Boolean Models Can Explain Bistability in the *lac* Operon

ALAN VELIZ-CUBA[1] and BRANDILYN STIGLER[2]

ABSTRACT

The *lac* operon in *Escherichia coli* has been studied extensively and is one of the earliest gene systems found to undergo both positive and negative control. The *lac* operon is known to exhibit bistability, in the sense that the operon is either induced or uninduced. Many dynamical models have been proposed to capture this phenomenon. While most are based on complex mathematical formulations, it has been suggested that for other gene systems network topology is sufficient to produce the desired dynamical behavior. We present a Boolean network as a discrete model for the *lac* operon. Our model includes the two main glucose control mechanisms of catabolite repression and inducer exclusion. We show that this Boolean model is capable of predicting the ON and OFF steady states and bistability. Further, we present a reduced model which shows that *lac* mRNA and lactose form the core of the *lac* operon, and that this reduced model exhibits the same dynamics. This work suggests that the key to model qualitative dynamics of gene systems is the topology of the network and Boolean models are well suited for this purpose. Supplementary Material is available online at www.libertonline.com/cmb.

Key words: bistability, Boolean model, catabolite repression, feedback loop, inducer exclusion, lac operon.

1. INTRODUCTION

THE *LAC* OPERON IN THE BACTERIUM *Escherichia coli* has been used as a model system of gene regulation since the landmark work by Jacob and Monod (1961). This system of genes is responsible for the metabolism of lactose in the absence of glucose. Its study has led to numerous insights into sugar metabolism, including how the presence of a substrate could trigger induction of its catabolizing enzyme, yet in the presence of a preferred energy source, namely glucose, the substrate is rendered ineffective. Originally termed the "glucose effect," catabolite repression became known as one of the mechanisms by which glucose regulates the induction of sugar-metabolizing operons. A second glucose-dependent control mechanism was also identified through analysis of the *lac* operon. This mechanism, called inducer exclusion, refers to the suppression of the *lac* permease protein in the presence of extracellular glucose, thereby preventing the transport of lactose into the cell. Early work on the *lac* operon led to the discovery that transcription of an operon's genes is subject to positive (activator proteins initiate transcription) or negative (repressor proteins prevent transcription) control via induction (inducer molecules bind to regulatory proteins

---

[1]Department of Mathematics, University of Nebraska–Lincoln, Lincoln, Nebraska.
[2]Department of Mathematics, Southern Methodist University, Dallas, Texas.

to initiate transcription) or repression (corepressors are used to prevent transcription). The *lac* operon is one of the earliest examples of a inducible system of genes being under both positive and negative control.

There are many formulations modeling the behavior and interaction of the *lac* genes. The first model was proposed by Goodwin two years after the discovery of the *lac* operon (Goodwin, 1963). Since then there has been a steady flow of models following the advances in biological insight of the system, with the majority describing operon induction using artificial nonmetabolizable compounds such as IPTG and TMG (Ozbudak et al., 2004; van Hoek and Hogeweg, 2006, 2007; Santillán et al., 2007). The first model to consider catabolite repression and inducer exclusion with cells grown in both glucose and lactose and lactose used as the inducer was presented by Wong et al. (1997). Their model consisted of up to 13 ordinary differential equations involving 65 parameters. Further, Santillán and coauthors have presented mathematical models and analysis purporting bistability in that the operon is either induced or uninduced (Yildirim et al., 2004; Santillán and Mackey, 2004; Santillán et al., 2007; Santillán, 2008), which has been observed experimentally (Novick and Weiner, 1957; Ozbudak et al., 2004). These findings have given rise to the analogy of the *lac* operon acting as a biological switch (Ozbudak et al., 2004; Halász et al., 2007).

Most mathematical models of the *lac* operon, as well as other genetic systems, are given as systems of differential equations which are appropriate for understanding mechanisms. However, discrete modeling frameworks are receiving more attention in systems biology for their use in analyzing entire state spaces and facilitating global-level predictions. In fact it has been argued that network topology and interaction type (activation/inhibition) are sufficient for capturing dynamics of gene networks (for proofs of concept in other model organisms, see Albert and Othmer, 2003; Mendoza and Xenarios, 2006). A consequence of this observation is that the use of qualitative networks eliminates the need to estimate parameters, thereby reducing model complexity.

Finite dynamical systems (FDSs), which are discrete modeling frameworks suitable for describing qualitative aspects of a network including topology, provide a computational advantage in that entire state spaces can be computed and explored analytically using methods such as generalized logical analysis (Mendoza and Xenarios, 2006; Thomas et al., 1995), in contrast to continuous modeling frameworks. The advantage of having the state space given explicitly is that features such as steady states and limit cycles can be essentially read from the state space. Such predictions, as well as others provided by FDSs, are often easier to verify experimentally than their continuous counterparts, due to their qualitative description; that is, precise experimental conditions are not required to verify qualitative predictions. Boolean networks (BNs), a special class of FDSs, permit an intuitive, yet formal mathematical description of network dynamics by encoding network topology and interaction type in Boolean expressions (Albert and Othmer, 2003; Davidich and Bornholdt, 2008).

Currently few discrete models of the *lac* genes exist. Gianchandani and coauthors introduced a Boolean model of the *lac* operon, excluding the regulatory effects of glucose and lactose and demonstrated that an alternate method for identifying all steady states using a matrix representation of the Boolean model (Gianchandani et al., 2006). Setty et al. (2003) defined a logical function, a specialization of FDSs, for the transcription of the *lac* genes in terms of the proteins regulating the operon, namely CRP and LacI. Although the authors initially aimed to construct a simple Boolean function to mimic the switching behavior of the operon, they discovered that AND-like and OR-like expressions could not reproduce the complexity that the *lac* genes exhibited. Instead they found that a logical function on four states (as opposed to two states—0 and 1) was more biologically relevant. Mayo et al. (2006) tested and showed that this logical function was robust with respect to point mutations, that is, given the formulation in Setty et al. (2003), the operon is still functional after point mutations.

One limitation in both models is that they do not predict bistability. We propose a logical model for the *lac* operon which predicts bistability (when stochasticity is included) and includes the two main control mechanisms of glucose, namely catabolite repression and inducer exclusion. In order to facilitate interpretation, we have added variables so as to present the model as a BN. Therefore, we are able to show that the *lac* operon has only two steady states, corresponding to the operon being either ON or OFF (induced or uninduced).

An important question is to identify the key players in a network, in this case for the purpose of determining the drivers of the dynamics. Aguda and Goryachev (2007) provided a systematic approach for reducing a network derived from literature to a subnetwork which can be thought of as the core of the essential qualitative behavior. Albert and Othmer (2003) showed that the topology and interaction type are the determining factors in producing the steady-state behavior. We corroborate these findings by reducing the Boolean model to one involving only the *lac* genes and lactose. We show that the dynamics of the reduced model matches that of the full model. Our results further support the hypothesis that the topology is the key to dynamical properties.

This article is organized as follows. In Sections 2 and 3, we provide the biological and modeling background. We present a BN as a model for the *lac* operon in Section 4, and we show its network topology, associated dynamics, and bistability analysis. In Section 4.2, we use a novel model reduction algorithm and show the core subnetwork supporting bistability. We close with a discussion of potential future directions in Section 5.

2. THEORY

2.1. Biological background

Here we describe the components and features of the *lac* operon which we include in the model. This description is summarized largely from the material provided in the online book (Griffiths et al., 1999). Additional citations are given as necessary.

The *lac* operon contains three structural genes, *lacZ*, *lacY*, and *lacA*, and is a negative inducible system: the repressor protein LacI prevents transcription of the *lac* genes, and the operon is induced by allolactose, an isomer of lactose. Extracellular lactose is thought to be readily available, but can diffuse into the cell at low concentrations. Once in the cell, lactose can induce the operon, though with lower probability than allolactose. Transcription of the *lac* genes produces a single mRNA, whose translation gives rise to the following proteins: *β*-galactoside permease (LacY), a membrane-bound protein which transports lactose into the cell; *β*-galactosidase (LacZ), an intracellular protein which cleaves lactose into glucose and its stereoisomer galactose, and in a separate reaction converts lactose into allolactose; and *β*-galactoside transacetylase (LacA) which transfers an acetyl group from acetyl-CoA to *β*-galactosides.

Glucose is thought to regulate the *lac* operon through two key mechanisms: catabolite repression and inducer exclusion. In the absence of glucose, the catabolite activator protein CAP (also known as CRP for cAMP receptor protein) forms a complex with cAMP which binds to a site upstream of the *lac* promoter region. Binding of the cAMP-CAP complex makes a conformational change in the DNA, thereby allowing RNA polymerase to bind to the DNA and enhancing transcription of the *lac* genes. Transcription continues until extracellular glucose is available. However, when glucose is abundant, cAMP synthesis is inhibited (Wong et al., 1997) and the repressor protein LacI can bind to the operator region of the operon, preventing transcription of the *lac* genes. Furthermore, the transport of lactose into the cell by permease is inhibited by external glucose, a mechanism referred to as *inducer exclusion*. Finally, the presence of (sufficient amounts of) intracellular glucose shuts off the operon, a phenomenon referred to as *catabolite repression*.

2.2. Modeling background

A BN $F = (F_1, \ldots, F_n)$ on n variables is a tuple of functions defined over the set $\{0, 1\}$ such that for each $i = 1, \ldots, n$, the function F_i determines the next state of variable i and is written in terms of the Boolean operators $\vee$, $\wedge$, $\neg$ (logical OR, AND, and NOT, respectively). The values 0 and 1 are the *states* of the variables.

"Network topology" refers to the connectivity structure of a network and is typically represented as a directed graph. For a BN, a *wiring diagram* is a directed graph on the variables of the system (in this case, mRNA, proteins, and sugars) with edges defined in the following way: there is a directed edge from variable x_i to x_j if the function F_j for x_j depends on x_i. An edge from x_i to x_j has a circle at its head x_j if NOT x_i, denoted by $\neg x_i$, appears in F_j (we consider this edge to correspond to an inhibitory interaction); otherwise, edges have arrows at their heads. We call directed cycles *feedback loops*. The *parity* of a feedback loop (or a path) can be either $+1$ or -1 and is calculated as follows: assign -1 to an edge if it is inhibitory and $+1$ otherwise. The parity of a feedback loop is the product of the signs on the edges of the loop. If the parity of a feedback loop is $+1$, we call the loop *positive*; otherwise, it is *negative*.

"Dynamics" refers to the state transitions of the network as a whole. To generate the dynamics of a BN F on n variables, we evaluate its functions on all possible combinations of 0/1 n-tuples. The dynamics can be viewed as a directed graph, called the *state space* of F. In this graph each node is an n-tuple, called a *state of the system F*; there is a directed edge from a to b if F evaluated at the current state a gives state b; that is, if $F(a) = b$. Hence, b represents that next state of the system. Directed cycles are called *limit cycles*. If length of the cycle is 1, then it is called a *fixed point*. In the context of scientific applications, fixed points are also referred to as *steady states*, which we use in this discourse. We draw the state space using the visualization software DVD (Jarrah et al., 2004).

3. METHODS

The Boolean network model that we present in Section 4 consists of variables and functions, each representing mRNAs, proteins and sugars. We assume that each biomolecule can be either 0 or 1 (absent/inactive or present/active). In order to perform a comparison with an ODE model we considered that extracellular lactose can be in three states (low/0, medium/1, high/2) which we model using the variables L_e and L_{em}; they indicate high concentration and medium concentration of extra cellular lactose, resp. (L_e, L_{em}) = (0, 0), (0, 1), (1, 1) means that the concentration of extracellular lactose is low, medium and high, resp. The same was done for lactose, allolactose and the repressor protein.

The key Boolean variables are labeled as follows:

$M = lac$ mRNA
$P = lac$ permease
$L =$ lactose
C, $R =$ CAP and LacI, resp.
$B = lac\beta$-galactosidase
$A =$ allolactose

where the subscript m indicates medium concentration and e indicates extracellular concentrations.

3.1. Deriving Boolean networks

Here we illustrate how to derive a Boolean function for mRNA based on the information in Section 2.1. The operator $\wedge$ indicates that both variables need to be present, that is, activation occurs by the synergistic action of its regulators; on the other hand, the operator $\vee$ indicates that either variable is enough for activation to occur, that is, activation occurs independently of the other variable. The operator $\neg$ indicates that the variable is a repressor and that it has to be absent for activation to occur. For any variable a, the function F_a determines the value of a after one time unit.

Boolean function for M: To have mRNA being transcribed we need the activator protein (CAP), but also the repressor to be absent. That is, the future Boolean value of M is given by the PRESENCE of C AND the ABSENCE of R AND the ABSENCE of R_m. Hence, the Boolean function for M is $F_M = C \wedge \neg R \wedge \neg R_m$.

Boolean functions for P, B: To have the *lac* proteins being produced we need mRNA. Hence, the Boolean function for P, B are $F_P = F_B = M$.

Boolean function for C: To have CAP in high concentrations we need glucose to be absent. That is, the future Boolean value of C is given by the ABSENCE of G_e. Hence, the Boolean function for C is $F_C = \neg G_e$.

Boolean function for R: The repressor will be active as long as there is no allolactose. That is, the future Boolean value of R is given by the ABSENCE of A AND the ABSENCE of A_m. Hence, the Boolean function for R is $F_R = \neg A \wedge \neg A_m$.

Boolean function for R_m: The repressor will be active at (at least) a medium level as long as there is no allolactose or if the repressor was active to begin with. That is, the future Boolean value of R_m is given by {the ABSENCE of A AND the ABSENCE of A_m} OR the PRESENCE of R. Hence, the Boolean function for R is $F_{Rm} = (\neg A \wedge \neg A_m) \vee R$.

Boolean function for A: To have allolactose in high concentrations we need both β-galactosidase and lactose to be present. That is, the future Boolean value of A is given by the PRESENCE of B AND L. Hence, the Boolean function for A is $F_A = B \wedge L$.

Boolean function for A_m: To have allolactose in medium concentrations we need lactose to be present in at least a medium concentration (it will be transformed to allolactose by a basal number of β-galactosidase molecules which we are not including in the model). That is, the future Boolean value of A_m is given by the PRESENCE of L OR L_m. Hence, the Boolean function for A_m is $F_{Am} = L \vee L_m$.

Boolean function for L: To have lactose in high concentrations we need both permease and extracellular lactose to be present, but also glucose to be absent. That is, the future Boolean value of L is given by the PRESENCE of P AND L_e AND the ABSENCE of G_e. Hence, the Boolean function for L is $F_L = P \wedge L_e \wedge \neg G_e$.

Boolean function for L_m: To have lactose in medium concentrations we need extracellular lactose to be present in at least a medium concentration together with permease, or extracellular lactose to be present in concentrations high enough so that it will enter the cell (by diffusion or a basal number of permease molecules which we are not including in the model); we also need glucose to be absent. That is, the future

Boolean value of L_m is given by the {{PRESENCE of L_{em} AND P} OR the PRESENCE of L_e} AND the ABSENCE of G_e. Hence, the Boolean function for L_m is $F_{L_m} = ((L_{em} \wedge P) \vee L_e) \wedge \neg G_e$.

We use F to denote the Boolean model comprising all of the above functions.

3.2. Bistability

To prove that F exhibits bistability we have to show that for a range of parameters a population of "cells" may exhibit both stable steady states at the same time, corresponding to the *lac* operon being ON and OFF (Santillán and Mackey, 2004); that is, there exists a region of bistability. We will show that if we consider stochasticity in the uptake of the inducer, then bistability can occur.

We performed experiments similar to those in Ozbudak et al. (2004). In order to do this, we considered a stochastic version of our model, and we considered glucose to be absent ($G_e = 0$):

$$\mathcal{F}_M = C \wedge \neg R \wedge \neg R_m$$
$$\mathcal{F}_B = M$$
$$\mathcal{F}_R = \neg A \wedge \neg A_m$$
$$\mathcal{F}_A = L \wedge B$$
$$\mathcal{F}_L = P \wedge L_e$$
$$\mathcal{F}_P = M$$
$$\mathcal{F}_C = 1$$
$$\mathcal{F}_{R_m} = (\neg A \wedge \neg A_m) \vee R$$
$$\mathcal{F}_{A_m} = L \vee L_m$$
$$\mathcal{F}_{L_m} = (L_{em} \wedge P) \vee L_e$$

where extracellular lactose is given as the discretization of a random variable, $\mathcal{L}_e \sim N(\mu, \sigma)$ taken from a normal distribution with μ equal to the mean of {0.6, …, 2.8} incremented by 0.2 and $\sigma = 0.3$. Then, extracellular lactose (L_e) is given as a function of $\mathcal{L}_e$ defined by

$$\text{Extracellular lactose} = \begin{cases} \text{low}/0 & (\text{i.e. } L_e = L_{em} = 0), & \text{if } \mathcal{L}_e < 1 \\ \text{medium}/1 & (\text{i.e. } L_e = 0, L_{em} = 1), & \text{if } 1 \leq \mathcal{L}_e < 2 \\ \text{high}/2 & (\text{i.e. } L_e = L_{em} = 1), & \text{if } 2 \leq \mathcal{L}_e \end{cases}$$

3.3. Reducing Boolean networks

We provide a method to reduce a BN and its corresponding wiring diagram. The idea behind the reduction method is the following: the wiring diagram and Boolean functions should reflect direct regulation and hence nonfunctional edges and variables should be removed; furthermore, vertices can be deleted, without losing important information, by allowing its functionality to be "passed on" to other variables.

1. We simplify the Boolean functions and wiring diagram:
   (a) Reduce Boolean expressions using Boolean algebra. This will delete variables that are nonfunctional.
   (b) Delete edges that do not correspond to Boolean expressions. That is, we delete edges that are nonfunctional.
2. We delete vertices with no self loop, that is, vertices whose Boolean function does not depend on it. Let x_i be a vertex such that f_{x_i} does not depend on x_i.
   (a) For all vertices $x_i \rightarrow y$, that is, for all vertices whose Boolean function depends on x_i, replace the Boolean function for y, $f_y(x_1, \ldots, x_i, \ldots, x_k)$, by $f_y(x_1, \ldots, f_{x_i}, \ldots, x_k)$.
   (b) Replace edges $y \rightarrow x_i \rightarrow z$ by $y \rightarrow z$ and delete x_i (and edges from/to x_i).

4. RESULTS

4.1. Boolean network

In this subsection, we present the BN for the *lac* operon that models gene regulation such as the two main control mechanisms of glucose, namely catabolite repression and inducer exclusion.

The complete BN is given as follows ($\wedge$, $\vee$, and $\neg$ are the logical AND, OR, and NOT operators, respectively):

$$F_M = C \wedge \neg R \wedge \neg R_m$$
$$F_B = M$$
$$F_R = \neg A \wedge \neg A_m$$
$$F_A = L \wedge B$$
$$F_L = P \wedge L_e \wedge \neg G_e$$
$$F_P = M$$
$$F_C = \neg G_e$$
$$F_{R_m} = (\neg A \wedge \neg A_m) \vee R$$
$$F_{A_m} = L \vee L_m$$
$$F_{L_m} = ((L_{em} \wedge P) \vee L_e) \wedge \neg G_e$$

where L_e and G_e (lactose and glucose) are considered as parameters in the model. We use F to refer to the model consisting of this BN.

4.1.1. Network topology.

The network topology for the model F is shown in Figure 1 and is displayed as a wiring diagram (L_e, L_{em} are shown as a single node, as well as R, R_m, A, A_m, and L, L_m); see Section 2.2 for definitions. From the diagram we can identify topological features such as the feedback loops in F. We see that there are at least two positive feedback loops involving M, namely $M \rightarrow P \rightarrow L \rightarrow A \rightarrow R \rightarrow M$ and $M \rightarrow B \rightarrow A \rightarrow R \rightarrow M$. Note that there are no negative feedback loops.

4.1.2. Dynamics.

The dynamics of F can be computed by evaluating the functions on all possible combinations of vectors (M, P, B, C, R, R_m, A, A_m, L, L_m) with 0-1 entries (see Section 2.2 for more details). We say that the operon is OFF when the value of the triple (M, P, B) is (0, 0, 0) and ON when (M, P, B) = (1, 1, 1). For simplicity, we will focus on the steady states.

When glucose is high ($G_e = 1$) we have the unique steady state (**0**, **0**, **0**, 0, 1, 1, 0, 0, 0, 0)[1] corresponding to the *lac* operon being OFF (all initializations eventually reach this steady state). On the other hand if glucose is low ($G_e = 0$) we have the following cases:

1. For low/0 extracellular lactose (that is, $L_e = L_{em} = 0$), there is only one steady state, (**0**, **0**, **0**, 1, 1, 1, 0, 0, 0, 0), that corresponds to the operon being OFF (all of the 2^{10} initializations eventually reach this steady state).
2. For medium/1 extracellular lactose (that is, $L_e = 0$, $L_{em} = 1$), there are two steady states, (**0**, **0**, **0**, 1, 1, 1, 0, 0, 0, 0), that corresponds to the operon being OFF and (**1**, **1**, **1**, 1, 0, 0, 0, 1, 0, 1), that corresponds to the operon being ON (all of the 2^{10} initializations eventually reach one of the steady states). That is, the model is bistable.
3. For high/2 extracellular lactose (that is, $L_e = L_{em} = 1$), there is only one steady state, (**1**, **1**, **1**, 1, 0, 0, 1, 1, 1, 1), that corresponds to the operon being ON (all of the 2^{10} initializations eventually reach this steady state).
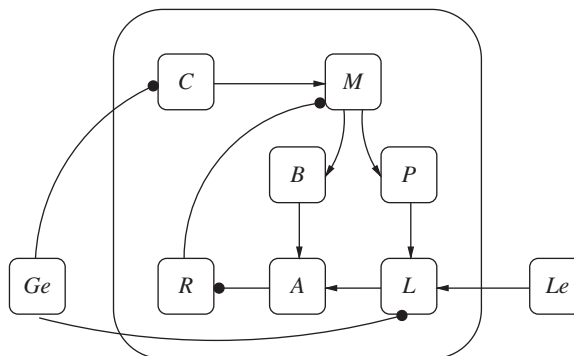
FIG. 1. Wiring diagram for the model F. Edges in this diagram represent interactions between variables. Arrows indicate positive interactions, and circles indicate negative interactions.

---

[1]We make the first three entries bold as these are the entries of interest.

In summary, the model predicts that the *lac* operon is OFF when extracellular glucose is available. When glucose is not available, the model predicts that the operon is OFF, bistable, or ON when extracellular lactose is low, medium, or high, respectively. This is consistent with the reports of bistability, see for example (Santillán, 2008; Ozbudak et al., 2004). The comparison of our Boolean model with ODE models can be illustrated with the bifurcation diagram shown in Figure 2. This figure shows that both models have the same qualitative behavior.

4.1.3. Bistability.   The results of our experiments are in Figure 3 (left). We can see the same qualitative behavior as in the experiments performed in Ozbudak et al. (2004), shown in Figure 3 (right). We can see that as we decrease the concentration of the inducer (top panel) the population of cells start turning their operon OFF; also, as we increase the concentration of the inducer (bottom panel) the population of cells start turning their operon ON.

4.2. Reduced model

An important question is whether the dynamical properties of the model is a direct consequence of its topology and the type of interactions in the model. If network topology and interaction type are sufficient for maintaining certain dynamical properties, we expect a reduced model to have dynamics equivalent to the original model. This would suggest that the qualitative dynamical properties of the *lac* operon are dependent on intrinsic characteristics such as topological features and interaction type and not on extrinsic characteristics such as the specific proteins, sugars involved or the size of the network. This in turn implies that in order to obtain an accurate qualitative description of the *lac* operon we need to focus on using the correct network topology and interaction type. The importance of network topology and interaction type has been proposed by Albert and Othmer (2003), among others, and studied by Aguda and Goryachev (2007). Alternatively, if the dynamics of the model is highly dependent on the size, the model would not be useful to study the dynamics of the biological system: changing the size of the network could significantly alter the predictions.

4.2.1. Reducing the Boolean network.   We provide a reduced network that retains topological features, such as feedback loops, present in the original model. Since many existing models with bistable behavior include glucose and lactose, we will construct a reduced model using M, L, L_e, G_e only. We denote the reduced model by f.

Boolean function for M: To have mRNA being transcribed we need the activator protein (CAP) which will be present if glucose is absent; we also need the repressor to be absent which will happen if there is lactose in the cell. That is, the future Boolean value of M is given by the ABSENCE of G_e AND {the PRESENCE of L_m OR L}. Hence, the Boolean function for M is $f_M = \neg G_e \wedge (L \vee L_m)$.

Boolean function for L: To have lactose in high concentrations we need both permease (which means we need mRNA) and extracellular lactose to be present, but also glucose to be absent. That is, the future Boolean value of L is given by the PRESENCE of P AND L_e AND the ABSENCE of G_e. Hence, the Boolean function for L is $f_L = M \wedge L_e \wedge \neg G_e$.

Boolean function for L_m: To have lactose in medium concentrations we need extracellular lactose to be present in at least a medium concentration together with permease (which means we need mRNA), or extracellular lactose to be present in concentrations high enough so that it will enter the cell; we also need glucose to be absent. That is, the future Boolean value of L_{em} is given by the {{PRESENCE of L_{em} AND
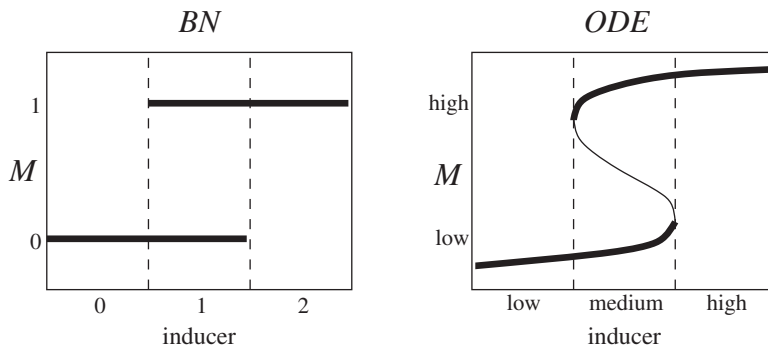
FIG. 2.   Bifurcation diagram of our Boolean network and a typical bifurcation diagram showing bistability of an ODE (two saddle-node bifurcation).

FIG. 3. Results of the bistability experiments performed in the Boolean model F (**left**) and those in Ozbudak et al. (2004) (**right**). Figure in right panel is reprinted by permission from Macmillan Publishers Ltd: *Nature* 427, 737–740, Copyright 2004.
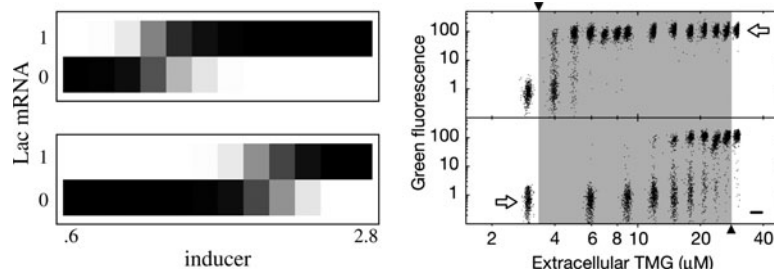
M} OR the PRESENCE of L_e} AND the ABSENCE of G_e. Hence, the Boolean function for L_m is $f_{L_m} = ((L_{em} \wedge M) \vee L_e) \wedge \neg G_e$.

The final reduced model is presented in the next section.

4.2.2. Reduced model. We reduced the Boolean model F by deleting P, B, C, R, A_l, L_l, A (for details, see Supplementary Material online at www.liebertonline.com/cmb) to show that a core subnetwork exists which exhibits bistability. The reduced model, denoted by f, contains the variables M, L, L_e and G_e. The model f is given by

$$f_M = \neg G_e \wedge (L \vee L_m)$$
$$f_L = M \wedge L_e \wedge \neg G_e$$
$$f_{L_m} = ((L_{em} \wedge M) \vee L_e) \wedge \neg G_e$$

where L_e and G_e are treated as parameters.

4.2.3. Network topology. The wiring diagram for the model f is shown in Figure 4 (L_e, L_{em} are shown as a single node, as well as L, L_m). From the reduced model, we can clearly see the roles of the parameters on the core subnetwork whose size is 30% of that of F. We can see that the paths from G_e and L_e to M are still present in the model; also, the signs of these paths have not changed. G_e inhibits M and L (catabolite repression and inducer exclusion, resp.) whereas L_e activates L. Furthermore, we still have a positive feedback loop involving M.

Here we have identified the core of the network to be M and L. From the reduced model, we can clearly see the roles of the parameters on the core subnetwork whose size is about 22% of that of F.

4.2.4. Dynamics. The dynamics of f can be computed by evaluating the functions on all possible combinations of vectors (M, L, L_m) with 0–1 entries. We say that the operon is OFF when the value of **M** is **0** and ON when **M = 1**.

When glucose is high ($G_e = 1$), we have the unique steady state (**0**, 0, 0) corresponding to the lac operon being OFF (all initializations eventually reach this steady state). On the other hand, if glucose is low ($G_e = 0$), we have the following cases:

1. For low/0 extracellular lactose, there is a single steady state, (**0**, 0, 0), that corresponds to the operon being OFF (all initializations eventually reach this steady state).

FIG. 4. Wiring diagram of f.
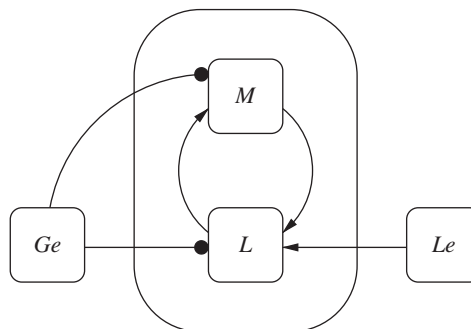
2. For medium/1 extracellular lactose, there are two steady states, $(\mathbf{0}, 0, 0)$, that corresponds to the operon being OFF and $(\mathbf{1}, 0, 1)$, that corresponds to the operon being ON (all initializations eventually reach one of the steady states). That is, the model is bistable.

3. For high/2 extracellular lactose, there is a single steady state, $(\mathbf{1}, 1, 1)$, that corresponds to the operon being ON (all initializations eventually reach this steady state).

In summary, the model predicts that the *lac* operon is OFF when extracellular glucose is available. When glucose is not available, the model predicts that the operon is OFF, bistable or ON when extracellular lactose is low, medium, or high, respectively. Hence, the reduced network has the same dynamics as the full network. We observe that the reduced model has only one positive feedback loop, whereas the model F has several more. Since the reduced model still exhibits bistability, this suggests that it does not depend on the number of positive feedback loops but simply on the existence of such a loop.

4.2.5. Bistability. The region of bistabilty for the model f is similar to that shown in the left panel of Fig. 3 (data not shown for f). This suggests that bistability does not depend on the number of positive feedback loops but simply on the existence of such a loop and stochasticity. Furthermore, because the feedback loop involves only M and L, this suggests that bistability is maintained by stochasticity and the interaction between the operon (represented by M) and lactose.

It is important to notice that the state space of the reduced model has 2^3 states which is about 0.8% of the size of the state space of F, that is, 2^{10} states. We also considered models with different numbers of variables, ranging from four to nine vertices (plus two parameters), and obtained the same results. Hence, we conclude that the dynamical properties of the model are independent of its size.

4.3. Alternative models

Since it has been argued that inducer exclusion is not necessary for bistability (van Hoek and Hogeweg, 2006), we considered alternative models without inducer exclusion and without catabolite repression (Fig. 5) and bistability was also observed.

Model without inducer inclusion:

$F_M = C \wedge \neg R \wedge \neg R_m$
$F_B = M$
$F_R = \neg A \wedge \neg A_m$
$F_A = L \wedge B$
$F_L = P \wedge L_e$
$F_P = M$
$F_C = \neg G_e$
$F_{R_m} = (\neg A \wedge \neg A_m) \vee R$
$F_{A_m} = L \vee L_m$
$F_{L_m} = (L_{em} \wedge P) \vee L_e$
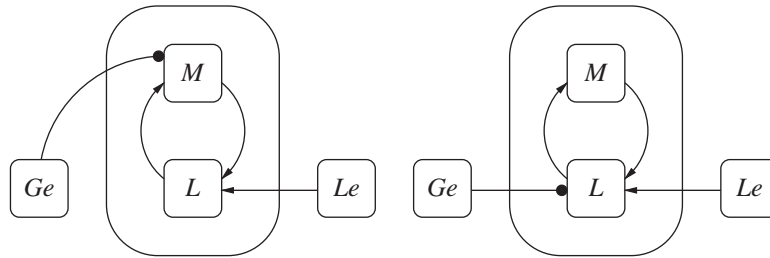
Model without catabolite repression:

$F_M = C \wedge \neg R \wedge \neg R_m$
$F_B = M$
$F_R = \neg A \wedge \neg A_m$
$F_A = L \wedge B$
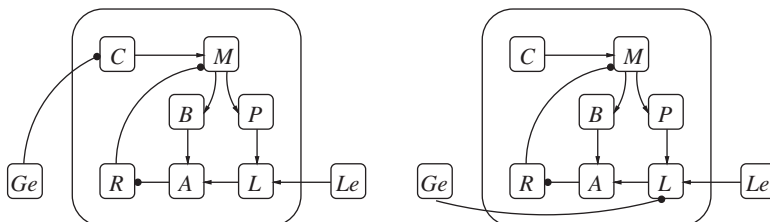
FIG. 5. Wiring diagram of the model without inducer inclusion and without catabolite repression.

FIG. 6. Wiring diagram of the reduced model without inducer inclusion and without catabolite repression.

$$F_L = P \wedge L_e \wedge \neg Ge$$
$$F_P = M$$
$$F_C = 1$$
$$F_{R_m} = (\neg A \wedge \neg A_m) \vee R$$
$$F_{A_m} = L \vee L_m$$
$$F_{L_m} = (L_{em} \wedge P) \vee L_e) \wedge \neg G_e$$

Furthermore, their reduced versions (Fig. 6) also showed bistability.

Reduced model without inducer inclusion:

$$f_M = \neg G_e \wedge (L \vee L_m)$$
$$f_L = M \wedge L_e$$
$$f_{L_m} = (L_{em} \wedge M) \vee L_e$$

Reduced model without catabolite repression:

$$f_M = L \vee L_m$$
$$f_L = M \wedge L_e \wedge \neg G_e$$
$$f_{L_m} = ((L_{em} \wedge M) \vee L_e) \wedge \neg G_e$$

5. CONCLUSION

Many authors have studied the problem of inferring dynamical properties of a system from the network structure, including Jarrah et al. (2010) and Sontag et al. (2008). Furthermore, for special classes of BNs and ODEs it has been proven that the network structure contains all the information needed to determine certain dynamical properties (Halász et al., 2007), such as bistability. On the other hand, it has been claimed that network topology and sign of interactions are more important than quantitative functionality of the components of a system (Albert and Othmer, 2003). To test this hypothesis, we applied the ideas of Albert and Othmer (2003) to lactose metabolism.

The *lac* operon has been studied extensively and is one of the earliest discovered gene systems that undergoes both positive and negative control. While there are numerous continuous models of the *lac* operon, few discrete models exist (Gianchandani et al., 2006; Setty et al., 2003). Gianchandani et al. (2006) introduced a Boolean model of the *lac* operon along with a method for identifying all steady states; however, their model does not include the two main glucose control mechanisms. The model in Setty et al. (2003) is a logical (on four states) function for the *lac* genes written in terms of the regulators CRP and LacI that is capable of accurately predicting induction of the operon based on concentration levels of the regulators. Further, this function has been shown to be robust with respect to point mutations (Mayo et al., 2006). One limitation of this model, however, is that it does not predict bistability, as has been reported and confirmed previously (Novick and Weiner, 1957; Yildirim et al., 2004; Santillán and Mackey, 2004; Ozbudak et al., 2004; Halász et al., 2007; Santillán et al., 2007; Santillán, 2008).

The contributions in this work are twofold. First, we proposed a BN as a discrete model for the *lac* operon and included the glucose control mechanisms of catabolite repression and inducer exclusion. We showed that our model exhibits the ON/OFF switching dynamics and that bistability is also observed, in accordance with the work of Santillán and coauthors (Yildirim et al., 2004; Ozbudak et al., 2004; Santillán

and Mackey, 2004; Santillán et al., 2007; Santillán, 2008). Since it has been argued that inducer exclusion is not necessary for bistability (van Hoek and Hogeweg, 2006), we considered alternative models without inducer exclusion and bistability was also observed.

Second, we presented a reduced model which shows that *lac* mRNA and lactose form the core of the *lac* operon, and that this reduced model also exhibits the same dynamics. This result suggests that the presence of positive feedback loops is a better indicator of dynamics than the individual constituents that comprise the feedback loops. Taken together, this work supports the idea that topology is the driver of dynamics (Albert and Othmer, 2003; Mendoza and Xenarios, 2006).

The use of BNs in modeling has many advantages, such as their mathematical formulation typically being more intuitive than that of differential-equations-based models. BNs are particularly useful in the case where one is interested in qualitatively behavior. For example, our Boolean models exhibit bistability when stochasticity is incorporated, and this depends on topological features rather than on the network itself.

A future work may be to extend this model to a multi-state framework, which has the potential to provide a more refined qualitative description of the *lac* operon. Such a framework may allow for inclusion of other features of the operon, such as multiple promoter and operator regions.

ACKNOWLEDGMENTS

DISCLOSURE STATEMENT

No competing financial interests exist.

REFERENCES

Aguda, B., and Goryachev, A. 2007. From pathways databases to network models of switching behavior. *PLoS Comput. Biol.* 3, 1674–1678.

Albert, R., and Othmer, H. 2003. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster. J. Theor. Biol.* 223, 1–18.

Davidich, M., and Bornholdt, S. 2008. The transition from differential equations to Boolean networks: A case study in simplifying a regulatory network model. *J. Theor. Biol.* 25, 269–277.

Gianchandani, E., Papin, J., Price, N., et al. 2006. Matrix formalism to describe functional states of transcriptional regulatory systems. *PLoS Comput. Biol.* 2, 902–0917.

Goodwin, B. 1963. *Temporal Organization in Cells*. Academic Press, New York.

Griffiths, A., Miller, J., Suzuki, D., et al. 1999. *Introduction to Genetic Analysis*. W.H. Freeman, New York.

Halász, A., Kumar, V., Imieliński, M., et al. 2007. Analysis of lactose metabolism in *E. coli* using reachability analysis of hybrid systems. *IET Systems Biol.* 1, 130–148.

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

Jarrah, A., Laubenbacher, R., Stigler, B., et al. 2004. DVD: Discrete Visualizer of Dynamics. Available at: http://dvd.vbi.vt.edu. Accessed March 1, 2011.

Jarrah, A., Laubenbacher, R., and Veliz-Cuba, A. 2010. The dynamics of conjunctive and disjunctive Boolean network models. *Bull. Math. Bio.* 72, 1425–1447.

Mayo, A., Setty, Y., Shavit, S., et al. 2006. Plasticity of the *cis*-regulatory input function of a gene. *PLoS Biol.* 4, 555–561.

Mendoza, L., and Xenarios, I. 2006. A method for the generation of standardized qualitative dynamical systems of regulatory networks. *Theor. Biol. Med. Model.* 3, 1–18.

Novick, A., and Weiner, M. 1957. Enzyme induction as an all-or-none phenomenon. *Proc. Nat. Acad. Sci. USA* 43, 553–566.

Ozbudak, E., Thattai, M., Lim, H., et al. 2004. Multistability in the lactose utilization network of *Escherichia coli. Nature* 427, 737–740.

Santillán, M. 2008. Bistable behavior in a model of the *lac* operon in *Escherichia coli* with variable growth rate. *Biophys. J.* 94, 2065–2081.

Santillán, M., and Mackey, M. 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., and Zeron, E. 2007. Origin of bistability in the *lac* operon. *Biophys. J.* 92, 3830–3842.

Setty, Y., Mayo, A., Surette, M., et al. 2003. Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. USA* 100, 7702–7707.

Sontag, E., Veliz-Cuba, A., Laubenbacher, R., et al. 2008. The effect of negative feedback loops on the dynamics of Boolean networks. *Biophys. J.* 95, 518–526.

Thomas, R., Thieffry, D., and Kaufman, M. 1995. Dynamical behaviour of biological regulatory networks—I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull. Math. Biol.* 57, 247–276.

van Hoek, M., and Hogeweg, P. 2006. *In silico* evolved *lac* operons exhibit bistability for artificial inducers, but not for lactose. *Biophys. J.* 91, 2833–2843.

van Hoek, M., and Hogeweg, P. 2007. The effect of stochasticity on the *lac* operon: an evolutionary perspective. *PLoS Comput. Biol.* 3, 1071–1082.

Wong, P., Gladney, S., and Keasling, J. 1997. Mathematical model of the *lac* operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose. *Biotechnol. Prog.* 13, 132–143.

Yildirim, N., Santillán, M., Horike, D., et al. 2004. Dynamics and bistability in a reduced model of the *lac* operon. *Chaos* 14, 279–292.

Address correspondence
Dr. Brandilyn Stigler
Department of Mathematics
208 Clements Hall
P.O. Box 750156
Southern Methodist University
Dallas, TX 75275

E-mail: bstigler@smu.edu