

## Modeling Gene Regulation in Graded Hypoxia Using Continuous-Time Markov Chains

Putu Indah Ciptayani<sup>1</sup>, Reza Pulungan<sup>2</sup>

<sup>1</sup>Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Gedung BJ Bukit Jimbaran, Badung, Indonesia, E-mail: putu.indah@cs.unud.ac.id

<sup>2</sup>Jurusan Ilmu Komputer dan Elektronika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada, FMIPA UGM Gedung Selatan, Sekip Unit III, Yogyakarta, Indonesia, E-mail: pulungan@ugm.ac.id

---

### Article Info

#### Article history:

Received Jul. 6<sup>th</sup>, 2013

Accepted Jul. 21<sup>st</sup>, 2013

---

#### Keyword:

Stochastic

Gene regulation

Graded hypoxia

Continuous-time Markov chain

Stochastic hybrid model

---

### ABSTRACT

*Hypoxia inducible factor* (HIF) is the main protein in hypoxia pathway. The response of HIF to changes of oxygen pressure is regulated by 2 oxygen sensors, *prolyl hydroxylase* (PHD) and *factor inhibiting HIF* (FIH). Studies have shown that biochemical reactions at molecular level actually exhibit stochastic and random behaviors. Modeling biochemical reactions using purely deterministic method, therefore, ignore these characteristics. Hence, we use stochastic modeling using CTMC to model this regulation. Nevertheless, the use of pure CTMC on complex biochemical reaction networks, such as hypoxia, results in an infeasible computation time and typically requires very large memory. Therefore, we use a numerical hybrid method that combines pure CTMC and deterministic methods. The purpose is to reduce time complexity and to obtain a better accuracy than deterministic method. Using this model, we can observe that an increase of oxygen pressure results in a decrease in the amount of HIF and that oxygen sensor FIH only inhibits C-TAD activity. The model is also able to classify 84% genes that were observed.

Copyright © 2013 Institute of Advanced Engineering and Science.  
All rights reserved.

---

### 1. INTRODUCTION

Hypoxia is a pathological condition in which the whole or a region of the body is deprived of adequate oxygen supply. Hypoxia is a characteristic of many diseases such as myocardial infarction, stroke and cancer [1]. Hypoxia is also found in tumors because the rapid growth of tumor cells creates hypoxic conditions [2]. Studies, such as [3], concluded that the main protein pathway in hypoxia can be a novel therapeutic target in reducing local tumor growth. The main protein in hypoxia pathway is hypoxia inducible factor (HIF) [4]. HIF has two transcriptional activation domains: N-TAD and C-TAD. The response of HIF to oxygen is caused by two oxygen sensor: prolyl hydroxylase domain (PHD) and factor inhibiting HIF (FIH). In the presence of oxygen, PHD hydroxylates HIF into two prolyl residues, which is a signal for interaction with the von Hippel-Lindau (VHL); while FIH hydroxylates HIF into an asparagynil residue, which causes inhibition of C-TAD activity. In hypoxic condition, these two sensors are inactive and HIF is activated and stabilized by its two TADs and then bound to an hypoxia response element (HRE).

Studies of gene regulation in hypoxia have been carried out by many researchers, such as [1], [5] and [6]. Biological experiments were performed in [5], while in [1] the gene regulation was analyzed through a mathematical model that is based on differential equations. In both studies, it is shown that FIH only inhibits the activity of C-TAD, not N-TAD. The study in [1] also showed that an increase in oxygen pressure causes a decrease in the amount of HIF. Yu et al. in [6] studied the switch-like behavior in gene regulation using extreme pathway analysis (EPA).

In this paper, we report our investigation on the expression of HIF in oxygen gradient and the role of FIH using stochastic modeling. Studies reported in [7] and [8] indicated that biochemical reactions at molecular level actually exhibit stochastic and random behaviors, not only statistical behaviors. Mere deterministic models, namely those models that purely use differential equations, cannot adequately capture

these behaviors. They also stated that mathematical modeling that simply uses differential equations ignores the discrete characteristics of molecules involved and the probabilistic nature of microscopic molecular collisions. To model these characteristics, stochasticity is needed. In this paper, we propose to use continuous-time Markov chain to model the biochemical reactions along the pathway. In this case, a state in the CTMC represents the number of each molecule-type, called a species, involved in the reaction at any given time, while transitions in the CTMC represent the chemical reactions that change the configuration of all species from time to time. The evolution of a certain state in the model is given by a system of linear ordinary differential equations called chemical master equations (CME) which equate the change probability of the state and the sum of all inflow and outflow probabilities.

Modeling chemical reaction networks by CTMCs usually results in large state spaces. Hence, efficient algorithms in transient and steady-state analysis are needed to reduce computation times. A simple biochemical reaction called enzyme-catalyzed substrate conversion, involving 4 species and 3 reactions, was modeled using CTMCs in [8] and [9]. Wolf in [8] used stochastic automata networks (SANs) to produce the CTMC models, while Busch et al. in [9] used numerical aggregation technique to solve the stiffness of the resulting CTMC models. In [8], the author employed a compact representation, but nevertheless the number of states grows very fast such that the time complexity is large. In [9], the time complexity could be reduced but the accuracy of the model, due to its stiffness, decreases. Another approach to analyze the resulting CTMC model from biochemical reaction using adaptive uniformisation was proposed by [10]; this approach can solve more complex chemical reaction networks. The largest network in [10] consists of 19 reactions and 12 species. This approach avoids matrix transition construction and uses threshold to neglect insignificant states. As a result, the approximation error is small and, in only small number of population (in the order of hundreds), this method efficiently approximates the probability distribution of the system.

Henzinger et al. in [11] combined deterministic and stochastic methods dynamically based on a certain population threshold to form a numerical hybrid method for solving CME, by exploiting the fact that in a large population, deterministic and stochastic approaches give the same accuracy. Discrete stochastic variables were used to represent the population number of species under the threshold, and the remaining species were represented by continuous deterministic variables. Besides population threshold, [11] also applied threshold for significant states based on the idea of [10]; so the number of states could be reduced. By using this approach, the computation time is better than the purely stochastic approach and the accuracy is better than the purely deterministic approach. All of the studies described so far concern with the transient analysis of continuous-time Markov chains. For the steady-state analysis of CTMC models of biochemical reaction networks, a method that combines the Lyapunov theory with numerical approximation and bounding techniques was proposed by Dayar et al. in [12].

Several tools exist to compute the transient analysis of chemical reaction networks. SABRE [13] implements a method proposed in [10], while SHAVE [14] implements a method proposed in [11]. The input for SABRE is guarded commands, while SHAVE can receive input in the forms of guarded commands, chemical reactions or SBML file. The comparison of various forms of languages for chemical reaction networks using CTMCs, such as matrix representation, guarded commands, stoichiometric equations and stochastic Petri net was studied by [15]. This study showed that the guarded command form gives many advantages over the others.

To perform transient analysis of our resulting CTMC model, we use SHAVE [14]. We use MIM from [1] to determine all species and chemical reactions involved in this system. We then observe the effect of oxygen pressure to HIF expression. We also observe the role of oxygen sensor FIH in our model. In the end, we compare our result to that of [1], which used deterministic method in their study. To test the accuracy of our model, we use data of 25 genes found in [5]. All genes will be classified into 2 groups: C-TAD responsive and N-TAD dominant. The classification produced by our model will be compared to that of [1] to determine the better model.

## 2. RESEARCH METHOD

### 2.1. Analysis of Molecular Interaction Map (MIM)

MIM is a network diagram illustrating the reactions that occur in gene regulations. All proteins, compounds, enzymes and chemical products are called species in MIM. The MIM of the gene regulation during graded hypoxia is shown in Figure 1. There are 5 types of processes in the MIM: transcription, degradation, association (binding), disassociation (unbinding), and production. Each transcription produces one species, while degradation will degrade the number of species by one. Figure 1 shows that species HIF, PHD, FIH, and VHL will be formed through the process of transcription in which individual precursor sequences are  $S_0$ ,  $S_1$ ,  $S_2$ , and  $S_3$ , respectively. Each transcription depends on the reaction constant  $k_i$ . Each of HIF, PHD, FIH, VHL and HIFOH<sub>a</sub> will be degraded with rate that depends on constants  $l_0$ ,  $l_1$ ,  $l_2$ ,  $l_3$ , and  $l_2$ ,

respectively. Association is a reaction between two species to form one more complex species. The example of association reaction is PHD and HIF binds to form a new complex species, namely HIF:PHD with reaction constant  $a_1$ . The complex species HIF:PHD can either disassociate with constant  $d_1$  or produce HIF(OH)<sub>2</sub><sub>p</sub> and PHD by oxygen stimulation. The production can be stimulated by oxygen or not. The example of production that is not stimulated by oxygen is the production of VHL from complex species 10 ( $x_{10}$ ) with constant  $b_3$ .

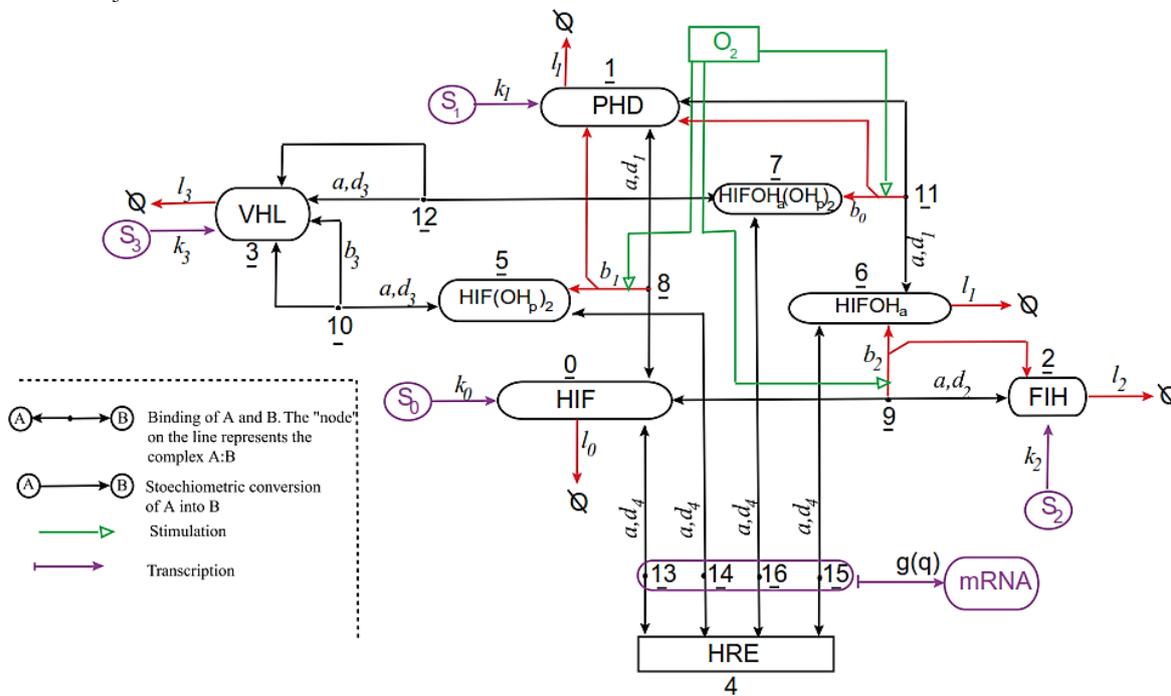


Figure 1. Molecular interaction map of gene regulation in hypoxia [1]

All of the reaction constants in the MIM are shown in Table 1. There are 4 transcription constants ( $k_i$ ), 4 precursors ( $S_i$ ), 4 degradation constants ( $l_i$ ), 4 association constants ( $a_i$ ), 4 disassociation constants ( $d_i$ ) and 3 production constants ( $b_i$ ).

Table 1. Constants and precursors for MIM [1]

Transcription constant ( $k_0-k_3$ )	Degradation constant ( $l_0-l_3$ )	Transcription precursor ( $S_0-S_3$ )	Association constant ( $a_1-a_4$ )	Disassociation constant ( $d_1-d_4$ )	Production constant ( $b_1-b_3$ )
0.0035 $t^{-1}$	1 $t^{-1}$	100 M	1 $t^{-1}M^{-1}$	1 $t^{-1}$	0.000065 $t^{-1}M^{-1}$
5.77 $t^{-1}$	1 $t^{-1}$	100 M	1 $t^{-1}M^{-1}$	1 $t^{-1}$	0.000154 $t^{-1}M^{-1}$
635.34 $t^{-1}$	1 $t^{-1}$	100 M	1 $t^{-1}M^{-1}$	1 $t^{-1}$	1 $t^{-1}M^{-1}$
37.04 $t^{-1}$	1 $t^{-1}$	100 M	1.007 $t^{-1}M^{-1}$	0.1 $t^{-1}$	

The translation of MIM into chemical reactions is shown in Table 2. There are 31 reactions and 19 species involved. Reactions are classified into 4 transcriptions, 4 degradations, 9 associations, 9 disassociations and 5 productions. Three of all 5 productions are stimulated by oxygen. The symbol  $x_i$  in Table 2 means that species  $i$  is formed by either an association of two species or production stimulated by oxygen.

Table 2. Biochemical reactions in hypoxia

Transcription	Degradation	Association/Disassociation	Production
$S_0 \xrightarrow{k_0} HIF$	$HIF \xrightarrow{l_0} \emptyset$	$HIF + PHD \rightleftharpoons x_8$ $d_1$	$x_8 + O_2 \xrightarrow{b_1} x_5 + PHD$
$S_1 \xrightarrow{k_1} PHD$	$PHD \xrightarrow{l_1} \emptyset$	$HIF + FIH \rightleftharpoons x_9$ $d_2$	$x_9 + O_2 \xrightarrow{b_2} x_6 + FIH$
$S_2 \xrightarrow{k_2} FIH$	$FIH \xrightarrow{l_2} \emptyset$	$x_6 + PHD \rightleftharpoons x_{11}$ $d_1$	$x_{11} + O_2 \xrightarrow{b_1} x_7 + PHD$
$S_3 \xrightarrow{k_3} VHL$	$VHL \xrightarrow{l_3} \emptyset$	$x_5 + VHL \rightleftharpoons x_{10}$ $d_3$	$x_{10} \xrightarrow{b_3} VHL$
		$x_7 + VHL \rightleftharpoons x_{12}$ $d_3$	$x_{12} \xrightarrow{b_3} VHL$
		$HRE + HIF \rightleftharpoons x_{13}$ $d_4$	
		$HRE + x_5 \rightleftharpoons x_{14}$ $d_4$	
		$HRE + x_6 \rightleftharpoons x_{15}$ $d_4$	
		$HRE + x_7 \rightleftharpoons x_{16}$ $d_4$	

## 2.2. Model Design

Model design was carried out in 4 steps, as follow:

### 1. Identification of species

There are 17 species involved in this model as shown in Table 3.  $S_0$ ,  $S_1$ ,  $S_2$ ,  $S_3$ , and oxygen are not species because their numbers are constant all the time. We denote the species by  $x_i$  to simplify the state representation. The order of the symbols is based on the MIM depicted in Figure 1.

Table 3. Species involved in the model

Species	Symbol	Species	Symbol
HIF	$x_0$	HIF:FIH	$x_9$
PHD	$x_1$	HIF(OH <sub>p</sub> ) <sub>2</sub> :VHL	$x_{10}$
FIH	$x_2$	HIFOH <sub>a</sub> :PHD	$x_{11}$
VHL	$x_3$	HIFOH <sub>a</sub> (OH <sub>p</sub> ) <sub>2</sub> :VHL	$x_{12}$
HRE	$x_4$	HIF:HRE	$x_{13}$
HIF(OH <sub>p</sub> ) <sub>2</sub>	$x_5$	HIF(OH <sub>p</sub> ) <sub>2</sub> :HRE	$x_{14}$
HIFOH <sub>a</sub>	$x_6$	HIFOH <sub>a</sub> :HRE	$x_{15}$
HIFOH <sub>a</sub> (OH <sub>p</sub> ) <sub>2</sub>	$x_7$	HIFOH <sub>a</sub> (OH <sub>p</sub> ) <sub>2</sub> :HRE	$x_{16}$
HIF:PHD	$x_8$		

### 2. Representation of state

A state represents the number of each species at particular times; so the representation of state is the sequence  $(x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16})$  where  $i=0,1,\dots,16$  and  $x_i \geq 0$  indicates the amount of species  $i$ . For example state  $(0, 75, 3000, 2500, 2, 1, 2, 1, 0, 1, 1, 0, 2, 2, 1, 0, 0)$  denotes that the amount of species HIF, HIF:PHD, HIFOH<sub>a</sub>:PHD, HIFOH<sub>a</sub>:HRE, and HIFOH<sub>a</sub>(OH<sub>p</sub>)<sub>2</sub>:HRE is 0 for each species; the amount of species PHD is 75; FIH 3000; VHL 2500; HRE, HIFOH<sub>a</sub>, HIFOH<sub>a</sub>(OH<sub>p</sub>)<sub>2</sub>:VHL, and HIF:HRE is 2 for each species; and HIF(OH<sub>p</sub>)<sub>2</sub>, HIFOH<sub>a</sub>(OH<sub>p</sub>)<sub>2</sub>, HIF:FIH, HIF(OH<sub>p</sub>)<sub>2</sub>:VHL, and HIF(OH<sub>p</sub>)<sub>2</sub>:HRE is 1 for each species.

3. Determining the initial state and probability

We use the initial state  $Q_1=(0,0,0,0,10,0,0,0,0,0,0,0,0,0,0,0)$  with probability 1. This means that we start our observation when the number of each species is 0, except for HRE. We set HRE non-zero because there are no reactions that form HRE.

4. State transition

A transition from one state to another is triggered by a chemical reaction, where the transition occurs at a specific rate that depends on the constant of the reaction multiplied by the number of species involved in reaction. An example of a transition is shown in Figure 2, where an association of species 0 and 2 (HIF and FIH) with reaction rate  $a_2 \cdot 1 \cdot 100$  occurs and this produces species 9 (HIF:FIH). This means that the number of species 0 and 2 each decreases by 1, while the number of species 9 increases by 1.

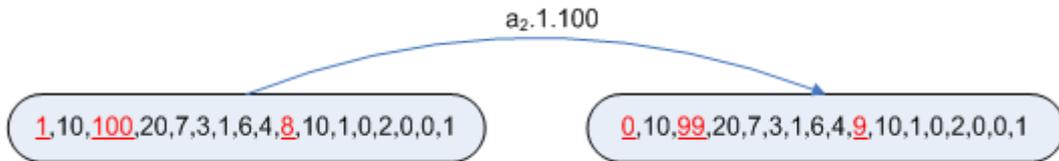


Figure 2. A state transition

A part of the state-transition diagram of this model is depicted in Figure 3. To simplify the diagram, we rename each state by  $Q_i$ . A dotted line next to a state in Figure 1 indicates that there are other incoming and outgoing transitions that are not depicted in the figure for that state. The initial state is  $Q_1$ . There are 4 incoming transitions to  $Q_1$  and 4 outgoing transitions from  $Q_1$ . The incoming transitions consist of a transition from  $Q_2$  denoting a degradation of  $x_0$  (HIF), a transition from  $Q_3$  denoting a degradation of  $x_1$  (PHD), a transition from  $Q_4$  denoting a degradation of  $x_2$  (FIH), and a transition from  $Q_5$  denoting a degradation of  $x_4$  (VHL). The outgoing transitions consist of a transition to  $Q_2$  denoting a transcription of  $x_0$  (HIF), a transition to  $Q_3$  denoting a transcription of  $x_1$  (PHD), a transition to  $Q_4$  denoting a transcription of  $x_2$  (FIH), and a transition to  $Q_5$  denoting a transcription of  $x_4$  (VHL). The actual state representation of each of  $Q_i$ 's is shown in Table 4.

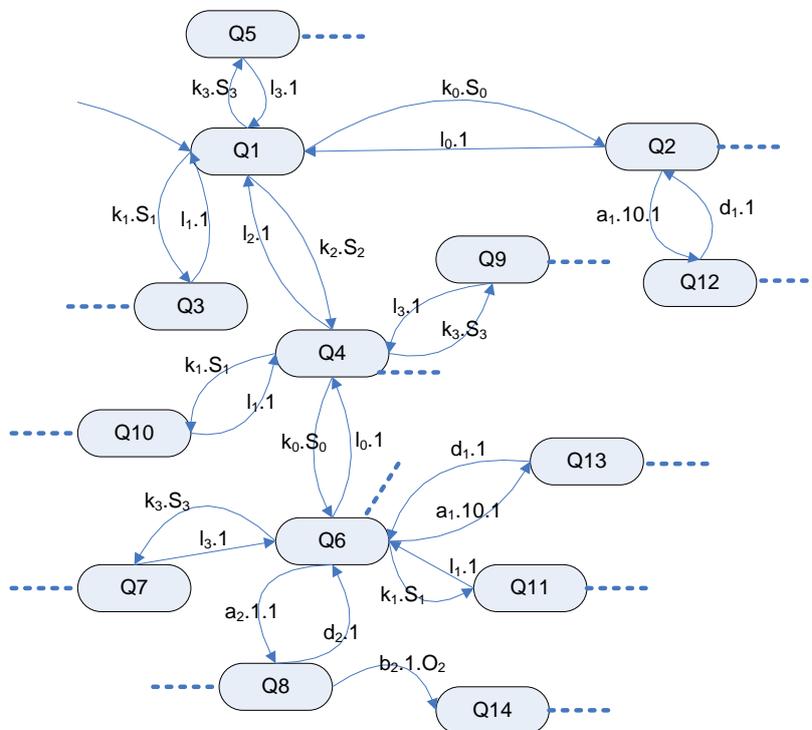


Figure 3. A portion of the state-transition diagram

Table 4. The actual state representations for Figure 3

State Symbol	State ( $x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}$ )
$Q_1$	(0,0,0,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_2$	(1,0,0,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_3$	(0,1,0,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_4$	(0,0,1,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_5$	(0,0,0,1,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_6$	(1,0,1,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_7$	(1,0,1,1,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_8$	(0,0,0,0,10,0,0,0,0,1,0,0,0,0,0,0)
$Q_9$	(0,0,1,1,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_{10}$	(0,1,1,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_{11}$	(1,1,1,0,10,0,0,0,0,0,0,0,0,0,0,0)
$Q_{12}$	(0,0,0,0,9,0,0,0,0,0,0,0,0,1,0,0)
$Q_{13}$	(0,0,1,0,9,0,0,0,0,0,0,0,0,1,0,0)
$Q_{14}$	(0,0,1,0,10,0,1,0,0,0,0,0,0,0,0,0)

### 2.3. Implementation and Data Analysis

We perform transient analysis to observe the state probabilities at a certain time by using SHAVE. SHAVE is a tool for transient analysis that implements the hybrid algorithm of [11]. A threshold  $\delta$  is used to limit the state that has a significant probability. As reported in [10], a good value of  $\delta$  is around  $10^{-14}$ . Threshold  $K$  is used to limit the number of population: if an expected value of any species exceeds  $K$ , then the hybrid algorithm is run. In the hybrid algorithm, a small population (namely, when the number of species is no more than  $K$ ) is represented by a discrete stochastic variable, whose probability is calculated by numerical integration Runge-Kutta 4, while the rest will be represented by a continuous deterministic variable whose value will be calculated by its conditional expectations based on the small population.

The results of the analysis are obtained from SHAVE in the form of csv file that describes the probability distribution of each state and the expected value of each species. We use threshold  $10^{-5}$  in the steady-state analysis; this means that we assume the system enters its steady state if the change of probability is under  $10^{-5}$ . Data obtained in the steady state are then used to observe the expression of HIF in oxygen gradient and the role of oxygen sensor FIH. In order to perform this observation, we build a tool using Borland Delphi 7.

### 2.4. Model testing

The model that has been constructed is tested using data from biological experiments previously described in [5] at 3% oxygen and at anoxia (oxygen 0.02%). There are 25 genes used in the testing. These genes were classified into C-TAD responsive and N-TAD dominant based on their FIH sensitivity value, denoted by  $q$ . FIH sensitivity value is a ratio between gene responses following FIH inhibition (3% oxygen) to gene responses following over-expression of FIH (anoxia). We obtain the real value of FIH sensitivity from [5]. We calculate the value of  $q$  using the steady-state value of HIF, PHD, and VHL obtained from SHAVE. The detailed description of  $q$  can be found [1]. The validity of the model that has been built will be indicated by the number of genes that are correctly classified. Our testing results will then be compared to that of [1].

## 3. RESULTS AND ANALYSIS

In this section, we describe the results of research and provide a comprehensive discussion. Results will be presented in figures, graphs, tables and others to make the reader understand easily.

Experiments are conducted with oxygen pressure ranging from 0% to 21% with increments of 3%. In each of these experiments we will perform transient analysis on the produced CTMC models, which give us information about the number of each species involved in the model at any given time for each oxygen pressure. The analysis will be organized as follows: first, we will explain the expression of HIF in gradient oxygen; second, we will discuss the role of the oxygen sensor FIH on this gene regulation pathway; and last,

we will explain the model testing using 25 genes derived from biological experiments described in [5] at 3% oxygen.

### 3.1. HIF Expression in Gradient Oxygen

Experiments conducted with SHAVE show that by increasing the oxygen pressure, the amount of HIF decreases. The study reported in [1] produced the same conclusion as this study on the effect of oxygen pressure on HIF. Nevertheless, we encounter differences in the values of HIF generated at each oxygen pressure. The values of HIF at various oxygen pressures are shown in Table 5, where the first column shows the levels of oxygen, the second column shows the value of HIF in this study and the third column shows the value of HIF in study [1]. The graph of HIF expression in gradient oxygen are shown in Figure 4, where Figure 4a shows the resulting graph of this study, while 4b shows the graph in study [1].

Table 5. The value of HIF in gradient oxygen

Oxygen	HIF in this study	HIF in [1]
0	0.005422	0.35
3	0.00295	0.26176
6	0.000989	0.149233
9	0.000445	0.087164
12	0.000253	0.055272
15	0.000164	0.037734
18	0.000115	0.027304
21	8.63E-05	0.020671

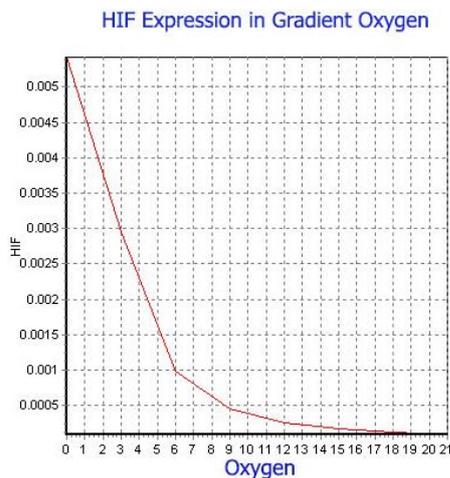


Figure 4a. HIF expression in gradient oxygen

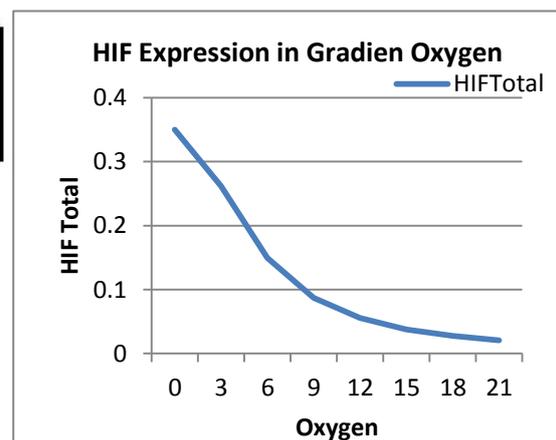


Figure 4b. HIF expression in gradient oxygen [1]

### 3.2. The Role of FIH

To observe the role of FIH in this model, we perform experiments for 6 different values of FIH, i.e. 0, 5, 50, 100, 500 and 1000. For each value, we examine 5 genes with different values (namely, 0, 1, 6, 10, and 100) of FIH sensitivity  $q$ . We then calculate genes expression using the probability distribution obtained in the experiments regarding the gradient oxygen.

The expressions of all genes in each value of FIH are shown in Figure 5. The abscissa stands for the oxygen pressure, while the ordinate stands for the expression of genes. The different values of  $q$  for every gene means that every gene has its sensitivity score to FIH. The value of  $q$  equals to 0 means that the gene is not sensitive towards FIH (N-TAD only gene). Figure 5a indicates that for FIH value equals to 0, the expression of all genes are the same for the same oxygen pressure. The study of [1] showed the same result as ours. Figures 6b, 6c, and 6d show the genes expression for FIH values 5, 10, and 100, respectively. Our result shows that for those values, the genes expression is different but minor. On the other hand, [1] showed that there are major differences between genes. Figures 6e and 6f indicate that there is a major difference in gene expression: from Figures 6e and 6f we can clearly observe that the larger value of  $q$  results in the smaller gene expression.

We can conclude that FIH does not affect N-TAD only genes ( $q=0$ ), as shown in Figures 6a to 6f, the expression when  $q=0$  is the same for different values of FIH. Furthermore, FIH inhibits the genes that have  $q$  more than 1 (C-TAD responsive), and the more value of  $q$ , the more inhibition by FIH.

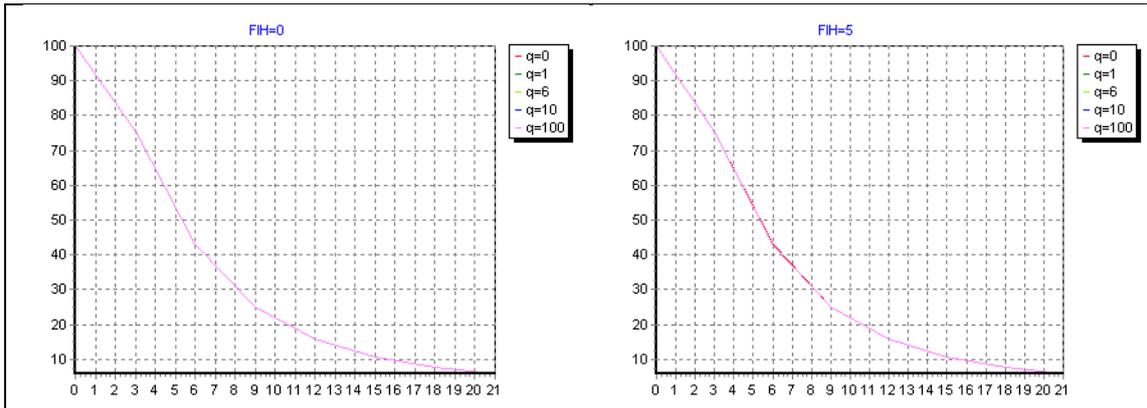


Figure 5a. Genes expression when FIH=0

Figure 5b. Genes expression when FIH=5

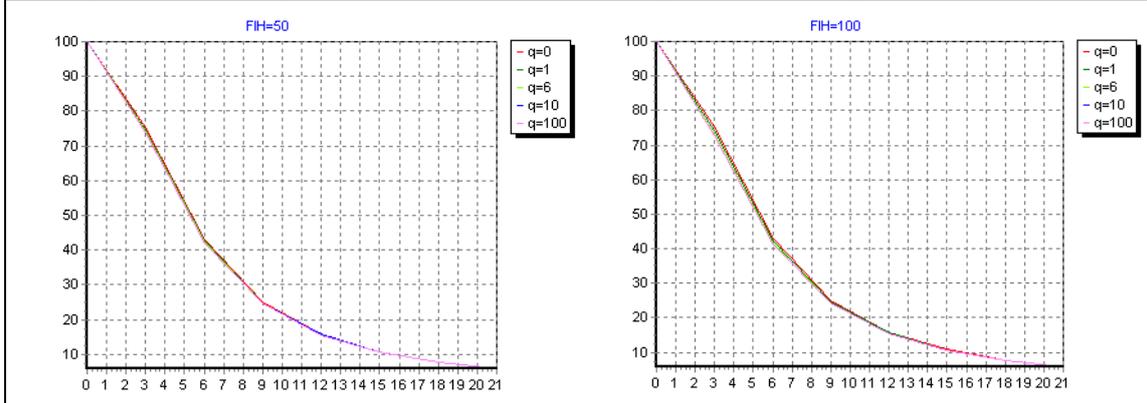


Figure 5c. Genes expression when FIH=50

Figure 5d. Genes expression when FIH=100

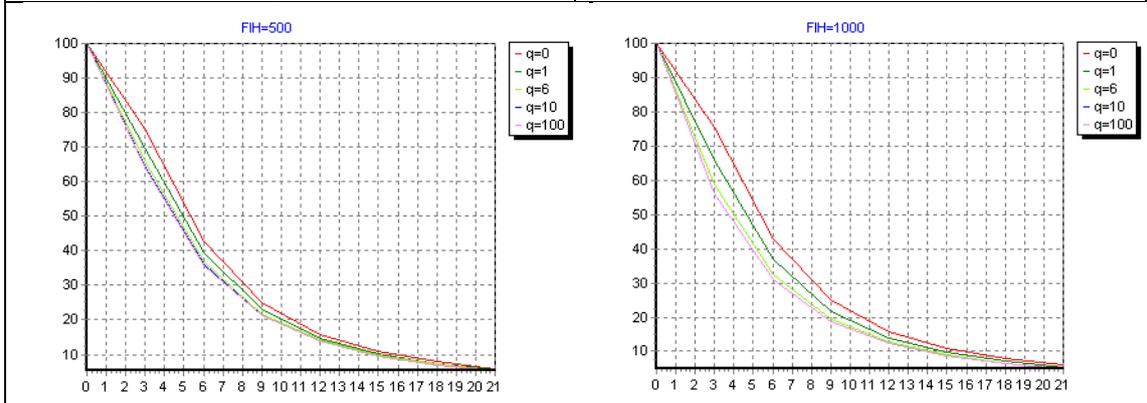


Figure 5e. Genes expression when FIH=500

Figure 5f. Genes expression when FIH=1000

### 3.3. Classification of Genes

Testing of the model is carried out by grouping genes into C-TAD responsive and N-TAD dominant based on their calculated values of FIH sensitivity  $q$ . We used the 25 genes from [5]. The result is shown in Figure 6, where *phd3* with a value of  $q$  109.93 is excluded from the graph because the value is too far from the actual value for the FIH sensitivity. In Figure 6, N-TAD dominant genes are indicated by the red ellipse; while the C-TAD responsive genes with intermediary value of  $q$  are inside the black circle and with a high value of  $q$  are inside the black ellipse. Another gene that cannot be grouped is *bnip3*. Its actual sensitivity score is below one, which means that the value of calculated  $q$  must be below 0 [1]. Two genes that are grouped incorrectly are *trefoil factor 3* and *p21*. They should have been inside the C-TAD responsive group, but they are grouped as N-TAD dominant because their calculated values of  $q$  are very low.

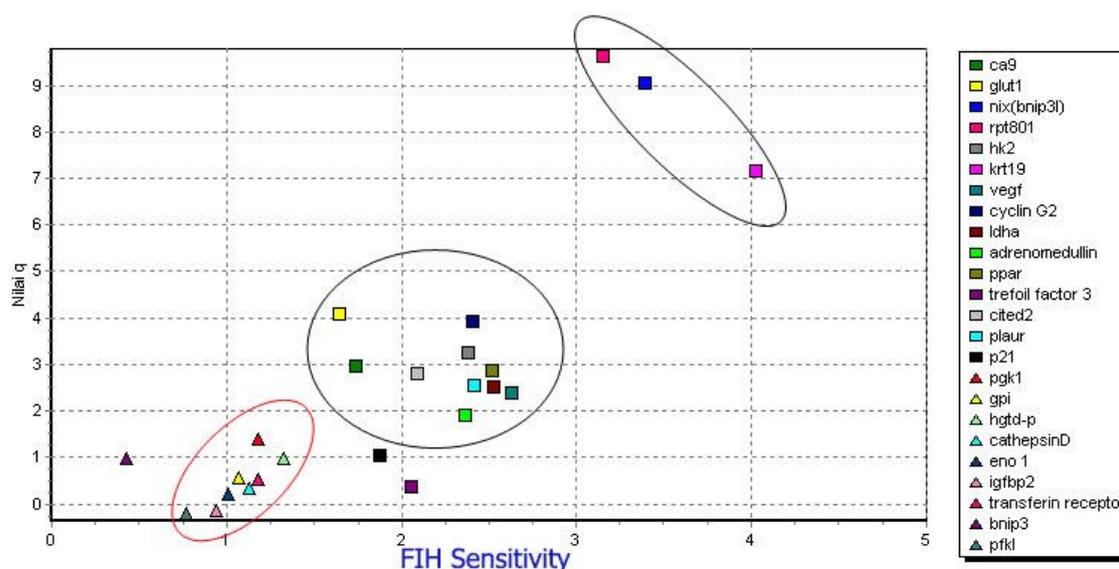


Figure 6. Classification of genes

#### 4. CONCLUSION

The stochastic model we have developed in this paper shows that increasing the oxygen levels will decrease the number of the major protein in hypoxia pathway, HIF. The model also shows that the oxygen sensor FIH only inhibits the activity of C-TAD responsive genes. The model accuracy is 84% and it is same as that of [1]. This result indicates that the stochastic approach can work synergistically with biological experiments by [5] and mathematical model by [1] in terms of providing a better understanding of the expression of HIF and the role of oxygen sensor FIH in gene regulation in response to hypoxia.

Although this study and [1] produce the same accuracy and result in the observation of HIF expression and FIH role, but there is little difference in the number of species. This is because in the biochemical reaction network of hypoxia, not all species have small population size. Species, such as FIH, PHD and VHL, have a big population size, and in this case, the stochastic approach is simply equal to the deterministic one. Other species, such as HIF and HRE, have a small population size, and the hybrid approach, in this case, produces different results, albeit in small amounts. It must be noted that our study produces the probability distributions that can be used to observe and to further study the dynamics of the system. For further research, the probability distribution produced in the transient analysis can be used to determine other measures of interest in the regulation of genes in graded hypoxia such as switch-like behavior and the distribution of protein production.

#### ACKNOWLEDGEMENTS

We thank to Linar Mikeev for helping us carrying out the transient analysis in SHAVE.

#### REFERENCES

- [1] Dayan, F., Monticelli, M., Pouyssegur, J., Pecou, E. Gene Regulation in Response to Graded Hypoxia: The Non-Redundant Roles of The Oxygen Sensors PHD and FIH in the HIF Pathway, *Journal of Theoretical Biology*. 2009; 259: 304-316.
- [2] Ziello, J. E., Jovin, I. S., Huang, Y. Hypoxia-Inducible Factor (HIF)-1 Regulatory Pathway and its Potential for Therapeutic Intervention in Malignancy and Ischemia, *Yale Journal of Biology and Medicine*. 2007; 80: 51-60.
- [3] Burrows, N., Babur, M., Resch, J., Williams, K. J., Brabant, G. Hypoxia-Inducible Factor in Thyroid Carcinoma. *Journal of Thyroid Research*. 2011; Volume 2011: 1-17.
- [4] Brahimi-Horn, Pouyssegur, J. Harnessing the Hypoxia-Inducible Factor in Cancer and Ischemic Disease. *Biochem Pharmacol*. 2007; 73(3): 450-457.
- [5] Dayan, F., Roux, D., Brahimi-Horn, M. C., Pouyssegur, J., Mazure, N. M. The Oxygen Sensor Factor-Inhibiting Hypoxia-Inducible Factor-1 Controls Expression of Distinct Genes Through the Bifunctional Transcriptional Character of Hypoxia-Inducible Factor-1 $\alpha$ . *Cancer Res*. 2006; 66(7): 3688-3698.
- [6] Yu, Y., Wang, G., Simha, R., Peng, W., Turano, F., Zeng, C., Pathway Switching Explains the Sharp Response Characteristic of Hypoxia Response Network. *PLoS Computational Biology*. 2007; 3(8): 1657-1667.

- [7] Sandmann, W., Wolf, V. *Computational Probability for Systems Biology*. Proceedings of the International Workshop on Formal Methods in Systems Biology. Cambridge. 2008; 5054:33-47.
- [8] Wolf, V., Modelling of Biochemical Reactions by Stochastic Automata Networks. *Electr. Notes Theor. Comput. Sci.* 2007; 171(2): 197-208.
- [9] Busch, H., Sandmann, W., Wolf, V. *A Numerical Aggregation Algorithm for the Enzyme-Catalyzed Substrate Conversion*. Proceedings of the 4<sup>th</sup> International Conference on Computational Methods in Systems Biology (CMSB'06). Trento-Italy. 2006; 4210: 298-311.
- [10] Didier, F., Henzinger, T.A., Mateescu, M., Wolf, V. Fast Adaptive Uniformization of the Chemical Master Equation. *IET Systems Biology*. 2010; 4(6): 441-452.
- [11] Henzinger, T.A., Mateescu, M., Mikeev, L., Wolf, V., Hybrid Numerical Solution of the Chemical Master Equation. *Proceedings of the 8<sup>th</sup> International Conference on Computational Methods in Systems Biology (CMSB'10)*. Trento-Italy. 2010; 55-65.
- [12] Dayar, T., Hermanns, H., Spieler, D., Wolf, V., Bounding the Equilibrium Distribution of Markov Population Models. *Numerical Linear Algebra with Applications*. 2011; 18(6): 931-946.
- [13] Didier, F., Henzinger, T. A., Mateescu, M., Wolf, V. *SABRE: A Tool for Stochastic Analysis of Biochemical Reaction Networks*. Proceedings of the 7<sup>th</sup> International Conference on the Quantitative Evaluation of Systems (QEST'10). Williamsburg-Virginia. 2010: 193-194.
- [14] Lapin, M., Mikeev, L., Wolf, V. *SHAVE: Stochastic Hybrid Analysis of Markov Population Models*. Proceedings of the 14<sup>th</sup> International Conference on Hybrid Systems: Computational and Control (HSCC'11). Chicago. 2011; 311-312.
- [15] Henzinger, T. A., Jobstmann, B., Wolf, V. Formalism for Specifying Markovian Population Models. *International Journal of Foundations of Computer Science*. 2001; 22(4): 823-841.

## BIOGRAPHY OF AUTHORS



Putu Indah Ciptayani received her bachelor's degree from Universitas Brawijaya, Indonesia in 2008; her master's degree from the University Gadjah Mada, Indonesia in 2012. Since 2012 she has been a contract lecturer at the department of Computer Science, Universitas Udayana, Indonesia.



Reza Pulungan received his bachelor's degree in 1999 from Universitas Gadjah Mada, Indonesia; his master's degree in 2004 from the University of Twente, the Netherlands; and his doctoral degree in computer science in 2009 from Saarland University, Germany. Since 2009 he has been a lecturer and a researcher at the department of Computer Science and Electronics, Universitas Gadjah Mada. His research interests lie in the field of stochastic processes, especially Markov processes and phase-type distributions. He is also interested in modeling and analysis of networked systems.