG levels.

Ozbudak *et al.* report that, for a given concentration of extracellular glucose, the *lac* operon is uninduced at low TMG concentrations and fully induced at high TMG concentrations regardless of the cell's history. Between these switching thresholds, however, the system response is hysteretic (history dependent). By measuring the fluorescence of single cells, Ozbudak *et al.* obtained bimodal distributions between the switching thresholds, confirming the existence of bistability. When the experiments were repeated with lactose, instead of TMG, no evidence of bistability was found. The results in this paper not only confirmed bistability in the *lac* operon when induced with TMG, but also provided new and novel quantitative data that raise questions that may be answered via a modelling approach.

### *4.2. A minimal model*

A number of mathematical models have been developed to investigate the dynamic behaviour of the *lac* operon. In §4.3, the characteristics of some of these models will be analysed by contrasting them with the minimal model introduced below.

Let $M$, $E$ and $L$, respectively, denote the intracellular concentrations of mRNA, LacZ polypeptide and lactose. The differential equations governing the dynamics of these variables are

$$\dot{M} = k_{\mathrm{M}}\mathcal{P}_{\mathrm{R}}(L, Ge) - \gamma_{\mathrm{M}}M,$$

$$\dot{E} = k_{\mathrm{E}}M - \gamma_{\mathrm{E}}E,$$

$$\dot{L} = k_{\mathrm{L}}E\mathcal{P}_{\mathrm{L}}(Le, Ge) - E/4\mathcal{M}(L) - \gamma_{\mathrm{L}}L.$$

In the above equations, $k_{\mathrm{M}}$ represents the maximum rate of mRNA production per cell, $k_{\mathrm{E}}$ is the maximum translation initiation rate per mRNA and $k_{\mathrm{E}}$ is the maximum lactose uptake rate per permease. Furthermore, $\gamma_{\mathrm{M}}$, $\gamma_{\mathrm{E}}$ and $\gamma_{\mathrm{L}}$, respectively, represent the degradation plus dilution rates of $M$, $E$ and $L$. The function $\mathcal{P}_{\mathrm{R}}(L,Ge)$ is an increasing function of $L$ (due to the positive regulatory feedback loop described above) and a decreasing function of the extracellular glucose concentration ($Ge$) owing to catabolite repression. $\mathcal{P}_{\mathrm{R}}(L,Ge)$ represents the probability that an mRNA polymerase is bound to the promoter and no repressor is bound to operator O1. The function $\mathcal{P}_{\mathrm{L}}(Le,Ge)$ represents the *lac* permease activity level. It is an increasing function of the external lactose concentration ($Le$) and a decreasing function of $Ge$ due to inducer exclusion. Finally, $\mathcal{M}(L)$ is the rate of lactose metabolism per β-galactosidase.

This simple model structure relies on the assumptions given below.

— The processes governing the dynamics of all the chemical species other than $M$, $E$ and $L$ are assumed to be fast enough to make quasi-steady state approximations to the corresponding dynamic equations.
— Since half of the lactose taken up is directly metabolized into glucose and galactose by β-galactosidase, while the rest is turned into allolactose (which is also later metabolized into glucose and galactose), it can be assumed that the intracellular concentrations of lactose and allolactose are very similar (see Santillán *et al.* 2007 for more details).

— The translation and degradation rates of genes *lacZ* and *lacY* are assumed to be identical. Thus since β-galactosidase (*lac* permease) is a tetramer (monomer), its concentration is assumed to be one-quarter of (equal to) *E*.

It is important to note that the published *lac* operon models differ in the way the function $\mathcal{P}_{\mathrm{R}}(L,Ge)$ is formulated. In some cases, heuristic reasoning is used to propose Hill-type equations for $\mathcal{P}_{\mathrm{R}}(L,Ge)$. Some other models take into account, with different levels of detail, the interactions between the mRNA polymerase and the repressor molecules with the DNA chain to model this function. We discuss these differences in our review of various models in §4.3.

### *4.3. Recent modelling approaches*

Wong *et al.* (1997) developed a 13-dimensional model for the *lac* operon. Besides the structural genes' mRNA and the intracellular lactose concentrations, the variables they consider are repressor mRNA and protein concentrations; β-galactosidase and permease concentrations (each governed by a different differential equation); the internal concentrations of allolactose, cAMP, glucose and phosphorylated glucose; and the external concentrations of glucose and lactose. Their model includes catabolite repression, inducer exclusion, lactose hydrolysis to glucose and galactose, synthesis and degradation of allolactose, and a variable growth rate. Wong *et al.* employed their model to study the diauxic growth of *E. coli* on glucose and lactose. For this, they tested different models for catabolite repression and the phosphorylation of the glucose produced from lactose hydrolysis, and analysed the influence of the model parameters on the two diauxic phases. Besides being quite detailed, this model has the virtue that most of the parameters in it were estimated from experimental data. However, even though Wong *et al.* considered the existence of the three known operator sites, they ignored their cooperative behaviour and incorrectly assumed that a repressor bound to any operator inhibits transcription initiation. Furthermore, Wong *et al.* also took into account the effect of the CAP activator by assuming that it must be bound to its specific site in the DNA regulatory region in order for the polymerase to bind the promoter.

Vilar *et al.* (2003) introduced a simple model of the *lac* operon to illustrate the applicability and limitations of mathematical modelling of the dynamics of cellular networks. In particular, they study the *lac* operon induction dynamics and its relation to bistability. Vilar *et al.* aimed at integrating three different levels of description (molecular, cellular and that of cell population) into a single model, and used it to investigate the system dynamics when an artificial (non-metabolizable) inducer is employed to activate it. In contrast to the minimal model, that of Vilar *et al.* lumps (through a quasi-steady state assumption) the mRNA dynamics into the equations governing the corresponding protein concentrations. It further accounts for the LacZ and LacY dynamics by means of two differential equations, assumes that LacY can be in either a non-functional or a functional state and includes one more equation for this last chemical species. Finally, this model takes into account neither catabolite repression nor inducer exclusion and accounts for the repression mechanism by means of a Hill-type equation.

Yildirim & Mackey (2003) investigated the bistable behaviour of the *lac* operon. For this, they introduced a five-dimensional mathematical model. The model of Yildirim and Mackey accounts for the dynamics of β-galactosidase and *lac* permease by means of two different differential equations, as well as for the dynamics of intracellular lactose and allolactose. This model also takes into account the delays due to transcription and translation processes. The authors paid particular attention to the estimation of the parameters in the model. They further tested their model against two sets of β-galactosidase activity versus time data, and against a set of data on β-galactosidase activity during periodic phosphate feeding. Their analytical and numerical studies indicate that for physiologically realistic values of external lactose and the bacterial growth rate, a regime exists where there may be bistability, and that this corresponds to a cusp bifurcation in the model dynamics. Deficiencies of Yildirim & Mackey's model are that it does not take into account catabolite repression or the inducer exclusion regulatory mechanisms. Furthermore, though they built the repression regulatory function by taking into account the repressor–operator and the polymerase–promoter interactions, they ignored the existence of three operators and considered operator O1 only.

In a later paper, Yildirim *et al.* (2004) attempted to identify as the origin of bistability one of the mechanisms involved in the regulation of the *lac* operon. To do this, they simplified the model presented in Yildirim & Mackey (2003) by ignoring permease dynamics and assuming a constant permease concentration. They numerically and analytically analysed the steady states of the reduced model and showed that it may indeed display bistability, depending on the extracellular lactose concentration and growth rate.

Santillán & Mackey (2004) developed a mathematical model of the *lac* operon, which accounts for all of the known regulatory mechanisms, including catabolite repression and inducer exclusion (both of which depend on external glucose concentrations), as well as the time delays inherent to transcription and translation. With this model, they investigated the influence of catabolite repression and inducer exclusion on the bistable behaviour of the *lac* operon. The model of Santillán & Mackey is six dimensional and the free variables are the *lacZ* and *lacY* mRNA concentrations, the β-galactosidase and *lac* permease concentrations, and the allolactose and cAMP concentrations. It is important to note that Santillán & Mackey's model considers all three known operators and the cooperativity among them, and that all the parameters in it were estimated from experimental data. In particular, they used a thermodynamic approach to model the interactions between the CAP activator, the repressor and the polymerase with their respective binding sites along the DNA chain, as well as the cooperative behaviour of the three known operators.

Van Hoek & Hogeweg (2006) carried out *in silico* simulation of the *lac* operon evolution in bacterial populations. From their results, the parameters that control the expression of the *lac* operon genes evolve in such a way that the system avoids bistability with respect to lactose, but does exhibit bistability with respect to artificial inducers. Thus, they argue from their computational experiments that the wild-type *lac* operon, which regulates lactose metabolism, is not a bistable switch under natural conditions. The model used by van Hoek & Hogeweg contains 10 independent differential equations and is based on the model of Wong *et al.* (1997). This model takes into account all known regulatory mechanisms. However, rather than considering the chemical details, van Hoek & Hogeweg modelled the repressor–DNA and the CAP activator–DNA interactions by means of Hill-type equations.

In a later paper, van Hoek & Hogeweg (2007) modified the *lac* operon model in Van Hoek & Hogeweg (2006) to incorporate stochasticity and study its effects from an evolutionary point of view. Through a mutation-selection process, they evolved the shape of the promoter function, and thus the effective amount of stochasticity. Van Hoek and Hogeweg concluded from their results that noise values for lactose, the natural inducer, are much lower than those for artificial, non-metabolizable inducers, because these artificial inducers experience a stronger positive feedback. They further showed that a high repression rate and hence high stochasticity increase the delay in lactose uptake in a variable environment. From this, the authors concluded that the *lac* operon has evolved such that the impact of stochastic gene expression is minor in its natural environment, but happens to respond with much stronger stochasticity when confronted with artificial inducers.

Santillán *et al.* (2007) investigated the origin of bistability in the *lac* operon. For this, they developed a mathematical model for the regulatory pathway in this system and compared the model predictions with the experimental results of Ozbudak *et al.* (2004). Santillán *et al.* examined the effect of lactose metabolism using this model, and showed that it greatly modifies the bistable region in the external lactose versus external glucose parameter space. The model also predicts that lactose metabolism can cause bistability to disappear for very low external glucose concentrations. The authors further carried out stochastic numerical simulations for several levels of external glucose and lactose and concluded from their results that bistability can help guarantee that *E. coli* consumes glucose and lactose in the most efficient possible way. Namely, the *lac* operon is induced only when there is almost no glucose in the growing medium, but if the external lactose is high, the operon induction level increases abruptly when the levels of glucose in the environment decrease to very low values. Finally, they demonstrated that this behaviour could not be obtained without bistability if the stability of the induced and uninduced states is to be preserved.

In a continuation of the work in Santillán *et al.* (2007), Santillán (2008) improved the mathematical model to account, in a more detailed way, for the interaction of the repressor molecules with the three *lac* operators. Besides, Santillán includes in the model a recently discovered cooperative interaction between the CAP molecule (an activator of the lactose operon) and operator 3, which influences DNA folding. Finally, this new model also includes the growth rate dependence on bacterial energy input rate in the form of transported glucose molecules and of metabolized lactose molecules. A large number of numerical experiments were carried out with the improved model, and the results are discussed along the same lines as in Santillán *et al.* (2007), including a detailed examination of the effect of a variable growth rate on the system dynamics. The models in both Santillán *et al.* (2007) and Santillán (2008) have the same structure as the minimal model above. Furthermore, both models take into account the chemical details of the repressor–DNA and CAP activator–DNA interactions, as well as the cooperativity observed between repressor molecules bound to different operators.

Table 1. Summary of the mathematical models of the *lac* operon that we have reviewed here. (The dimension number refers to the number of dependent variables. A model type can be either deterministic (D) or stochastic (S). Finally, the inducer column states whether model induction with lactose (L) or an artificial inducer (A) is taken into account).

| model | dimensions | type | inducer |
|---|---|---|---|
| Wong *et al.* (1997) | 13 | D | L |
| Vilar *et al.* (2003) | 3 | D | L |
| Yildirim & Mackey (2003) | 5 | D | L |
| Yildirim *et al.* (2004) | 3 | D | L |
| Santillán & Mackey (2004) | 6 | D | L |
| van Hoek & Hogeweg (2006) | 10 | D | A & L |
| van Hoek & Hogeweg (2007) | 10 | S | A & L |
| Santillán *et al.* (2007) | 3 | D & S | A & L |
| Santillán (2008) | 3 | D & S | A & L |

The models here reviewed are summarized in table 1. All of them deal with the bistable behaviour of the *lac* operon. However, most of the ones published prior to Ozbudak *et al.* (2004) only consider the use of lactose as inducer, and they predict that the *lac* operon shows bistability for physiological lactose concentrations. In this sense, the experimental work of Ozbudak *et al.* provided new data and opened new questions suitable for a mathematical modelling approach. One of these questions is why bistability cannot be observed when the *lac* operon natural inducer is employed. The most recent quantitative approaches have made use of Ozbudak *et al.*'s results to develop more accurate models, and two different answers to the above question have been proposed. Van Hoek & Hogeweg (2006, 2007) argue that bistability disappears altogether due to bacteria evolutionary adaptation to a fluctuating environment of glucose and lactose, while Santillán *et al.* (2007) and Santillán (2008) assert that bistability does not disappear but becomes extremely hard to identify with the experimental setup of Ozbudak *et al.* They, furthermore, discuss its significance from an evolutionary perspective. New experiments are needed to resolve this discrepancy.

Most of the models reviewed in this subsection involve ODEs. Given that chemical kinetics is the formalism behind ODE models, they are valid only when the molecule count ($N$) is such that $1/\sqrt{N}$ is small enough.

However, in the *lac* operon, the *lacZ* mRNA degradation rate is so high that the average number of mRNA molecules per bacterium is approximately 0.75, when the operon is fully induced (Santillán 2008). Furthermore, the *lac* repressor LacI is present in only approximately 10 tetramers per cell (Müller-Hill 1998). It follows from this that an essential aspect of modelling the *lac* operon is its stochastic nature. As seen in table 1, only the most recent models (Santillán *et al.* 2007; van Hoek & Hogeweg 2007; Santillán 2008) account for this inherent system stochasticity. Van Hoek & Hogeweg claim that noise has a large effect on the evolution of the *lac* operon: cells evolve such that noise has little effect on the system dynamical behaviour. By contrast, Santillán *et al.* do claim that noise has a large effect on the system dynamical behaviour. It is our belief that a huge amount of work remains to be done on this issue.

## 5. CONCLUSIONS

In this paper, we have presented a description of the regulatory mechanisms in the *lac* operon, taking into consideration the most recent discoveries. The system history has been surveyed as well, emphasizing the discovery of its main elements and the influence they have had on the development of molecular and systems biology. The operon bistable behaviour has also been analysed, including the discovery and origin of this complex phenomenon.

Multistability (of which bistability is the simplest example) corresponds to a true switch between alternate and coexisting steady states, and so allows a graded signal to be turned into a discontinuous evolution of the system along several different possible pathways. Multistability has certain unique properties not shared by other mechanisms of integrative control. These properties may play an essential role in the dynamics of living cells and organisms. Moreover, multistability has been invoked to explain catastrophic events in ecology (Rietkerk *et al.* 2004), mitogen-activated protein kinase (MAPK) cascades in animal cells (Ferrell & Machleder 1998; Bagowski & Ferrell 2001; Bhalla *et al.* 2002), cell cycle regulatory circuits in *Xenopus* and *Saccharomyces cerevisiae* (Cross *et al.* 2002; Pomerening *et al.* 2003), the generation of switch-like biochemical responses (Ferrell & Machleder 1998; Bagowski & Ferrell 2001; Bagowski *et al.* 2003), and the establishment of cell cycle oscillations and mutually exclusive cell cycle phases (Pomerening *et al.* 2003; Sha *et al.* 2003), among other biological phenomena.

Not only was the *lac* operon the first system in which bistability was discovered but, as the literature reviewed in this paper demonstrates, it has been and is still one of the ideal model systems to analyse this complex behaviour. According to the quantitative studies we have reviewed, the *lac* operon has helped to understand the origin, biological implications and subtleness of bistability; it may also help to tackle similar questions in other systems and organisms. For this and other reasons not addressed in this paper, it is our opinion that the lactose operon may be as influential in the development of the nascent field of systems biology as it was in the development of molecular biology.

The different mathematical models of the *lac* operon here reviewed present a good example of what the philosophy of model making is and how mathematical models can influence the development of a given scientific discipline. To discuss this issue we make extensive use of the excellent essay by Rosenblueth & Wiener (1945), which we shall quote a number of times in the forthcoming paragraphs.

According to Rosenblueth and Wiener, models are a central necessity of scientific procedure because no substantial part of the universe is so simple that it can be grasped and controlled without abstraction, abstraction being replacing the part of the universe under consideration by a model of similar but simpler structure. Rosenblueth and Wiener further classify scientific models as either material or formal. In their view, a material model is the representation of a complex system by a system that is assumed simpler and is also assumed to have some properties similar to those selected for study in the original complex system. By contrast, a formal model is a symbolic assertion in logical terms of an idealized relatively simple situation sharing the structural properties of the original factual system. Here we are concerned with formal models of which the mathematical ones are a subset.

Closed boxes (in which a finite number of output variables are causally related to a finite number of input variables, without the knowledge of the detailed mechanisms inside the box) are often employed in formal models. Indeed, according to Rosenblueth and Wiener, all scientific problems begin as closed-box problems (i.e. only a few of the significant variables are recognized), and scientific progress consists of a progressive opening of those boxes. The successive addition of variables gradually leads to more elaborate theoretical models, hence to a hierarchy in these models, from relatively simple, highly abstract ones to more complex, more concrete theoretical structures. The setting up of a simple model for a closed box is based on the assumption that a number of variables are only loosely coupled with the rest of those belonging to the system. As the successive models become progressively more sophisticated, the number of closed regions usually increases, because the process may be compared to the subdivision of an original single box into several smaller shut compartments. Many of these small compartments may be deliberately left closed, because they are considered only functionally, but not structurally, important.

Early closed-box models, such as that of Griffith (1968), acknowledged that a gene subjected to positive feedback regulation can show bistability. However, since having positive feedback does not guarantee bistability, these models could not predict whether the *lac* operon would show this behaviour and, if so, what the responsible mechanisms are. More detailed models were needed to address these questions, and they have been developed as the necessary experimental information is available now. As foreseen by Rosenblueth and Wiener more than 60 years ago, these more detailed models are heterogeneous assemblies of elements, some treated in detail (that is, specifically or structurally) and some treated merely with respect to their overall

performance (that is, generically or functionally). The models by Santillán *et al.* and van Hoek and Hogeweg are good examples. For instance, while the Santillán *et al.* models take into consideration the details of the polymerase–DNA, the repressor–DNA and the activator–DNA interactions, the models of van Hoek and Hogeweg lump them together into a closed box. On the other hand, van Hoek and Hogeweg model with more detail the dynamics of the lactose, allolactose, glucose, cAMP, and ATP intracellular concentrations, and this is reflected in the number of independent equations that the van Hoek and Hogeweg and the Santillán *et al.* models have. Finally, while van Hoek and Hogeweg use an evolutionary modelling approach to study bistability in the *lac* operon, Santillán *et al.* use a more static approach, trying to model the *lac* operon in full detail. To develop more elaborate models, more accurate quantitative data on the system components and their interactions are required. Ideally, a cycle of modelling and experimental efforts shall continue, with one of the outcomes being models progressively more sophisticated and capable of addressing more specific questions. However, there is no point in carrying this process out until its obvious limit. To explain this, we refer once more to Rosenblueth & Wiener (1945), who asserted that as a model becomes more detailed and accurate, it will tend to become identical with the original system. As a limit, it will become the system itself. The ideal model would then be one which agrees with the system in its full complexity and which leaves no closed boxes. However, any one capable of elaborating and comprehending such a model in its entirety would find it unnecessary, because he/she could then grasp the complete system directly as a whole.

We hope that readers of this paper will appreciate that mathematical modelling is a process that constantly evolves as the predictions of the models are iterated against laboratory data. The results of the past three decades in modelling the dynamics of the *lac* operon exemplify this. The reader will, no doubt, also realize that each model has its positive and negative aspects. The level of detail of the model depends on the availability and quality of the data and also on the questions we want to address. The more the detail, the more complicated the model will be. A mathematical analysis might then be hard to undertake and the conclusions may only be based on numerical experiments that many, including us, find less than satisfactory. On the other hand, a simple model may be easier to analyse and a mathematical analysis can give more insights into the dynamical properties or the underlying system; but it may oversimplify and fail to capture some important features of the reality.

The issue of model complexity is intimately tied to the issue of the dimensionality of the parameter space, and this is tied directly to one of the quandaries that face every modeller. The more complex the model, the more parameters must be estimated. It is a virtual truism in mathematical biology that one is almost never able to obtain all of the parameters in a model from the same laboratory setting using the same procedures and techniques and subjects. So, as mathematical model construction is something of an art in itself, the same can be said for parameter estimation. The senior author (M.C.M.) with over 45 years of experience in mathematical biology suggests that the hardest part of the modelling exercise is obtaining decent parameter estimations.

This research was partially supported by Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexico) under grant 55228, the Natural Sciences and Engineering Research Council (NSERC, Canada) and the Mathematics of Information Technology and Complex Systems (MITACS, Canada). We thank the anonymous referees whose comments and criticisms greatly helped us to improve this paper.

## REFERENCES

Abramson, J., Iwata, S. & Kaback, H. R. 2004 Lactose permease as a paradigm for membrane transport proteins (review). *Mol. Membr. Biol.* **21**, 227–236. (doi:10.1080/09687680410001716862)

Babloyantz, A. & Sanglier, M. 1972 Chemical instabilities of "all-or-none" type in beta-galactosidase induction and active transport. *FEBS Lett.* **23**, 364–366. (doi:10.1016/0014-5793(72)80317-X)

Bagowski, C. P. & Ferrell, J. E. 2001 Bistability in the JNK cascade. *Curr. Biol.* **11**, 1176–1182. (doi:10.1016/S0960-9822(01)00330-X)

Bagowski, C., Besser, J., Frey, C. R. & Ferrell Jr, J. E. 2003 The JNK cascade as a biochemical switch in mammalian cells: ultrasensitive and all-or-none responses. *Curr. Biol.* **13**, 315–320. (doi:10.1016/S0960-9822(03)00083-6)

Beckwith, J. 1987 The lactose operon. In *Escherichia coli and Salmonella thyphymurium: cellular and molecular biology*, vol. 2 (eds F. C. Neidhart, J. L. Ingraham, K.B. Low, B. Magasanik & H. E. Umbarger), pp. 1439–1443. Washington, DC: American Society for Microbiology.

Belmont, A. S. & Straight, F. 1998 *In vivo* visualization of chromosomes using *lac* operator–repressor binding. *Trends Cell Biol.* **8**, 121–124. (doi:10.1016/S0962-8924(97)01211-7)

Bhalla, U. S., Ram, P. T. & Iyengar, R. 2002 MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. *Science* **297**, 1018–1023. (doi:10.1126/science.1068873)

Cohn, M. & Horibata, K. 1959 Analysis of the differentiation and of the heterogeneity within a population of *Escherichia coli* undergoing induced beta-galactosidase synthesis. *J. Bacteriol.* **78**, 613–623.

Cross, F. R., Archambault, V., Miller, M. & Klovstad, M. 2002 Testing a mathematical model of the yeast cell cycle. *Mol. Biol. Cell.* **13**, 52–70. (doi:10.1091/mbc.01-05-0265)

Delbrück, M. 1949 Discussion. In *Unités Biologiques douées de Continuité Génétique*, pp. 33–34. Paris, France: Editions du CNRS.

Ferrell Jr, J. E. & Machleder, E. M. 1998 The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* **280**, 895–898. (doi:10.1126/science.280.5365.895)

Gilbert, W. & Maxam, A. 1973 The nucleotide sequence of the *lac* repressor. *Proc. Natl Acad. Sci. USA* **70**, 3581–3584. (doi:10.1073/pnas.70.12.3581)

Gralla, J. D. 1989 Bacterial gene regulation from distant DNA sites. *Cell* **57**, 193–195. (doi:10.1016/0092-8674(89)90955-0)

Griffith, J. S. 1968 Mathematics of cellular control processes. II. Positive feedback to one gene. *J. Theor. Biol.* **20**, 209–216. (doi:10.1016/0022-5193(68)90190-2)

Jacob, F. & Monod, J. 1961*a* Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356.

Jacob, F. & Monod, J. 1961*b* On the regulation of gene activity. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 193–211.

Jacob, F., Perrin, D., Sanchez, C. & Monod, J. 1960 Operon: a group of genes with the expression coordinated by an operator. *C. R. Hebd. Seances Acad. Sci.* **250**, 1727–1729.

Josephy, P. D. 1996 The *Escherichia coli lacZ* reversion mutagenicity assay. *Mutat. Res.* **455**, 71–80.

Kaback, H. R. 2005 Structure and mechanism of the lactose permease. *C. R. Biol.* **328**, 557–567. (doi:10.1016/j.crvi.2005.03.008)

Kuhlman, T., Zhang, Z., Saier, M. H. & Hwa, T. 2007 Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **104**, 6043–6048. (doi:10.1073/pnas.0606717104)

Lendenmann, U., Snozzi, M. & Egli, T. 1996 Kinetics of the simultaneous utilization of sugar mixtures by *Escherichia coli* in continuous culture. *Appl. Environ. Microbiol.* **62**, 1493–1499.

Lewis, M. 2005 The *lac* repressor. *C. R. Biol.* **328**, 521–548. (doi:10.1016/j.crvi.2005.04.004)

Monod, J. 1941 Recherches sur la croissance des cultures bactériennes. PhD thesis, Université de Paris, Paris.

Monod, J., Changeux, J.-P. & Jacob, F. 1963 Allosteric proteins and cellular control systems. *J. Mol. Biol.* **6**, 306–329.

Müller-Hill, B. 1998 The function of auxiliaty operators. *Mol. Microbiol.* **29**, 13–18. (doi:10.1046/j.1365-2958.1998.00870.x)

Narang, A. 2007 Effect of DNA looping on the induction kinetics of the *lac* operon. *J. Theor. Biol.* **247**, 695–712. (doi:10.1016/j.jtbi.2007.03.030)

Nicolis, G. & Prigogine, I. 1977 *Self-organization in non-equilibrium systems. From dissipative structures to order through fluctuations*. New York, NY: Wiley.

Novick, A. & Weiner, M. 1957 Enzyme induction as an all-or-none phenomenon. *Proc. Natl Acad. Sci. USA* **43**, 553–566. (doi:10.1073/pnas.43.7.553)

Oehler, S., Eismann, E. R., Krämer, H. & Müller-Hill, B. 1990 The three operators of *lac* operon cooperate in repression. *EMBO J.* **9**, 973–979.

Ozbudak, E. M., Thattai, M., Lim, H. N., Shraiman, B. I. & van Oudenaarden, A. 2004 Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**, 737–740. (doi:10.1038/nature02298)

Pomerening, J. R., Sontag, E. D. & Ferrell Jr, J. E. 2003 Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nat. Cell. Biol.* **5**, 346–351. (doi:10.1038/ncb954)

Reznikoff, W. S. 1992 The lactose operon-controlling elements: a complex paradigm. *Mol. Microbiol.* **6**, 2419–2422.

Reznikoff, W. S., Winter, R. B. & Hurley, C. K. 1974 The location of the repressor binding sites in the *lac* operon. *Proc. Natl Acad. Sci. USA* **71**, 2314–2348. (doi:10.1073/pnas.71.6.2314)

Rietkerk, M., Dekker, S. F., de Ruiter, P. C. & van de Koppel, J. 2004 Self-organized patchiness and catastrophic shifts in ecosystems. *Science* **305**, 1926–1929. (doi:10.1126/science.1101867)

Rosenblueth, A. & Wiener, N. 1945 The role of models in science. *Phil. Science* **12**, 316–321. (doi:10.1086/286874)

Santillán, M. 2008 Bistable behaviour in a model of the *lac* operon in *Escherichia coli* with variable growth rate. *Biophys. J.* **94**, 2065–2081. (doi:10.1529/biophysj.107.118026)

Santillán, M. & Mackey, M. C. 2004 Influence of catabolite repression and inducer exclusion on the bistable behavior of the *lac* operon. *Biophys. J.* **86**, 1282–1292.

Santillán, M., Mackey, M. C. & Zeron, E. S. 2007 Origin of bistability in the *lac* operon. *Biophys. J.* **92**, 3830–3842. (doi:10.1529/biophysj.106.101717)

Serebriiskii, I. G. & Golemis, E. A. 2000 Uses of *lacZ* to study gene function: evaluation of β-galactosidase assays employed in the yeast two-hybrid system. *Anal. Biochem.* **285**, 1–15. (doi:10.1006/abio.2000.4672)

Sha, W., Moore, J., Chen, K., Lassaletta, A. D., Yi, C. S., Tyson, J. J. & Sible, J. C. 2003 Hysteresis drives cell-cycle transitions in *Xenopus laevis* egg extracts. *Proc. Natl Acad. Sci. USA* **100**, 975–980. (doi:10.1073/pnas.0235349100)

Shuman, H. A. & Silhavy, T. J. 2003 The art and design of genetic screens: *Escherichia coli*. *Nat. Rev. Genet.* **4**, 419–431. (doi:10.1038/nrg1087)

Silhavy, T. J. & Beckwith, R. 1985 Use of *lac* fusions for the study of biological problems. *Microbiol. Rev.* **49**, 398–418.

Van Hoek, M. J. & Hogeweg, P. 2006 *In silico* evolved *lac* operons exhibit bistability for artificial inducers, but not for lactose. *Biophys. J.* **91**, 2833–2843. (doi:10.1529/biophysj.105.077420)

van Hoek, M. J. & Hogeweg, P. 2007 The effect of stochasticity on the *lac* operon: an evolutionary perspective. *PLoS Comput. Biol.* **3**, e111. (doi:10.1371/journal.pcbi.0030111)

Vilar, J. M., Guet, C. C. & Leibler, S. 2003 Modeling network dynamics: the *lac* operon, a case study. *J. Cell Biol.* **161**, 471–476. (doi:10.1083/jcb.200301125)

Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. 2005 Characterization of the folding landscape of monomeric lactose repressor: quantitative comparison of theory and experiment. *Proc. Natl Acad. Sci. USA* **102**, 14 569–14 574. (doi:10.1073/pnas.0505844102)

Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. 2007 The lactose repressor system, paradigms for regulation, allosteric behavior and protein folding. *Cell. Mol. Life Sci.* **64**, 3–16. (doi:10.1007/s00018-006-6296-z)

Wong, P., Gladney, S. & Keasling, J. D. 1997 Mathematical model of the *lac* operon: inducer exclusion, atabolite repression, and diauxic growth on glucose and lactose. *Biotechnol. Prog.* **13**, 132–143. (doi:10.1021/bp970003o)

Yildirim, N. & Mackey, M. C. 2003 Feedback regulation in the lactose operon: a mathematical modeling study and comparison with experimental data. *Biophys. J.* **84**, 2841–2851.

Yildirim, N., Santillán, M., Horike, D. & Mackey, M. C. 2004 Dynamics and bistability in a reduced model of the *lac* operon. *Chaos* **14**, 279–292. (doi:10.1063/1.1689451)