

Zhaoqian Steven XIE, Chao TANG

A more robust Boolean model describing inhibitor binding

© Higher Education Press and Springer-Verlag 2008

Abstract From the first application of the Boolean model to the cell cycle regulation network of budding yeast, new regulative pathways have been discovered, particularly in the G1/S transition circuit. This discovery called for finer modeling to study the essential biology, and the resulting outcomes are first introduced in the article. A traditional Boolean network model set up for the new G1/S transition circuit shows that it cannot correctly simulate real biology unless the model parameters are fine tuned. The deficiency is caused by an overly coarse-grained description of the inhibitor binding process, which shall be overcome by a two-vector model proposed whose robustness is surveyed using random perturbations. Simulations show that the proposed two-vector model is much more robust in describing inhibitor binding processes within the Boolean framework.

Keywords Boolean model, cell cycle, biological networks, inhibitor binding

1 Introduction

The rising experimental power of molecular cell biology and the resulting accumulation of data have shifted its focus from single genes and pathways to complex regulation networks. As the complexity of the problems increases, mathematical modeling becomes a necessary tool. With only partial information available for network components and interactions, modeling at coarse-grained

and semi-quantitative levels can be considerably useful and informative [1]. One of the most successful models is the Boolean network [2–6], which is a discrete nonlinear dynamic system where all network nodes take an ON/OFF behavior similar to many actual molecular states. The cell cycle network of budding yeast is a case in Ref. [2], where an eleven-node Boolean network model shows dynamic robustness of the cell cycle regulation.

As more information becomes available for networks, the validity of previous models can be checked for modification and refinement. Take the budding yeast cell cycle regulation network for an example. Its previously unclear mechanisms of activation/inhibition for two transcription factors MBF and SBF are discovered and shown to involve inhibitor proteins Nrm1 [7] and Whi5 [8,9]. Both pathways are located in the G1/S transition circuit, which is part of the regulation network that deals with the transition between G1 and S phases of the cell-cycle. Therefore, an experiment of adding new information to the previous Boolean model of the yeast cell cycle is conducted for possible improvements.

We focus on the G1/S transition circuit alone because it can be isolated as a relatively independent functional module. As we will show, the dynamic trajectory of the old Boolean model for the new network fails to simulate the actual biological process of the G1/S transition. The analysis indicates that it cannot describe the correct timing of the inhibitor removal processes in the network, of which the removal of Sic1-Clb5,6 binding is the most important.

We offer two solutions to this dilemma. First, we try to simulate different timescales of the interactions, which is similar to that in Ref. [10]. This method overcomes the uniform timescale limit of the interactions, and can simulate the actual biological process with some fine-tuning of the timescale parameters. However, this model is far from robust against random perturbations. Therefore, a two-vector model is proposed as the second solution. It tracks all inhibitor binding and removal actions separately to mimic the actual inhibitor binding-removal process. Moreover, the new model still qualifies as a Boolean network model [3], although the state space is doubled.

The dynamic robustness of both models is analyzed to see whether they can still produce correct results when

Received June 10, 2008; accepted June 12, 2008

Zhaoqian Steven XIE
Center for Theoretical Biology, Peking University, Beijing 100871, China
Departments of Physics and Mathematics, the University of Hong Kong, Hong Kong, China

Chao TANG (✉)
Center for Theoretical Biology, Peking University, Beijing 100871, China
Department of Biopharmaceutical Sciences, University of California, San Francisco, CA 94158, USA
E-mail: chao.tang@ucsf.edu

subjected to perturbations, with their parameter strengths also subjected to variation as in earlier research [2]. In addition, the interaction timescales are randomly perturbed similar to previous work [10] to investigate robustness from another point of view. The results show that the two-vector model displays superior dynamic robustness with respect to changes in parameter strengths as well as in timescales.

2 Boolean network model

2.1 G1/S transition circuit

The G1/S transition circuit of budding yeast [11] is part of its cell cycle regulation network in charge of the transition between G1 and S phases. It consists of 8 types of proteins: cyclins Cln3, Cln1,2 and Clb5,6, transcription factors MBF and SBF, and inhibitor proteins Whi5 [8,9], Nrm1 [7], and Sic1 [11,12]. The starting signal of the G1/S transition is the accumulation of the G1 cyclin Cln3. When its concentration reaches a certain threshold, it initiates the activation of the G1/S cyclins Cln1,2 and then the S cyclins Clb5,6 via a combination of inhibitor removal and transcription regulation. After the S cyclins are activated, the cell enters the S phase. Recent experiments identified two new inhibitors, Whi5 [8,9] and Nrm1 [7], which are incorporated into the previous network [2]. The new network is shown in Fig. 1, where the red arrows represent negative regulations (inhibition), the green arrows denote positive ones (activation), and the yellow loop represents self-degradation. The abbreviations on each arrow account for different types of actions.

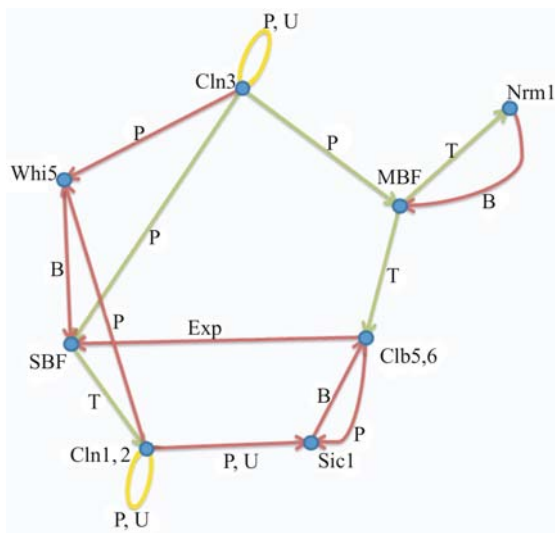


Fig. 1 G1/S transition circuit. The intermolecular interactions involved are identified as follows: P stands for phosphorylation, U ubiquitin-dependent proteolysis, T transcription, B inhibitor binding, Exp exportation out of the nucleus

2.2 Boolean network model

The previous Boolean network model [2] represents the states (concentration or activity) of the 8 kinds of proteins into an 8-vector (the state vector) $C_i(t)$, and their interactions into an 8×8 matrix (the action matrix) A_{ij} that is independent of time:

$$C_i(t) = \begin{cases} 1, & \text{if protein } i \text{ is active at time } t, \\ 0, & \text{if protein } i \text{ is inactive at time } t, \end{cases} \quad (1)$$

$$A_{ij} = \begin{cases} 1, & \text{if protein } i \text{ activates } j, \\ -1, & \text{if protein } i \text{ inhibits } j, \\ 0, & \text{if protein } i \text{ does not affect } j. \end{cases} \quad (2)$$

The dynamics of the system is represented by a step of matrix multiplication plus a step of Boolean decision as follows:

$$C_i(t+1) = \begin{cases} 1, & \text{if } \sum_j A_{ij} C_j(t) > 0, \\ 0, & \text{if } \sum_j A_{ij} C_j(t) < 0, \\ C_i(t), & \text{if } \sum_j A_{ij} C_j(t) = 0. \end{cases} \quad (3)$$

The resulting dynamic trajectory from START (the activation of Cln3) is expected to eventually enter the S phase (which is marked by activation of Clb5,6). However, in the new network, the system stops before the SBF has been activated (Table 1).

Table 1 Dynamics of unadjusted Boolean model

	Cln3	Whi5	MBF	SBF	Nrm1	Cln1,2	Clb5,6	Sic1
step 1	1	1	0	0	0	0	0	1
step 2	0	0	1	0	0	0	0	1
step 3	0	0	1	0	1	0	0	1
step 4	0	0	0	0	1	0	0	1

The system has reached a static point.

2.3 Parameter strengths that need fine tuning to simulate activation of SBF

The result above is far from the actual biological pathway of the G1/S transition. None of the cyclins nor the crucial transcription factor SBF is activated. The SBF activation is not correctly simulated because the activation by Cln3 is canceled out by the inhibition of Whi5, which is a result of taking all the interaction strengths (non-zero entries in the action matrix) to be unity in Eq. (2). Successful activation of the SBF requires either that the activation by Cln3 is stronger than the inhibition by Whi5, i.e., $|A_{14}| > |A_{24}|$ in the action matrix, or that the lifetime of Cln3 has to be longer than 1, which is the value set for all nodes with a self degradation loop in the previous model [2].

If we take $A_{14} = 2$ and keep the other entries in the action matrix unchanged as in Eq. (2), the following dynamics can be obtained, where the SBF is successfully activated but the system stops before the activation of Clb5,6 (see Table 2).

Table 2 Dynamics of Boolean model with adjusted interaction strengths

	Cln3	Whi5	MBF	SBF	Nrm1	Cln1,2	Clb5,6	Sic1
START	1	1	0	0	0	0	0	1
step 2	0	0	1	1	0	0	0	1
step 3	0	0	1	1	1	1	0	1
step 4	0	0	0	1	1	1	0	0
step 5	0	0	0	1	0	1	0	0

The system has reached a static point.

If we keep $A_{14} = 1$ instead but change the lifetime of Cln3 to be 2, the SBF can then also be activated after the deactivation of Whi5. This suggests that modeling different timescales of the interactions can also change the dynamics.

2.4 Timescale differences that need simulation and fine tuning for successful activations of Clb5,6

The foregoing result fails to reproduce the actual G1/S transition process because the MBF shuts itself away via the transcription of its own inhibitor Nrm1 before Sic1, the inhibitor of Clb5,6 is removed via the SBF pathway. This exposes the limitations of the uniform timescale of interactions of the Boolean framework. If we could simulate different timescales of the interactions that range from milliseconds for phosphorylation to tens of minutes for transcription and translation [11], then the Nrm1 expression (which is known to be a slow process) can be delayed after Sic1 has been removed from Clb5,6. In this way Clb5,6 can be successfully activated.

The simulation of different reaction timescales is done with the so-called timer mechanism. Previous work [10] has established a similar mechanism through differential equation considerations, but our conception of the timer is based on discrete structures, where a variable 8-vector $\text{Timer}_i(t)$ tracks how long the action of each type of protein (indexed by i) has been taking place, and a constant 8-vector Threshold_i marks the reaction timescales of each protein. These actions are considered effectively executed only after they have been persisting without interruption for a time longer than the timescale. This mechanism is essentially a low-pass filter, i.e., the signal has to be sustained to affect dynamics. The following is the mathematical formulation:

$$\begin{aligned} \text{If } (C_i(t) > 0) \text{ Then } (\text{Timer}_i(t+1) &= \text{Timer}_i(t) + 1), \\ \text{Else } (\text{Timer}_i(t+1) &= 0), \end{aligned} \quad (4)$$

$$\begin{aligned} \text{If } (\text{Timer}_i(t) > \text{Threshold}_i) \text{ Then } (\text{TurnOnAction}(i)), \\ \text{Else } (\text{TurnOffAction}(i)). \end{aligned} \quad (5)$$

TurnOnAction and TurnOffAction are functions that add or remove (change into zero) the corresponding entries in the action matrix.

If we set $\text{Threshold}_i = 0$ for all i , the model then functions as if no timer mechanism were involved. Thus, the timer mechanism is an extension inclusive of the original model.

This timer mechanism allows us to slow down the MBF pathway or speed up the SBF pathway by adjusting the Threshold_i . If we denote the time threshold of the MBF transcription (Threshold_3) to be 1, and the threshold of all the other interactions to be 0, and set the interaction strength Cln3-SBF activation A_{14} to be stronger than the Whi5-SBF inhibition A_{24} , i.e., $A_{14} > -A_{24} > 0$ (Sect. 2.3), we can then reproduce the original biological pathway of the G1/S transition (Table 3).

Table 3 Dynamics of Boolean model with adjusted parameter strengths and timescales

	Cln3	Whi5	MBF	SBF	Nrm1	Cln1,2	Clb5,6	Sic1
START	1	1	0	0	0	0	0	1
step 2	0	0	1	1	0	0	0	1
step 3	0	0	1	1	0	1	0	1
step 4	0	0	1	1	1	1	0	0
step 5	0	0	0	1	1	1	1	0
step 6	0	0	0	0	1	1	1	0
step 7	0	0	0	0	1	0	1	0

The system has reached a static point.

Similarly, if we make the lifetime of Cln3 longer, thereby assuring the successful activation of the SBF (Sect. 2.3), and let the transcription of the MBF be slower than any other action, we can also reproduce the original biological pathway of the G1/S transition. This suggests that the timescales and the interaction strengths have an interrelated effect on dynamics, although their effects occasionally overlap.

2.5 Fine-tuned Boolean model with timers not robust

The Boolean model requires a significant extension (the timer mechanism) as well as other fine-tuning measures (in parameter strengths or Cln3 lifetime, and in interaction timescales) to simulate the G1/S transition process correctly. Moreover, this fine-tuned model in Sect. 2.4 is not robust. In the yeast cell cycle network studied in Ref. [2], it is found that, given $A_{ij} = a$, if i activates j and $A_{ij} = -b$, if i inhibits j , the entire dynamics of the system remains unchanged so long as $b \geq a > 0$. In the G1/S transition circuit in Sect. 2.3, another requirement $|A_{14}| > |A_{24}|$ for the correct activation of the SBF is presented. Thus, $|A_{14}| - |A_{24}|$ is a bifurcation parameter for

the model in Sect. 2.3. In the model in Sect. 2.4, other bifurcation parameters are introduced together with the timer mechanism. For example, the transcription speed of the MBF Threshold₃ determines whether Clb5,6 can be successfully activated. Threshold₃ is not the only bifurcation parameter introduced with the timer mechanism. If the Threshold vector is such that the inhibition of Cln3 over Whi5 is faster compared with the activation of the SBF, then the requirement $|A_{14}| > |A_{24}|$ in Sect. 2.3 would not be necessary for the correct transition into the S phase.

To get a general idea of the robustness of the fine-tuned Boolean model with the timer mechanism, we randomly perturbed Threshold_{*i*} (by choosing its values from 1 to 10 with equal probability) and observed whether the model could reproduce the correct biological pathway. The results show that 23% of the random timescale choices can complete the G1/S transition if $a = b > 0$ and that 50% of the random timescale choices can complete the G1/S transition if $a > b > 0$ (our model in Sect. 2.4 is one case among these 50%). The percentages are obtained after they have converged within 1% after 30000 random timescale choices. Hence, it is concluded that the fine-tuned Boolean model with the timer mechanism proposed in Sect. 2.4 is not robust.

3 Two-vector model

3.1 A more realistic description of inhibitor binding-removal process

Difficulty in simulating Clb5,6 activation is due to failure to correctly describe the process of activation through inhibitor removal. During the G1/S transition, Clb5,6 are first inhibited directly by Sic1 [11,12] and then indirectly by Nrm1, since Nrm1 shuts down their transcription factor MBF [7]. Of these two inhibitors, Nrm1 is not present at START but Sic1 is. Thus, the uninhibited MBF initiates Clb5,6 transcriptions, but the products Clb5,6 are kept in a dormant state by their abundant inhibitors Sic1. Meanwhile, the MBF helps to produce its own inhibitor Nrm1 and is then shut down to stop further production of Clb5,6. Sic1 is destroyed via the Cln1,2 pathways, and the inhibited Clb5,6 are freed to push the cell into the S phase. The Boolean model fails to simulate this sequence of events because it cannot distinguish between the dormant presence of Clb5,6 and the total absence of Clb5,6, both of which are represented by a Boolean zero in the state vector. However, these two are essentially different because only the former responds to inhibitor removal.

To correctly depict the above-mentioned situation we need to record both the active amount of Clb5,6 and the inactive amount (which are bound by Sic1), or equivalently both the active amount and the total amount. We also need to distinguish between two broad classes of

negative regulations, which are both represented by a Boolean negative in the action matrix: the total annihilation of a certain protein species and the temporary repression that can be lifted at some time later via the inhibitor removal, such as the inhibition of Clb5,6 by Sic1. This idea is formulated mathematically as follows.

Suppose subscripts $i = 1, 2, 3, \dots$ stand for different types of regulative molecules. The total concentration vector \mathbf{T} is defined as $T_i =$ (total amount of molecule i that exist in the cell), and the effective concentration vector \mathbf{E} is defined as $E_i =$ (the amount of molecule i in the cell that is active and not bound by inhibitors). Reversible inhibitor binding processes (the three which involve Sic1, Nrm1, and Whi5, respectively) are taken out of the action matrix and jointly represented by the relation matrix \mathbf{R} .

$$E_i(t) = \begin{cases} 1, & \text{if protein } i \text{ is active at time } t, \\ 0, & \text{if protein } i \text{ is inactive at time } t. \end{cases} \quad (6)$$

$$T_i(t) = \begin{cases} 1, & \text{if protein } i \text{ exists at time } t, \\ 0, & \text{if protein } i \text{ does not exist at time } t. \end{cases} \quad (7)$$

$$A_{ij} = \begin{cases} 1, & \text{if protein } i \text{ activates } j, \\ -1, & \text{if protein } i \text{ inhibits } j, \\ & \text{but not via inhibitor binding,} \\ 0, & \text{if protein } i \text{ does not affect } j. \end{cases} \quad (8)$$

$$R_{ij} = \begin{cases} -1, & \text{if protein } i \text{ is an inhibitor which can bind to } j, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

\mathbf{T} and \mathbf{E} take time-dependent forms $\mathbf{T}(t)$ and $\mathbf{E}(t)$, while \mathbf{A} and \mathbf{R} remain constant in time.

The dynamic rule is as follows:

$$T_i(t+1) = \begin{cases} 1, & \text{if } \sum_j A_{ij} E_j(t) > 0, \\ 0, & \text{if } \sum_j A_{ij} E_j(t) < 0, \\ T_i(t), & \text{if } \sum_j A_{ij} E_j(t) = 0, \end{cases} \quad (10)$$

$$E_i(t+1) = \begin{cases} 1, & \text{if } \sum_j R_{ij} T_j(t+1) > 0, \\ 0, & \text{if } \sum_j R_{ij} T_j(t+1) < 0, \\ 0, & \text{if } \sum_j R_{ij} T_j(t+1) = 0. \end{cases} \quad (11)$$

Note that the two-vector model can be reduced to the original Boolean model by making the action matrix the same as the original one and the relation matrix a unity. Thus, the two-vector model is also an extension inclusive of the original model.

The dynamic trajectory of the two-vector model from START correctly depicts the real biology as shown in Table 4.

Each column stands for a protein species similar to previous research. Each entry of a, b means that for the

Table 4 Dynamics of two-vector model

	Cln3	Whi5	MBF	SBF	Nrm1	Cln1,2	Clb5,6	Sic1
START	1,1	1,1	0,0	0,1	0,0	0,0	0,0	1,1
step 2	0,0	0,0	1,1	1,1	0,0	0,0	0,0	1,1
step 3	0,0	0,0	0,1	1,1	1,1	1,1	0,1	1,1
step 4	0,0	0,0	0,1	1,1	1,1	1,1	1,1	0,0
step 5	0,0	0,0	0,1	0,0	1,1	1,1	1,1	0,0
step 6	0,0	0,0	0,1	0,0	1,1	0,0	1,1	0,0

The system has reached static point.

time specified by the row number, the protein species has an effective concentration a and a total concentration b .

3.2 Dynamically robust two-vector model

Random perturbations to the interaction strengths are applied to the two-vector model of the G1/S transition network. Our results show that the two-vector model is robust against all types of perturbations. It can reproduce the biological pathway of the G1/S transition for all choices of interaction strengths such that $a, b > 0$. For timescale perturbations, we employ a timer mechanism similar to the one in Sect. 2.4, with the concentration vector replaced by the effective concentration vector:

If ($E_i(t) > 0$) Then (Timer $_i(t+1) = \text{Timer}_i(t) + 1$),
 Else (Timer $_i(t+1) = 0$), (12)

If (Timer $_i(t) > \text{Threshold}_i$) Then (TurnOnAction(i)),
 Else (TurnOffAction(i)). (13)

It is found that the two-vector model can simulate the biological pathway of G1/S transition for all randomly perturbed timescales.

It is noted that $a \geq b > 0$, the robustness condition in Ref. [2] is stronger. The robustness condition for the two-vector model $a, b > 0$ is also the weakest possible reasonable condition. The mechanism behind the superior robustness of the two-vector model is as follows: Even if the activation strength of Cln3 on the SBF is small compared to the inhibition of Whi5, the SBF can still be activated later when Whi5 is finally suppressed by Cln3, because the two-vector mechanism allows the SBF to be inhibited at first and reactivated later. This is the case for Sic1-Clb5,6 binding, where the two-vector mechanism compensates for the limitations of uniform timescales of the interactions. It is worth noting that the motivation of the two-vector model has been the inhibitor binding process of Sic1 on Clb5,6 alone. That the model can capture more of the essential mechanism of Whi5-SBF binding is an extra bonus.

These results suggest that the two-vector model, where inhibitor binding-removal process is better depicted, displays superior robustness against perturbations in reaction strengths as well as timescales.

4 Conclusions

Inhibitor binding-removal process is abundant in regulatory networks. Results show that the simplest Boolean model with one bit information for each node may not be able to capture this process faithfully enough to correctly model the underlying biology, at least without fine-tuning some parameters. The two-vector model we proposed is an alternative to resolving this potential problem. When applied to the G1/S transition of yeast cell cycle regulation, the two-vector system is much more robust to perturbations in interaction parameters and timescales, suggesting that it is more reliable and needs much less detailed information about the molecular interactions within the network.

Acknowledgements The authors thank the members of the Center for Theoretical Biology at Peking University for hospitality and helpful discussions. This work was supported by the Chinese Ministry of Science and Technology (2003CB715900) and the National Natural Science Foundation of China (Grant No. 20640120446).

References

- Bornholdt S. Systems biology: Less is more in modeling large genetic networks. *Science*, 2005, 310(5747): 449–451
- Li F T, Long T, Lu Y, et al. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(14): 4781–4786
- Kauffman S A. *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press, 1993
- Albert R, Othmer H G. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *Journal of Theoretical Biology*, 2003, 223(1): 1–18
- Davidich M I, Bornholdt S. Boolean network model predicts cell cycle sequence of fission yeast. *PLoS ONE*, 2008, 3(2): e1672
- Rubinstein A, Gurevich V, Kasulin-Boneh, et al. Faithful modeling of transient expression and its application to elucidating negative feedback regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 2007 104(15): 6241–6246
- de Bruin R A, Kalashnikova T I, Chahwan C, et al. Constraining G1-specific transcription to late G1 phase: the MBF-associated corepressor Nrm1 acts via negative feedback. *Molecular Cell*, 2006, 23(4): 483–496
- de Bruin R A, McDonald W H, Kalashnikova T I, et al. Cln3 activates G1-specific transcription via phosphorylation of the SBF bound repressor Whi5. *Cell*, 2004, 117(7): 887–898
- Costanzo M, Nishikawa J L, Tang X, et al. CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast. *Cell*, 2004, 117(7): 899–913
- Braunewell S, Bornholdt S. Superstability of the yeast cell-cycle dynamics: ensuring causality in the presence of biochemical stochasticity. *Journal of Theoretical Biology*, 2007, 245(4): 638–643
- Morgan D O. *The Cell Cycle: Principles of Control*. New York: Oxford University Press, 2006, Chapter 3
- Verma R, Annan R S, Huddleston M J, et al. Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. *Science*, 1997, 278(5337): 455–460