A genetic algorithm-based boolean delay model of intracellular signal transduction in inflammation

Chu Chun Kang†, Yung Jen Chuang³, Kai Che Tung³, Chun Cheih Chao³, Chuan Yi Tang¹,⁴, Shih Chi Peng¹, David Shan Hill Wong²*

From The Ninth Asia Pacific Bioinformatics Conference (APBC 2011)
Inchon, Korea. 11-14 January 2011

Abstract

Background: Signal transduction is the major mechanism through which cells transmit external stimuli to evoke intracellular biochemical responses. Understanding relationship between external stimuli and corresponding cellular responses, as well as the subsequent effects on downstream genes, is a major challenge in systems biology. Thus, a systematic approach to integrate experimental data and qualitative knowledge to identify the physiological consequences of environmental stimuli is needed.

Results: In present study, we employed a genetic algorithm-based Boolean model to represent NF-κB signaling pathway. We were able to capture feedback and crosstalk characteristics to enhance our understanding on the acute and chronic inflammatory response. Key network components affecting the response dynamics were identified.

Conclusions: We designed an effective algorithm to elucidate the process of immune response using comprehensive knowledge about network structure and limited experimental data on dynamic responses. This approach can potentially be implemented for large-scale analysis on cellular processes and organism behaviors.

Background

Resolving the complex cellular signal transduction is a grand challenge in systems biology. Signal transduction involves cascade of protein-protein interaction and complex feedback loops [1] across proteomic and genomic levels. Models of the dynamics of the combined regulatory networks provide in-depth analysis temporal characteristics of targeted biological process. Furthermore, in silico knockout experiments by these models could help biologists to prioritize target genes of interest and reduce time and cost of real experiments.

Types of dynamic network models include kinetic models [2,3], hidden Markov models [4], and logic-based models [5,6]. Kinetic models based on differential equations have been used to elaborate dynamics on numerous systems [7]. However, they need detailed information about network structure, reaction mechanism and the respective kinetic parameters; which, unfortunately, are not easily obtainable. Hidden Markov model (HMM) is a statistical model in acyclic pathway [8]. The state is hidden, but the outcome dependent on the state is visible. Hence, HMMs are usually used to model known results with unknown process mechanism. Boolean network is a qualitative logic-based model that was introduced in the 1960s [9]. In the past few decades, scientists have frequently used Boolean network to model gene regulatory networks (GRN), apoptosis, metabolic network, immune response and signaling pathways [5,10-12]. Since logic-based or qualitative

* Correspondence: escpeng@gmail.com; dshwong@che.nthu.edu.tw
† Contributed equally
¹Department of Computer Science, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC
³Department of Chemical Engineering, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC
Full list of author information is available at the end of the article

© 2011 Kang et al; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
knowledge of interaction is abundant, the network structure can be easily established. Moreover, only minimal information is required to describe the dynamics of Boolean transfer function, they can be obtained using limited experimental data. Therefore Boolean model is an effective and extendable way of modeling the dynamics of signal transduction.

The transcription factor NF-κB controls various inflammation mediators to orchestrate interwoven cellular responses to inflammatory stimuli such as TNF, IL-1 and TLR4 etc. In this study, Boolean model with time-delay was used to described the NF-κB signaling. The objective is to integrate qualitative information on network interactions from published datasets and dynamic response data in literatures to reveal the regulatory mechanism of infection and inflammation.

**Results and discussion**

**Model**

Figure 1 shows the workflow of building our Boolean model with time delay. First, we generated the Boolean transfer function of our model. Oda et al [16] provided a comprehensive intracellular molecular interaction map. The information was integrated with pathways from KEGG database to obtain a comprehensive network. We focus on the network between three receptors: Interleukin 1 (IL-1), Toll-like receptor 4 (TLR4) and Tumor necrosis factor (TNF) and observable outputs are IKK, IkBa, TNFa. External stimuli considered are TNF and LPS. The network contains a lot of sequential relations which can be compressed using the method shown in Figure 2. In the compression, an intermediate node with only one input and one output is removed; Any branching or meeting nodes are preserved. Figure 3 illustrates the simplified cascade and feedback of signals. There are three feedback routes involving TNF, A20, and IL1. A kernel pathway involving IKK, IkBa and TNFa can be identified through which all the signals have to go through.

Boolean transfer functions were used for each edge. Each transfer function has two dynamic parameters. The delayed activation $\theta$ denotes the duration that the input to a node must turned on before the reception node is turned on. The sustained response $r$ is the time that the output of a node can be sustained once it is turned on. These parameters were obtained by fitting a training data set published in [2,17,18] using GA.

Even with a simplified network and limited number of dynamic parameters, convergence to a set of reasonable parameters was not easy. In order to improve the modeling process, we trained model parameters in kernel pathway first with parameters of the rest of the edges set equal to 1, using wild type data containing measurements of both IKK and NFKB measurements. When kernel's parameters were decided, the parameters of the remaining edges were determined by adding additional data involving A20 knockout, stimulus of various strengths and measurements of either IKK or NFKB only.

Figure 4 shows the comparison between our model outcome and the experimental data in the learning sets. The upper boxes are each active pattern with specific treatment (the description was written on the graph) and compared with real data (the western blot data under corresponded box). The MSE value between our model and data is 0.0919.

**Dynamics implications**

The parameters obtained for the Boolean model were shown in Table 1. The delay of transcription factor NFKB induced components, such as A20, IkBa, IL1 and TNF are longer than other reactions in Table 1. It reveals transcription costs more time than phosphorylation. TNF related pathways will response immediately in our model. In contrast, LPS costs longer delay time to start immune response. The activation of IKK by LPS through TAB1/TAB2/TAK1 is much slower than activation through NIK by TNF.

**Negative regulator**

Because NFkB regulates pro-inflammatory cytokines, such as IL1, TNF, and negative regulator A20 [19], these effectors can generate positive and negative feedback control. In our model, IL1 and TNF were positive feedback. Excessive cytokine production is harmful to the host due to its effects on blood circulation system. Thus, “Endotoxin tolerance” is a critical negative feedback mechanism to protect host from endotoxic shock [20]. In our model, negative regulator A20 and IkBa could suppress NFKB’s activity. Specifically, NFKB induced IkBa (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha) will suppress NFKB activity in a self-regulatory cycle. As shown in Figure 5, inflammation response will be prolonged by IkBa deletion in TNF and LPS stimulation.

Zinc finger protein A20, also called TNFAIT3 (Tumor necrosis factor, alpha-induced protein 3), can also produce RIP- or TRAF2- mediated signal to indirectly block the NFKB activity. We have learned from the literatures that the negative regulator A20 blocked the NFKB activation while protecting the host cell from TNF-mediated apoptosis. To mimic an A20 knockout assay done in the wet-bench experiments, we set the output of the A20 nod to 0 and keep it in off state during simulation as the deletion. The resulting signaling pattern of wild type and A20 mutation in our model were shown in Figure 4. TNF induced IKK(left side of Figure 4.A) active from 5 min to 60 min in wild type and persist with A20 deletion. NF-κB(right side of
Figure 4.A) caused secondary activation when A20 is knocked out. For LPS-induced IKK (left side of Figure 4.B) and NFkB (right side of Figure 4.B), there is no difference between active patterns obtained with wild-type and A20 deletion. This is because in a LPS induced response, the secondary TNF response will be triggered after the transcription of NFkB. Hence A20 is apparently a key component in TNF-induced pathway but with no significant influence in LPS-induced pathway.

**Clinical implication**

Figure 4.A and Figure 6 can help us understand acute and chronic inflammatory response. In acute infection, IKK will active for short span of time to initiate the immune response and return the system to steady state quickly. On the other hand, under chronic inflammatory response...
IKK has an oscillatory profile that can generate greater cytokines production to protect the host cell. With our model, it is hence possible to separate the time course into two phases: the pro-inflammatory and anti-inflammatory phases. As mentioned before, NFkB is a key regulator for inflammation. It up-regulates pro-

In Figure 3, the LPS/IL1/TNF network after simplification is shown. LPS, IL-1, and TNF are inputs. A-20, IL-1, and TNF are three key feedbacks. NF-kB, IKK, and IkBa are observable outputs which can be detected from experiment. An edge between nodes indicates interaction. Arrow implies activation, blunted line denotes inhibition. Within the red box is the kernel pathway of between IKK, IkBa and NFkB.

In Table 1, the components and their Boolean transfer functions are listed. Each component has a Boolean transfer function that contains logic gates and two types of parameters: delay activation (inside bracket) and sustained response (before * sign). Each time step corresponds to 30 seconds in real time.

Table 1 The model obtained by our approach

| Component       | Boolean transfer function |
|-----------------|---------------------------|
| IRAK1(t)        | = 5*IL1(t-1) OR 3*LPS(t-87) |
| TAB1/TAB2/TAK1(t) | = IRAK1(t-13) OR LPS(t-125) |
| NIK(t)          | = TAB1/TAB2/TAK1(t-25) OR TNFR1(t-1) |
| IKK(t)          | = TAB1/TAB2/TAK1(t-100) OR NIK(t-1) |
| IkBa(t)         | = NOT 2*IKK(t-3) OR -62*NFkB(t-96) |
| NFkB(t)         | = NOT IkBa(t-1) |
| A20(t)          | = 7*NFkB(t-107) |
| IL1(t)          | = -62*NFkB(t-373) |
| TNF(t)          | = -62*NFkB(t-69) |
| TNFR1(t)        | = 4*TNF(t-1) AND NOT A20(t-10) |

Each component has a Boolean transfer function that contains logic gates and two types of parameters: delay activation (inside bracket) and sustained response (before * sign). Each time step corresponds to 30 seconds in real time.

(shown in Figure 6), IKK has an oscillatory profile that can generate greater cytokines production to protect the host cell. With our model, it is hence possible to separate the time course into two phases: the pro-inflammatory and anti-inflammatory phases. As mentioned before, NFkB is a key regulator for inflammation. It up-regulates pro-

In Figure 4, the pattern of model simulation compared with experimental data is shown. Simulations profile of our model (Upper boxes) and corresponding original data [2,17,18]. Blue bars represent on and white space means off, respectively. X-axis is the time course (minutes) and y- axis means the activities of target in our model in distinct conditions: (A) IKK (left) and NFkB’s (right) activities induced by TNF 45 minutes treatment with wild type and A20 mutant. (B) IKK (left) and NFkB’s (right) activities with wild type and A20 knockout condition of LPS 45 minutes treatment. (C) NFkB’s activities induced by transient 15 minutes IL-1 in wild type and IkBa mutant. (D) IkBa’s activities with TNF (left) and LPS (right) 45 minutes stimuli.
inflammatory cytokines such as TNF and IL-1 as well as trigger negative regulator such as A20 and IkBa to suppress the IKK activities. Over a short span of 60 minutes, IKK’s activity will decrease by the repression of IkBa, this process can thus be defined as the anti-inflammatory phase. Further, the time profile of TNF-induced IKK can cause secondary activation around 6 hour.

Conclusions

In this work, a dynamic Boolean model was generated by integrating and comprehensive qualitative knowledge about network structure and fitting a minimal amount of dynamic response data. The model is capable of capturing feedback and crosstalk dynamics between diverse signaling pathways. Using this model, mechanisms of and factors affecting periodic pro-inflammatory and anti-inflammatory responses can be elucidated.

The proposed approach integrated intracellular and intercellular process. Hence it is possible for us to use this approach to develop system models for host defense against the shock from environmental or pathogen stimuli and predict the inflammatory response. Such a model will potentially be able to provide insight to a feedback treatment scheme for clinical therapy.
Methods

Boolean transfer function

The Boolean model with time delay can be described mathematically by the graph \( G = (V, E) \), in which \( V = \{x_1, \ldots, x_N\} \) is a set of nodes and \( E = \{e_{ij}\} \) is a set of edges, with \( e_{ij} \) equals 1 if there is a linking edge starting from the \( j \)-th node to the \( i \)-th node and \( e_{ij} \) equals 0 otherwise [21-23]. The transfer function that determines activation of the node \( x_i \) at time \( t \) is given by:

\[
x_i(t) = \sum_{e_{ij}} e_{ij} h \left( \left\{ x_j(t-\theta_{ij}) \right\}, \theta_{ij} \right)
\]

The effect of accumulated activation is given by

\[
h \left( \left\{ x_j(t-\theta_{ij}) \right\}, \theta_{ij} \right) = \begin{cases} 1 & x_j(t-\theta_{ij}) = 1 \\ 0 & \text{otherwise} \end{cases}
\]

In other words, the activation of the \( i \)-th node by \( j \)-th node is on only if the \( j \)-th node has been on continuously for a period of \( \theta_{ij} \). The effect of sustained and delayed response can be described by the following pseudocode:

- when \( h(\{x_j(t)\}, \theta_{ij}) = 1 \)
  - if \( r_{ij} \geq 1 \)
    - \( \tau_{ij} = \tau_{ij} + r_{ij} - 1 \)
    - \( g_{ij} = 1 \)
  - else
    - \( \tau_{ij} = \tau_{ij} + 1 \)
    - if \( r_{ij} \geq |r_{ij}| \)
      - \( g_{ij} = 1 \)

Figure 8 Experimental data processing normalization. The data processing by ImageJ software. The original data is from [2]. We subtracted background and separated profile bipartitely by Max entropy threshold approach. The binary data profile comes after data processing.

Figure 9 Flowchart of the Genetic algorithm workflow. The first population was generated randomly. The fitness criterion is defined as minimizing mean squared error (MSE) between model predictions and experimental results. Roulette wheel selection is used to select candidates for parents in crossover. A fixed mutation rate of 2% was used to prevent premature convergence. The Elitism algorithm was used to determine the survival of parents and children in the new generation.
This article has been published as part of BMC Bioinformatics Volume 12 Supplement 1: Selected articles from the Ninth Asia Pacific Bioinformatics Conference (APBC 2011). The full contents of the supplement are available online at http://www.biomedcentral.com/1471-2105/12?issue=S1.

Author details
1Department of Computer Science, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC. 2Department of Chemical Engineering, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC. 3Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC. 4Department of Computer Science and Information Engineering, Providence University, Taichung, 43301 Taiwan, ROC.

Authors’ contributions
CCK and SCP developed the method, performed the analyses and wrote the manuscript. YJC, CCC and KCT interpreted the data and wrote the manuscript. CYT advised on method design. DSHW investigated the principle and wrote the manuscript.

Competing interests
The authors declare that they have no competing interests.

Published: 15 February 2011

References
1. Neil A, Campbell JBR. Biology. 6 2002.
2. Werner SL, Kearns JD, Zadoverzhnya V, Lynch C, O’dea E, Boldin MP, Ma A, Baltimore D, Hoffmann A: Encoding NF-κappaB temporal control in response to TNF: distinct roles for the negative regulators IkappaBalpha and A20. Genes Dev 2008, 22:2093-2101.
3. Pettigrew MF, Rosat H: Modeling signal transduction networks: a comparison of two stochastic kinetic simulation algorithms. J Chem Phys 2005, 123:114707.
4. Mark de Been CF, Roy Moezelaar, Tjakko Abee, Siezen R: Comparative analysis of two-component signal transduction systems of Bacillus cereus, Bacillus thuringiensis and Bacillus anthracis. Microbiology 2006, 152:3035-3048.
5. Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, Sorger PK: Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. Mol Syst Biol 2009, 5:331.
6. Julio Saez-Rodriguez LS, Jonathan ALindquist, Rebecca Hemenway, Ursula Bommhardt, Boerge Arndt, Utz-Uwe Haus, Robert Weismantel, Ernst DGilles, Steffen Klamt, Burkhard Schraven. A Logical Model Provides Insights into T Cell Receptor Signaling. PLoS Comput Biol 2007, 3:e163.
7. Arii I, Cattaneo A, Rosato V: Parameter estimate of signal transduction pathways. BMC Neurosci 2006, 7(Suppl 1):S6.
8. Geiger D, Meek C, Wester V: Speeding up HMM algorithms for genetic linkage analysis via chain reductions of the state space. Bioinformatics 2009, 25:i96-203.
9. Kaufman SA: Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of Theoretical Biology 1969, 22:437-467.
10. Kwon YK, Choi SS, Cho KH: Investigations into the relationship between feedback loops and functional importance of a signal transduction network based on Boolean network modeling. BMC Bioinformatics 2007, 8:384.
11. Mai Z, Liu H: Boolean network-based analysis of the apoptosis network: Irreversible apoptosis and stable surviving. Journal of Theoretical Biology 2009, 259:760-769.
12. Thakar J, Pillene M, Kiranmpesvara G, Harvill ET, Albert R: Modeling Systems-Level Regulation of Host Immune Responses. PLoS Comput Biol 2007, 3:e109.
13. Holland JH: Genetic Algorithms and the Optimal Allocation of Trials. SIAM Journal on Computing 1973, 2:88-105.
14. Zheng Y, Yeh CW, Yang CD, Jang SS, Chu IM: On the local optimal solutions of metabolic regulatory networks using information guided genetic algorithm approach and clustering analysis. J Biotechnol 2007, 131:159-167.
15. Wu FX, Poirier GG, Zhang WJ: Inferring gene regulatory networks with time delays using a genetic algorithm. Syst Biol (Stevenage) 2005, 15267-74.

Table 2 Settings in genetic algorithm

| Setting                      | Value                  |
|------------------------------|------------------------|
| Chromosome representation    | $\theta_{ij}$          |
| Generation number            | 800                    |
| Population size              | 1000                   |
| Crossover rates              | 1                      |
| Mutation rates               | 0.02                   |

$\tau_{ij} = \tau_{ij} - 1$

end

Thus if $r_{ij} \geq 1$, the activation of $x_i$ by $x_j$ will be sustained; alternatively the activation of $x_j$ by $x_i$ will be delayed.

The function $f_1$ is a series of logical gates connecting input nodes to output nodes. The relations of these logical gates to biological processes are shown in Figure 7.

Data processing

To generate on-off response of the observed nodes, we collected the Western blot experimental data in [2,17,18] and employed processing steps in ImageJ as shown in Figure 8. First the background is subtracted. Then the maximum entropy threshold approach was used to filter the data. Finally, a binary data profile is obtained.

Model fitting by genetic algorithm

We implemented genetic algorithm to optimize model by MATLAB. Figure 9 shows work flow of the genetic algorithm utilized in this study. We generated first population randomly. The fitness is defined as mean squared error (MSE) between model predictions and experimental results:

$$Fitness = \frac{1}{M} \sum_{t} \sum_{i=IKK,RRA,NFkB} \left( x_{i}^{model}(t) - x_{i}^{exp}(t) \right)^2$$

The solution will be achieved by minimizing the fitness function through genetic operations. First the parameters in the Boolean transfer function $\theta_{ij}$ and $r_{ij}$ of the active links are transformed into chromosomes representation. Roulette wheel selection is used to select candidates for parents in crossover. A fixed mutation rate of 2% was used to prevent premature convergence. The Elitism algorithm [24] was used to determine the survival of parents and children in the new generation. A population size of 1000 was used. The GA is terminated at 800 generations. Parameter settings of the GA algorithm were in Table 2.
16. Kanae Oda TK, Yukiko Matsuoka, Akira Funahashi, Masaki Muramatsu, Hiroaki Kitano: Molecular Interaction Map of a Macrophage. Book Molecular Interaction Map of a Macrophage 2004, 2.
17. Werner SL, Barken D, Hoffmann A: Stimulus Specificity of Gene Expression Programs Determined by Temporal Control of IKK Activity. Science 2005, 309:1857-1861.
18. Shih VF-S, Keams JD, Basak S, Savinova OV, Ghosh G, Hoffmann A: Kinetic control of negative feedback regulators of NF-κB/RelA determines their pathogen- and cytokine-receptor signaling specificity. Proc Natl Acad Sci U S A 2009, 106:9619-9624.
19. Renner F, Schmitz ML: Autoregulatory feedback loops terminating the NF-kappaB response. Trends Biochem Sci 2009, 34:128-135.
20. Bell E: Mediating endotoxin tolerance. Nat Rev Immunol 2004, 4:750-750.
21. Akutsu T, Miyano S, Kuhara S: Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. Pac Symp Biocomput 1999, 17-28.
22. Fang-Xiang W, Kusalik AJ, Wen-Jun Z: A genetic algorithm for inferring time delays in gene regulatory networks. Computational Systems Bioinformatics Conference, 2004 CSB 2004 Proceedings 2004 IEEE 2004, 610-611.
23. Assmann SM, Albert R: Discrete dynamic modelling with asynchronous update, or how to model complex systems in the absence of quantitative information. Methods Mol Biol 2009, 553:207-225.
24. Grefenstette J: Optimization of control parameters for genetic algorithms. IEEE Trans Syst Man Cybern 1986, 16(1):122-128.

doi:10.1186/1471-2105-12-S1-S17
Cite this article as: Kang et al.: A genetic algorithm-based boolean delay model of intracellular signal transduction in inflammation. BMC Bioinformatics 2011 12(Suppl 1):S17.