Journal of Bioinformatics and Computational Biology
Vol. 4, No. 5 (2006) 1119–1140

Imperial College Press
www.icpress.co.uk

# STRUCTURAL MODELING AND ANALYSIS OF SIGNALING PATHWAYS BASED ON PETRI NETS

CHEN LI*, SHUNICHI SUZUKI†, QI-WEI GE‡ and MITSURU NAKATA§

*Graduate School of Science and Engineering, ‡,§Faculty of Education
Yamaguchi University
1677-1 Yoshida, Yamaguchi 753-8511, Japan

†Tokyo Head Office, TIS System Service Inc.,
Toyo Center Bldg., 29-15, Toyo-cho 5-chome, Koto-ku, Tokyo 135-0016, Japan
*li@ib.sci.yamaguchi-u.ac.jp
†toshikazu@ib.sci.yamaguchi-u.ac.jp
‡gqw@inf.edu.yamaguchi-u.ac.jp
§nakata@inf.edu.yamaguchi-u.ac.jp

HIROSHI MATSUNO

Graduate School of Science and Engineering, Yamaguchi University
1677-1 Yoshida, Yamaguchi 753-8512, Japan
matsuno@sci.yamaguchi-u.ac.jp

SATORU MIYANO

Human Genome Center, Institute of Medical Science, University of Tokyo
4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan
miyano@ims.u-tokyo.ac.jp

Received 5 February 2006
Revised 10 June 2006
Accepted 12 June 2006

The purpose of this paper is to discuss how to model and analyze signaling pathways by using Petri net. Firstly, we propose a modeling method based on Petri net by paying attention to the molecular interactions and mechanisms. Then, we introduce a new notion "activation transduction component" in order to describe an enzymic activation process of reactions in signaling pathways and shows its correspondence to a so-called elementary T-invariant in the Petri net models. Further, we design an algorithm to effectively find basic enzymic activation processes by obtaining a series of elementary T-invariants in the Petri net models. Finally, we demonstrate how our method is practically used in modeling and analyzing signaling pathway mediated by thrombopoietin as an example.

*Keywords*: Signaling pathway; Petri net; elementary T-invariant; enzymic activation process; activation transduction component.

## 1. Introduction

Cellular activities are precisely maintained in good condition including various biochemical interactions and processes, such as signaling pathways, metabolic pathways, and gene regulatory networks. So far, signaling pathways have been widely studied in cell biology. They are information cascades of enzyme reactions from transmembrane receptors to the nucleus DNA, which ultimately regulate intracellular responses. Till now, the modeling and analysis of biological networks have been investigated from quantitative and qualitative aspects by using various types of Petri nets: low level Petri nets,[1–3] stochastic Petri nets,[4] hybrid Petri nets,[5,6] colored Petri nets,[7–9] functional Petri nets,[10] and so on. By using qualitative method, researchers could gain lots of important insights into the behaviors of the models at a relatively low cost in terms of effort and computational time, even without quantitative data. The qualitative analysis for even large scale and complex biological networks can be handled with the intuitive structural and behavioral properties defined by Petri nets.

Many studies[1,2,5,7,9–12] on modeling and analyzing metabolic pathways using Petri net have been developed from the first paper by Reddy *et al.*[2] in 1993. Among them, functional Petri nets were used to calculate dynamic biocatalytic processes of metabolic pathways with functions for specifying the arc-weights.[10] However, there have been no attempts yet to simulate signaling pathways by functional Petri nets, since the signaling pathways are generally more complex. Accordingly, a bit of investigation focusing on these structural properties but not dynamic behaviors of signaling pathways have been provided. Heiner *et al.*[3] have proposed a method for developing and analyzing models of biological pathways in a systematic manner by calculating T-invariants to obtain all paths in signaling pathways. However, some modeling inconsistencies are noticed in the Petri net example shown; further the analysis method is not sufficient to discuss the general systematic behavior since they did not consider the effect of enzymes. Therefore, we propose a new Petri net based method to consistently model and analyze a signaling pathway with a focus on enzymes by which all the chains consisting of enzymic activation processes in signaling pathways can be provided.

The paper is organized as follows: First, we present a brief introduction of elementary flux modes into metabolic pathways that are known to correspond to elementary T-invariants of Petri net. Then, we propose a modeling method based on Petri net by taking notice of molecular interactions and mechanisms. Further, we introduce a new notion "activation transduction component" to express an enzymic activation process that has an elementary T-invariant in Petri net model as a counterpart. In the next section we design an algorithm to find such basic systematic components of signaling pathways by calculating a series of elementary T-invariants. In the final section, an application of proposed method is given with the example of thrombopoietin (TPO) signaling pathways.

## 2. Elementary Flux Mode and Elementary T-Invariant

Ideas to use Petri nets for modeling and analyzing metabolic pathways have been popular from the quantitative and qualitative points of view. A Petri net is a bipartite graph with two different types of vertices: places and transitions, which have the metabolites and reactions as counterparts respectively in metabolic pathways.

So far, besides the attempts to represent systems of metabolic pathways by setting up differential equations to investigate dynamics of concentration change in metabolites, the modeling using linear algebraic equations has also been launched under the assumption that the metabolites have reached a dynamic concentration equilibrium (steady state) due to the augmentation of the number of speed parameters. Schuster *et al.*[13,14] have proposed a concept of *elementary flux mode* (elementary mode or EM for short) developed from convex analysis to accurately reflect complete behaviors of a metabolic pathway with a set of linear paths. Elementary modes, which were defined as a minimal set of enzymes that could operate at steady state with all irreversible reactions proceeding in the appropriate direction, provide a mathematical tool to define and describe all metabolic routes. That is, for elementary modes, any disturbance to one enzyme belonging to this minimal set will result in a cessation of any flowing and a disruption of a dynamic concentration equilibrium of metabolites in the system.[13,14]

It is interesting to note that elementary flux modes corresponds to elementary T-invariants of Petri net.[11,12] In Ref. 9, elementary T-invariants are calculated from Petri nets that model glycolytic pathway and pentose phosphate pathway with showing the process to discover elementary modes.

As for intracellular interaction pathways, there also exist signaling pathways besides metabolic pathways. We have known that metabolic pathways can be described by elementary modes. In next section, we will inquire the characteristic behaviors of signaling pathways by elementary T-invariants.

## 3. A Petri Net Based Model for Signaling Pathways

Here, a new modeling method for signaling pathways is proposed on account of potential advantages of Petri net whose representation is easy to understand due to its graphical and precise nature. The aims of the modeling by Petri net for signaling pathways are: (i) to make the biologists intuitively understand the intrinsic structure and features of signaling pathways, and (ii) to make it possible to mechanically model larger and more complicated signaling pathway networks.

### 3.1. *Basic definitions*

In this subsection, we only give the necessary definitions used in this paper. For detailed definitions of Petri net, refer to Ref. 15. The followings are the mathematic

definitions for Petri nets as described below:

**[Definition 1]** A Petri net is denoted as $PN = (T, P, E, \alpha, \beta)$ that is a bipartite graph,[15] where $E = E^+ \cup E^-$ and

$T$: a set of transitions $\{t_1, t_2, \cdots, t_{|T|}\}$
$P$: a set of places $\{p_1, p_2, \cdots, p_{|P|}\}$
$E^+$: a set of arcs from transitions to places $e = (t, p)$
$E^-$: a set of arcs from places to transitions $e = (p, t)$
$\alpha$: $\alpha(e)$ is the weight of arc $e = (p, t)$
$\beta$: $\beta(e)$ is the weight of arc $e = (t, p)$. □

**[Definition 2]** Let $PN$ be a Petri net.

(1) $^\circ t$ (or $t^\circ$) is a set of the input (or output) places of $t$ and called the *pre-set* (or *post-set*) of transition $t$.
(2) The structure of $PN$ can be represented by a matrix, called *place–transition incidence matrix* (incidence matrix for short) and denoted by $C = C^+ - C^-$, where

$$C^+(i,j) = \begin{cases} \beta(e) & \text{if } e = (t_j, p_i) \in E^+ \\ 0 & \text{otherwise;} \end{cases}$$

$$C^-(i,j) = \begin{cases} \alpha(e) & \text{if } e = (p_i, t_j) \in E^- \\ 0 & \text{otherwise.} \end{cases}$$

(3) Token distribution to places is called *marking* and expressed by $M = (m_1, m_2, \cdots, m_{|P|})^t$, where, $m_i$ is the number of tokens at $p_i$.
(4) A transition sequence $\sigma = t_1 t_2 \cdots t_k$ is called *firing sequence* from $M_I$ to $M_F$, if the firing simulation of $\sigma$ on $M_I$ can be carried out all the way to the last element of $\sigma$, which leads to the marking $M_F$. The marking transition is expressed by $M_I[\sigma\rangle M_F$ and the firing numbers of all the transitions are expressed by a *firing count vector* $J = (j_1, j_2, \cdots, j_{|T|})^t$. The relationship among $C$, $J$, $M_I$ and $M_F$ can be expressed by $M_F = M_I + CJ$. □

**[Definition 3]** Let $PN$ and $C$ be a Petri net and its incidence matrix, respectively.

(1) A non-negative integer vector $J$ satisfying $CJ = 0$ is called *T-invariant* and the set of transitions $T_J = \{t_i \in T | j_i \neq 0\}$ is called the *support* of $J$.
(2) For a T-invariant $J$ with the support $T_J$, if there exists no such T-invariant $J'$ whose support $T_{J'}$ satisfies $T_{J'} \subset T_J$, then $T_J$ is called *minimum support*. Further for a T-invariant $J$ with minimum support $T_J$, if all the values $\{j_i | t_i \in T_J\}$ have no common divisor then $J$ is called *elementary T-invariant*.
(3) A subnet $N_J$ is called "generated by a set of transition $T_J$" if $N_J$ is such a subnet that $N_J$ is composed of all the transitions $t$ included in $T_J$ and all the places included in the pre-set and post-set of any $t \in T_J$.

(4) An inhibitor arc $e_i$ represents inhibitor function which is depicted as a line with a hollow circle at the end where the arrowhead normally appears. An inhibitor arc disables a transition to fire if the upstream place is occupied by a token, but does not consume the token. □

As defined above, a T-invariant is a non-negative integer vector that returns a Petri net's marking to its initial marking. That is, a T-invariant is a firing sequence of transitions that expresses a periodic behavior of a Petri net and an elementary T-invariant expresses such a minimum periodic behavior that can not be divided furthermore.

### 3.2. *Modeling rules*

In the following, we give the modeling rules for signaling pathways based on Petri net representation.

(1) Places denote static elements including chemical compounds, conditions, states, substances, and cellular organelles participating in the biological pathways. Tokens indicate the presence of these elements. The number of tokens is given to represent the amount of chemical substances. Current assignment of tokens to the places are expressed in form of a vector, namely a *marking* as defined above.
(2) Transitions denote active elements including chemical reactions, events, actions, conversions, and catalyzed reactions. A transition fires by taking off tokens from its individual input places and creating new tokens that are distributed to its output places if its input places has at least as many tokens in it as arc weight from the place to the transition.
(3) Directed arcs connecting the places and the transitions represent the relations between corresponding static elements and active elements. Arc weights $\alpha$ and $\beta$ (defined in **Definition 1**) describe the quantities of substances required before and after reaction, respectively. Especially in case of modeling a chemical reaction, arc weights represent quantities given by stoichiometric equations of the reaction itself. Note that, weight of an arc is omitted if the weight is 1.
(4) Since an enzyme itself plays a role of catalyzer in biological pathways and there occurs no consumption in biochemical reactions, an enzyme is exceptionally modeled in **Definition 4** below.
(5) An inhibition function in biological pathways is modeled by an inhibitor arc.

Generally in Petri net theory, weights of arcs are supposed to be positive integer; however, in this paper, we assume weights of all arcs could be positive rational in representing degradation of compounds. All rational could be multiplied by a common denominator to obtain integers when analyzing signaling pathways to be shown in Sec. 5. Note that, using rational arc-weights involves taking risks of high computational costs and computational overflow.

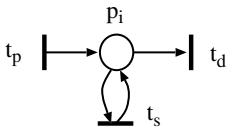



Fig. 1. An enzyme place in Petri net model.

[**Definition 4**] An enzyme in a biological pathway is modeled by a place, called *enzyme place*, as shown in Fig. 1.

(1) Enzyme place $p_i$ has a self-loop with same weight connected from and to $t_s$. Once an enzyme place is occupied by a token, the token will return to the place again to keep the firable state, if the transition $t_s$ is fired.
(2) Let $t_p$ and $t_d$ denote a token provider of $p_i$ and a sink output transition of $p_i$, respectively, where the firing of $t_p$ represents an enzyme activation reaction and the firing of $t_d$ implies an extremely small natural degradation in a biological pathway. $p_i$ holds up token(s) after firing transition $t_p$ and the weights of the arcs satisfy $\alpha(p_i, t_d) \ll \alpha(p_i, t_s)$. □

Numerous reaction types of molecular interaction mechanisms have been described by Petri net model,[16] which suffices to give the description of the metabolic pathway presently.[2] For signaling pathways, it has been pointed out that the additional information among the molecular interactions also should be extraordinarily distinguished according to different types of interactions.[17] To explicitly understand the structural complicated signaling pathways, the modeling of each essential molecular interaction by using Petri net is the first step in modeling the network of signaling pathways as a qualitative event system. Emphasizing a focus on possible molecular interactions as long as we have known, we summarize various molecular interactions of signaling pathways (left side of dashed line) and their corresponding Petri net model (right side of dashed line) in Fig. 2. Both of them in a reaction type are described as a "block" labeled with roman numeral in this paper. Since the majority of molecular interactions can be naturally and directly modeled (refer to Fig. 2), we describe the rest interactions and corresponding model in the following:

**IV.** Generally, continued activated ligand-receptor complex regulates varied majority of cellular pathways transmitting the signals within the cell. Few methods using Petri nets have been proposed to model such activated complex place possessing more than one transition that can trigger down-stream signaling pathways.[3] Their methods are easily understood, but have some problems that, if the transition of such place fires to remove the token(s) in shared input place at one time epoch, it will disable rest transitions simultaneously although the token will return back the same input via a self-loop. Hence, we need a more appropriate model to express this system's behavior. Our basic consideration

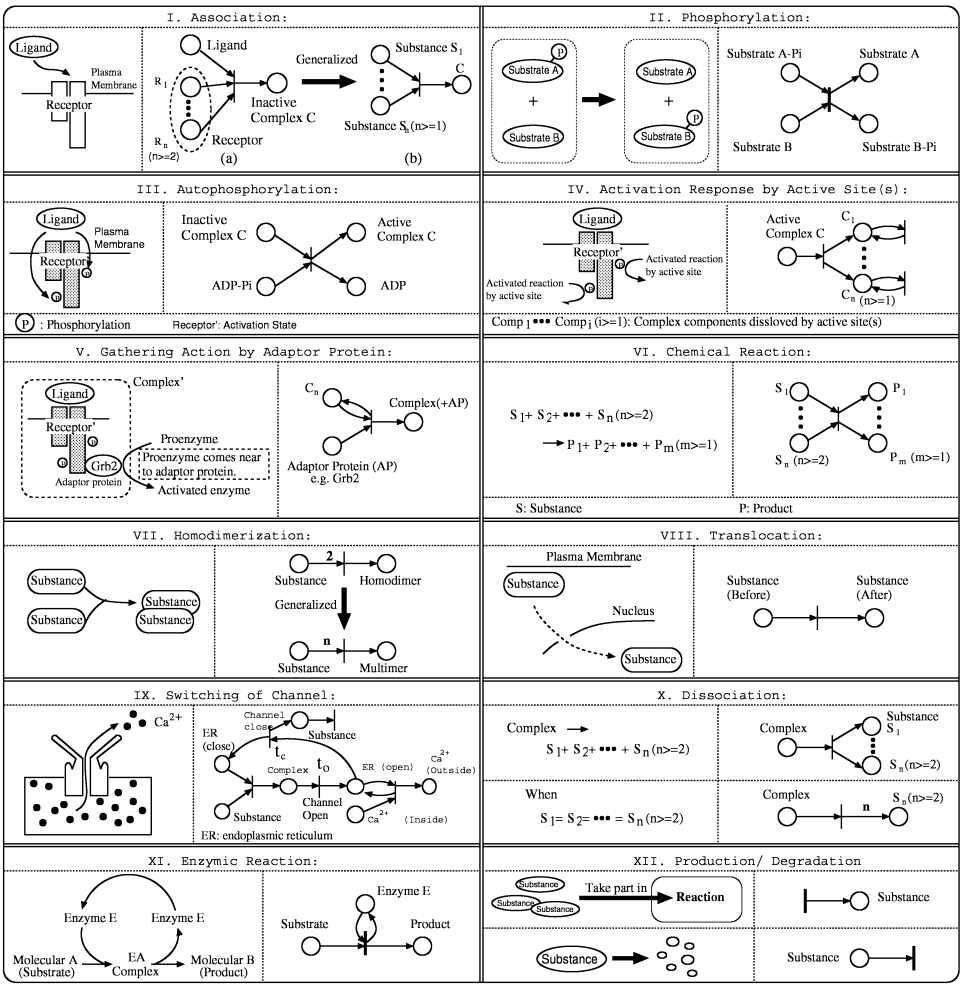



Fig. 2. Petri net models of various reaction types in signaling pathways. **I:** The transition is unfirable in the absence of place of ligand although receptors exist; **X** is the opposite of **I**. **II:** phosphorylation is a reaction to add a phosphate ($PO_4$) group to a protein or a small molecule and dephosphorylation that is the backward reaction of phosphorylation removing phosphate groups from a compound by hydrolysis. **III:** autophosphorylation is a transphosphorylation reaction frequently following the binding of a ligand to a receptor with intrinsic protein kinase activity. **V:** gathering action by adaptor protein is distinguished from association reaction; the main participator adaptor protein is an accessory protein to main proteins. These proteins lack the intrinsic enzymic activity themselves but instead mediate specific protein-protein interactions driving the formation of protein complexes. **VI:** in chemical reactions, the conversion of substances to products ordinarily modeled as input places to output places, both belonging to the same transition. **VII:** a substance is modeled as an input place connected with a 2-weighted arc. It is easy to expand the conception to model the formation of multimer holding $n$-weight such as trimer and a tetramer. For the underpart Petri net model of **X**, it is the opposite consideration of modeling homodimerization reactions. **VIII:** a transition is modeled to indicate the movement action of substances before and after. **XI:** substrates of enzymic reactions are catalyzed to products by enzymes modeled to enzyme places, whereas the reactions are modeled to a transition. **XII:** a source transition denotes an activity that substances take part in respective reactions, whereas a sink transition denotes an extremely small natural degradation of substances.

is that, if there is plurality of successive signaling pathways depending on distinct active site(s) (subunits) of activated complex, all the active site(s) shall be regarded as complex component(s) $C_1, \cdots, C_n (n \geq 1)$ as shown in block **IV**.

**IX.** Intracellular signal pathways are largely carried out by second messenger molecules. $Ca^{2+}$ acts as a second messenger molecule to carry out large intracellular signal inside the cell. Usually the concentration of free $Ca^{2+}$ within the cell is very low; it is stored inside of organelles, mostly the endoplasmic reticulum. To become active, $Ca^{2+}$ has to be released from the organelles into the cytosol. Two transitions $t_o$ and $t_c$ are introduced to denote channel activity of "open" and "close," respectively. $t_o$ is enabled when input place holds up token(s) after the association of organelles and substances, whereas $t_c$ is enabled as long as some stop mechanisms shutoff the channel.

## 4. Characterizing Signaling Pathways

### 4.1. *Signal propagation of signaling pathways*

A signaling pathway is a set of chains of intracellular signaling events which starts by attaching ligands at receptors and ends by altering target proteins, which are responsible for modifying the behaviors of a cell. These signaling events are mediated by intracellular signaling proteins (enzymes as usual) that relay the signal into the cell by activating the next enzyme from inactive state to active state on receipt of signal in the chain.

Many of the enzymes controlled by reactions such as phosphorylation are enzymes themselves. In the enzymic cascades, an enzyme activated by phosphorylation phosphorylates the next enzyme in sequence. That is, the signal in signaling pathways propagates itself in the form of a series of chains consisting of sequential enzymic activation processes where a certain protein changes from "inactivate" state to "activate" state depending on the function of an upstream enzyme. Therefore, it is important to inquire into the behaviors of sequential enzymic activation processes of signaling pathways.

### 4.2. *Activation transduction component and elementary T-invariant*

We call a set of reactions and related substances that make an enzyme active as *activation transduction component*. Figure 3(a) shows an activation transduction component of Ras activity regulation mechanism for enzyme Ras-GTP. Modeling this activation transduction component according to our modeling rules, it is clear that the activation transduction component corresponds to a subnet in which there is a T-invariant.

As we have mentioned, T-invariant is expressed by a vector and if each transition fires as many times as the vector indicates, the initial marking will be restored. In

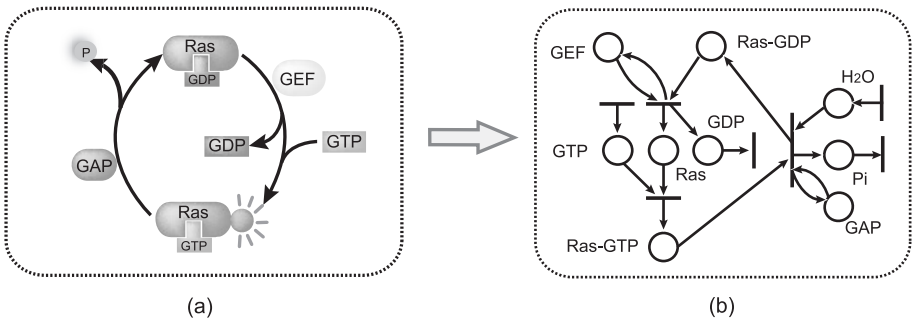



Fig. 3. The association between activation transduction component and elementary T-invariant with mapping a pathway to a Petri net model. (a) Activation transduction component of Ras activity regulation mechanism. (b) A Petri net model of (a) corresponding to a subnet of an elementary T-invariant.

Petri net modeled systems, a periodic behavior is represented by a T-invariant $J$ and the corresponding net is such one $N_J$ generated by $T_J$. This net $N_J$ has such features that (i) before and after any firing sequence corresponding to a T-invariant, the tokens on each place in $N_J$ are kept constant, and (ii) all the transitions in $N_J$ take part in the firing sequence. An elementary T-invariant expresses fundamental periodic behaviors and is a T-invariant that cannot be decomposed by nonnegative rational linear combination of the other T-invariants.

In Fig. 3, the Petri net model is exactly a corresponding net of an elementary T-invariant. This is because (1) an activation transduction component always behaves periodically in order to transmit the signals from the precedent steps to the next as long as tokens on the enzyme place exists and during its process, no token will decrease or increase on any places, and (2) it has no any other behavior except (1), i.e. an activation transduction component cannot be decomposed furthermore. Therefore, we can treat activation transduction components as corresponding nets of elementary T-invariants.

## 5. Algorithm to Find Activation Transduction Components

In this section, we show an algorithm to give relations among activation transduction components in signaling pathways in order to clarify how enzymic activation processes occur.

In a Petri net with inhibitor arcs, tokens in the place connected with an inhibitor arc never vary with the firing of the transition connected with the inhibitor arc, and thus we can simply delete the inhibitor arc. Therefore, in our algorithm Petri nets are supposed to have no inhibitor arcs.

We show an algorithm to identify a series of chains consisting of sequential activation transduction components by finding a series of elementary T-invariants from sink transition(s) except the sink transitions of enzyme places. Computation

of elementary T-invariants has been studied for decades. Linear programming technique has been taken into account in computing some elementary T-invariants.[18] Based on this, an approach has been done trying to obtain all the elementary T-invariants.[19] An algorithm by repeating pivoting operations has been proposed by Avis *et al.*[20] Fourier-Motzkin method[21] and its improved method[22] are also well-known. In this paper, we adopt the method of Ge *et al.*[23] to compute all the elementary T-invariants of a Petri net by applying Linear Programming (LP) technique. In Ref. 23, Ge *et al.* have proposed an efficient method to generate all the elementary T-invariants by applying LP technique. They have proposed an algorithm ≪**Searching Basic-Feasible Solution with** $x_s>0$≫ to search all the elementary T-invariants in each of which the element $x_s$ related to transition $t_s$ is always non-zero. Note that the term of basic-feasible solution of LP is a basic solution that satisfies all the constraints. For the details of this algorithm, the readers are suggested to refer to Ref. 23. In the following, incidence matrix $C$ of $PN$ is supposed to be rewritten as $C \leftarrow C * \frac{1}{\alpha(p_i,t_d)}$ in order to compute elementary T-invariants, where $\alpha(p_i, t_d) \ll 1$.

**≪Searching Activation Transduction Components≫**

1° Let $PN$ be a given Petri net, and $L_s$ be a list of sink transitions (except sink output transitions of enzyme places) in $PN$. Do $SN_J \leftarrow \phi$, $T_{sink} \leftarrow \{t | t \in L_s\}$, $T_{gen} \leftarrow \phi$ and initialize $FIFO$ queue $Q \leftarrow \phi$.

2° If $Q \neq \phi$, pull a subnet $N_{J_i}$ from $Q$ and do the followings:
(i) let $P_e$ and $T_e$ be a set of enzyme places in $N_{J_i}$ and a set of transitions providing tokens to the places of $P_e$, respectively;
(ii) let $L_e$ be a list of transitions in $T_e - T_{gen}$, and do $L_s \leftarrow L_s \cdot L_e$, $T_{gen} \leftarrow T_{gen} \cup T_e$ and $PROV(N_{J_i}) = \{t | t \in T_e\}$.

3° If $L_s = \phi$ go to 4°, otherwise take out a transition $t$ from the beginning of $L_s$ and do $gen(t) \leftarrow \phi$. Obtain all the elementary T-invariants $\{J_i\}$ with $J_i(t) > 0$ by applying ≪**Searching Basic-Feasible Solution with** $x_s > 0$≫[23]. For each $J_i$, do the followings:
(i) obtain its corresponding subnet $N_{J_i}$ (generated by the support $T_{J_i}$ of $J_i$);
(ii) do $gen(t) \leftarrow gen(t) \cup \{N_{J_i}\}$;
(iii) if $N_{J_i} \notin SN_J$ is satisfied, then $SN_J \leftarrow SN_J \cup \{N_{J_i}\}$ and push $N_{J_i}$ to $Q$.

4° If $Q = \phi$ then output $T_{gen}$, $T_{sink}$, $gen(t)$ for $t \in T_{gen} \cup T_{sink}$ and $PROV(N_{J_i})$ for $N_{J_i} \in SN_J$, and stop; otherwise go to 2°. □

In step 1° of the above algorithm, we mainly construct a transition list $L_s$ to be used to search subnet chains (the sequential activation transduction components) from the bottom of the Petri net model, where $SN_J$ denotes a set of subnets to be generated, and $T_{gen}$ and $T_{sink}$ denote the transitions that induce generation of these subnets. In 2°, for an obtained subnet $N_{J_i}$, we find the transitions $T_e$

that provide tokens to the enzyme places in it, where $L_e$ is used to update $L_s$ and is so constructed as to avoid repeated appearance in $L_s$, and $PROV(N_{J_i})$ expresses the set of transitions that provide tokens to its enzyme places. In 3°, we compute all the elementary T-invariants that are determined by a transition $t$ taken out from $L_s$ and get all the corresponding subnets, where $gen(t)$ indicates the subnets derived from the transition $t$, $SN_J$ is updated by all the obtained subnets. $Q$ is so updated that the subnets pushed to it cannot appear more than once. It is not difficult to confirm that the time complexity to perform the algorithm is $O((|T_{gen}|+|T_{sink}|)(K_m|T||P|^2+LP_s))$, where $K_m = \max\{|gen(t)|\ |t \in T_{gen} \cup T_{sink}\}$ and $LP_s$ is the time complexity of linear programming.

By applying the following operations to the result of algorithm ≪**Searching Activation Transduction Components**≫, $gen(t)$ and $PROV(N_{J_i})$, we can schematize the connection relations between the subnets that correspond to activation transduction components in signaling pathways.

(1) Do $T_s \leftarrow T_{sink}$. For each $t \in T_s$, do $SN \leftarrow gen(t)$ and draw arrows from all subnets in $SN$ to $t$. Do $T_s \leftarrow \phi$.
(2) For each subnet $N \in SN$, do the followings:
    (i) do $T'_s \leftarrow PROV(N)$, $T_s \leftarrow T_s \cup T'_s$;
    (ii) for each $t \in T'_s$, draw arrows from $t$ to $N$.
(3) Do $SN \leftarrow \phi$. For each $t \in T_s$, do the followings:
    (i) do $SN' \leftarrow gen(t)$ and $SN \leftarrow SN \cup SN'$;
    (ii) for each $N \in SN'$, draw arrows from $N$ to $t$.
(4) Do $T_s \leftarrow \phi$. If $SN \neq \phi$ goto (2); otherwise stop.

Figure 4 illustrates an example of connection relations among subnets by doing the above operations. Note that, hollow rectangle transitions denote the transitions in $T_{sink}$ and $T_{gen}$ to show the relationship between transitions and subnets.

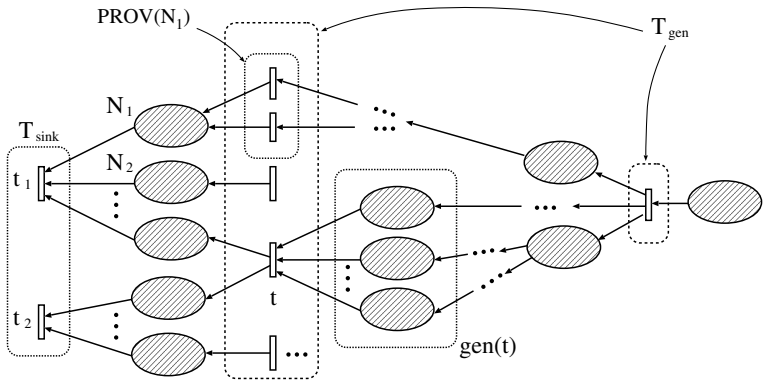



Fig. 4. An example of connection relations among subnets.

## 6. An Example

In this section, we give an example to demonstrate our modeling and analyzing method. The example is the signaling pathway mediated by TPO that is a cytokine regulating hematogenesis and production of hematoblast. TPO signals its growth regulating effects to the cell through several major pathways including JAK/Stat, MAPK as shown in Fig. 5.[24–26]

(1) Tyrosine phosphorylation of Jak2 in membrane proximal domain of TPO receptor activates JAK/Stat pathways consisting of the activation of STATs, STATs' homo- and hetero- dimerization, translocating to nucleus, where they modulate expression of target genes.
(2) Ras-MAPK pathways are activated by recruiting Shc in a membrane distal domain of TPO receptor. Grb2 is activated as an adaptor protein by binding to Shc and triggers subsequent activations of Sos, Ras, Raf-1, MEK, ERK, and RSK. Activated RSK translocates to nucleus and activates CREB that will bind to specific area of DNA to promote the transcription of genes.

We model the TPO signaling pathway to a Petri net model as shown in Fig. 6 based on the modeling rules, whose incidence matrix $C$ is rewritten to an integer matrix as follows:

$$C = \begin{pmatrix} -2 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 2 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & -2 & 2 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & -2 \end{pmatrix}$$

Note that, $\alpha(p_i, t_d)$ of the example is set to 0.5. In the following, we demonstrate how the algorithm is carried out.

First, we do step 1°, i.e. $L_s \leftarrow t_{20} \cdot t_{39} \cdot t_{95}$, $SN_J \leftarrow \phi$, $T_{sink} \leftarrow \{t_{20}, t_{39}, t_{95}\}$, $T_{gen} \leftarrow \phi$ and $Q \leftarrow \phi$. Then do 2°, $Q = \phi$ does not satisfy the condition $Q \neq \phi$ and goto 3°. Do 3°, since $L_s \neq \phi$ is not satisfied, take out a transition $t = t_{20}$ from the beginning of $L_s$. Then do $gen(t_{20}) \leftarrow \phi$ and $CJ_i = 0$ with $J_i(t_{20}) > 0$ by applying $\ll$**Searching Basic-Feasible Solution with** $x_s > 0 \gg$,

$$\begin{pmatrix} -2 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 2 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & -2 & 2 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{93} \\ x_{94} \\ x_{95} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$x_1, x_2, \ldots, x_{94}, x_{95} \geq 0.$$

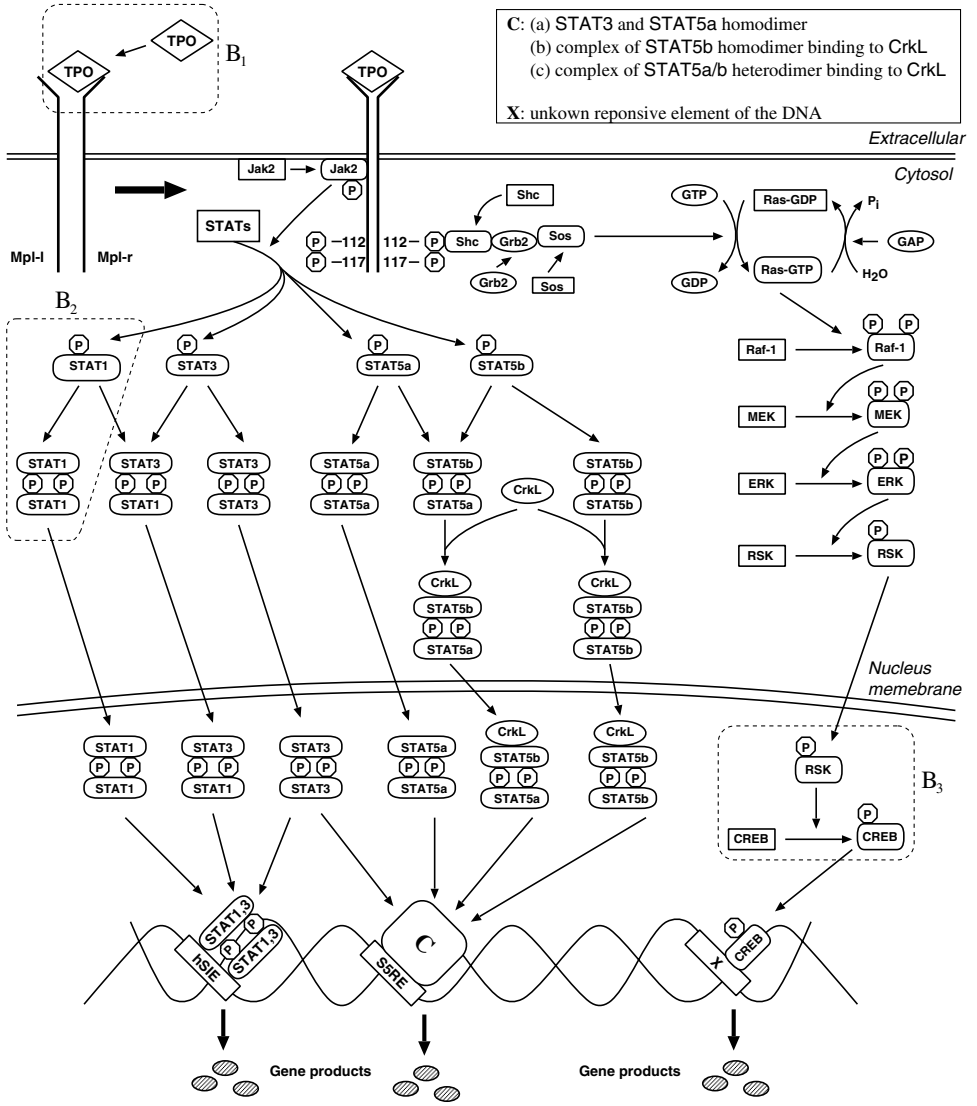



Fig. 5. TPO signaling pathway. The parts $B_1$, $B_2$, and $B_3$ surrounded by dashed lines are the three reaction examples, reactions of association, homodimerization and enzymic activation, which are the blocks **I**, **VII**, and **XI** of Fig. 2, respectively. Corresponding Petri net models of $B_1$, $B_2$, and $B_3$ are given in Fig. 6 by the dashed-line-surrounded parts $B_1'$, $B_2'$, and $B_3'$, respectively.

and obtain the elementary T-invariants $\{J_1, J_2, J_3\}$ as shown in Fig. 7. The table in Fig. 7 summarizes all elementary T-invariants of Petri net in Fig. 6, in which we only show the non-zero transitions due to the space limitation. For each elementary T-invariant, do the followings (here, we only explain the case of $J_1$): (i) obtain corresponding subnet $N_{J_1}$; (ii) do $gen(t_{20}) \leftarrow \{N_{J_1}\}$; (iii) since $N_{J_1} \notin SN_J$ is satisfied,

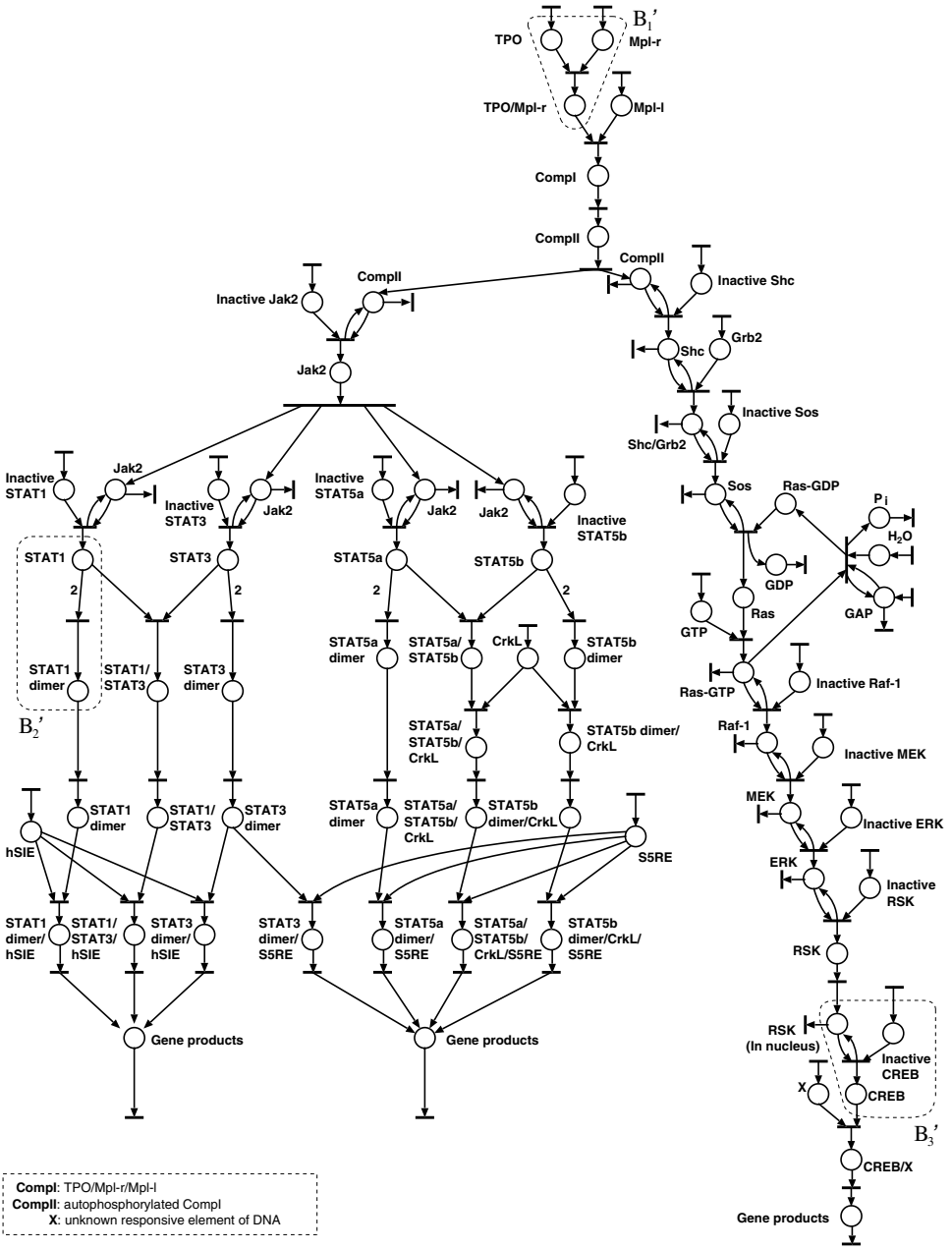



Fig. 6. Petri net model of the TPO signaling pathway shown in Fig. 5. Parts $B_1'$, $B_2'$, and $B_3'$ correspond to $B_1$, $B_2$, and $B_3$ of Fig. 5, respectively. $B_2'$ reflects the complex formation of STAT1 making a homodimer from two monomers, and so the arc-weight is set to 2.

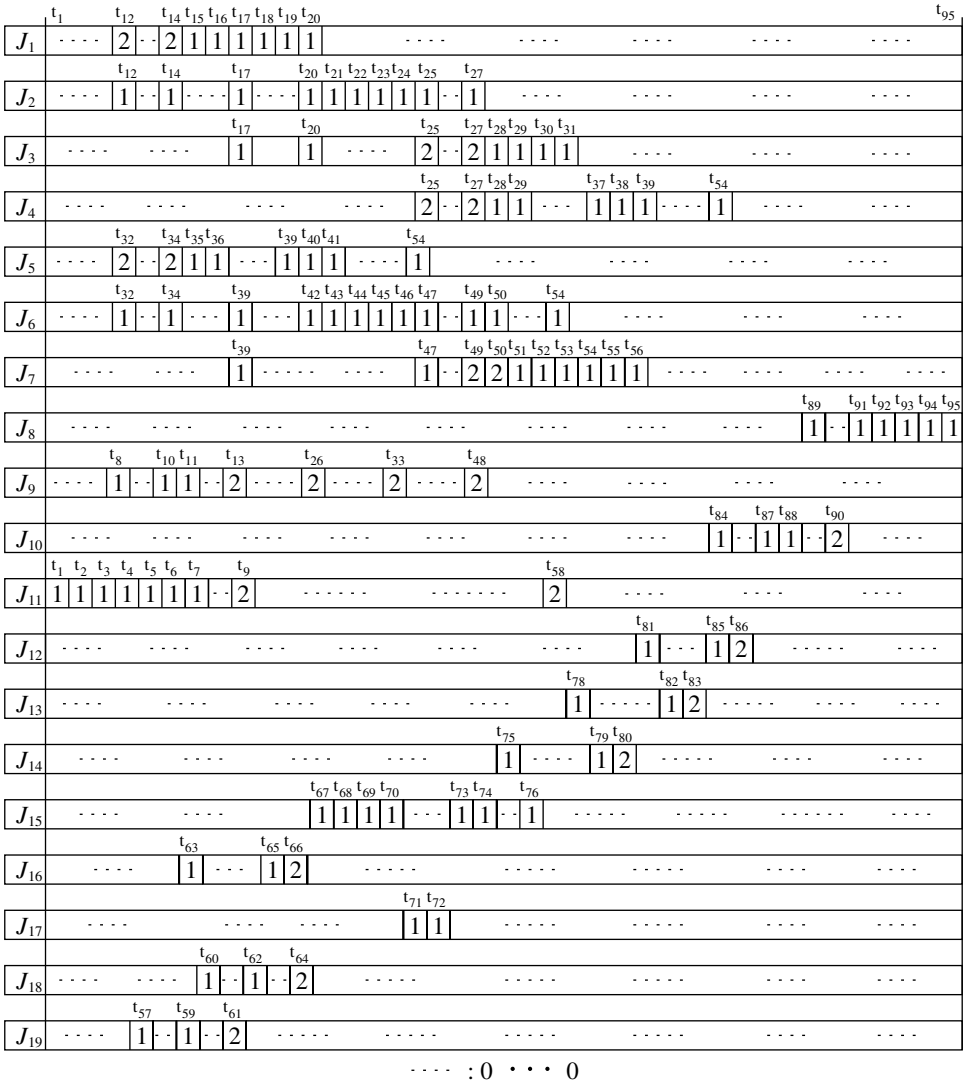



Fig. 7. All elementary T-invariants of the Petri net in Fig. 6.

do $SN_J \leftarrow \{N_{J_1}\}$ and push $N_{J_1}$ to $Q$. In 4°, $Q \neq \phi$ does not satisfy $Q = \phi$ and go to 2°.

In 2°, since $Q \neq \phi$ satisfies the condition, pull a subnet $N_{J_1}$ from $Q$ and do the followings: (i) $P_e \leftarrow \{p_{11}\}$, $T_e \leftarrow \{t_{11}\}$; (ii) do $L_s \leftarrow t_{39} \cdot t_{95} \cdot t_{11}$, $T_{gen} \leftarrow \{t_{11}\}$ and $PROV(N_{J_1}) = \{t_{11}\}$. Do 3°, since $L_s = \phi$ is not satisfied, take out $t = t_{39}$ from $L_s$ and do $gen(t_{39}) \leftarrow \phi$. Obtain all the elementary T-invariants $\{J_4, J_5, J_6, J_7\}$, corresponding subnets $\{N_{J_4}, N_{J_5}, N_{J_6}, N_{J_7}\}$, $gen(t_{39}) \leftarrow \{N_{J_4}, N_{J_5}, N_{J_6}, N_{J_7}\}$ and $SN_J \leftarrow \{N_{J_1}, N_{J_2}, N_{J_3}\} \cup \{N_{J_4}, N_{J_5}, N_{J_6}, N_{J_7}\}$. Push $N_{J_4}, N_{J_5}, N_{J_6}, N_{J_7}$ to $Q$

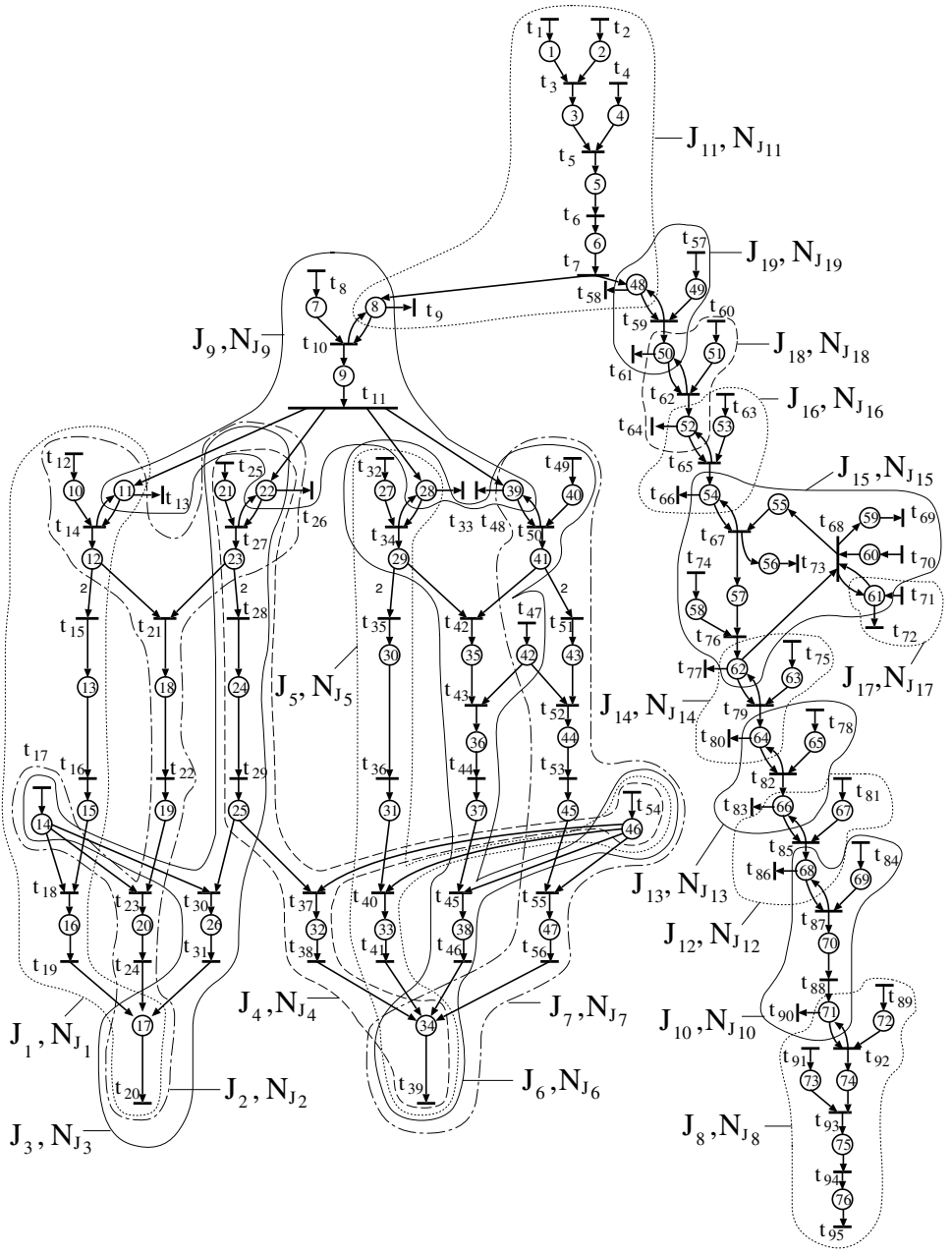
Table 1. The results $T_{sink}$, $T_{gen}$, $gen(t)$, and $PROV(N_{J_i})$ by performing proposed algorithm.

| $T_{gen}$ | $\{t_7, t_{11}, t_{59}, t_{62}, t_{65}, t_{71}, t_{76}, t_{79}, t_{82}, t_{85}, t_{88}\}$ | | |
|---|---|---|---|
| $T_{sink}$ | $\{t_{20}, t_{39}, t_{95}\}$ | | |
| $gen(t_7)$ | $\{N_{J_{11}}\}$ | $PROV(N_{J_1})$ | $\{t_{11}\}$ |
| $gen(t_{11})$ | $\{N_{J_9}\}$ | $PROV(N_{J_2})$ | $\{t_{11}\}$ |
| $gen(t_{20})$ | $\{N_{J_1}, N_{J_2}, N_{J_3}\}$ | $PROV(N_{J_3})$ | $\{t_{11}\}$ |
| $gen(t_{39})$ | $\{N_{J_4}, N_{J_5}, N_{J_6}, N_{J_7}\}$ | $PROV(N_{J_4})$ | $\{t_{11}\}$ |
| $gen(t_{59})$ | $\{N_{J_{19}}\}$ | $PROV(N_{J_5})$ | $\{t_{11}\}$ |
| $gen(t_{62})$ | $\{N_{J_{18}}\}$ | $PROV(N_{J_6})$ | $\{t_{11}\}$ |
| $gen(t_{65})$ | $\{N_{J_{16}}\}$ | $PROV(N_{J_7})$ | $\{t_{11}\}$ |
| $gen(t_{71})$ | $\{N_{J_{17}}\}$ | $PROV(N_{J_8})$ | $\{t_{88}\}$ |
| $gen(t_{76})$ | $\{N_{J_{15}}\}$ | $PROV(N_{J_9})$ | $\{t_7\}$ |
| $gen(t_{79})$ | $\{N_{J_{14}}\}$ | $PROV(N_{J_{10}})$ | $\{t_{85}\}$ |
| $gen(t_{82})$ | $\{N_{J_{13}}\}$ | $PROV(N_{J_{11}})$ | $\{\phi\}$ |
| $gen(t_{85})$ | $\{N_{J_{12}}\}$ | $PROV(N_{J_{12}})$ | $\{t_{82}\}$ |
| $gen(t_{88})$ | $\{N_{J_{10}}\}$ | $PROV(N_{J_{13}})$ | $\{t_{79}\}$ |
| $gen(t_{95})$ | $\{N_{J_8}\}$ | $PROV(N_{J_{14}})$ | $\{t_{76}\}$ |
| | | $PROV(N_{J_{15}})$ | $\{t_{65}, t_{71}\}$ |
| | | $PROV(N_{J_{16}})$ | $\{t_{62}\}$ |
| | | $PROV(N_{J_{17}})$ | $\{\phi\}$ |
| | | $PROV(N_{J_{18}})$ | $\{t_{59}\}$ |
| | | $PROV(N_{J_{19}})$ | $\{t_7\}$ |

sequentially. In 4°, $Q \neq \phi$ does not satisfy $Q = \phi$ and go to 2°. In this way, do the steps sequentially until $Q = \phi$, then output $T_{gen}$, $T_{sink}$, $gen(t)$ for $t \in T_{gen} \cup T_{sink}$ and $PROV(N_{J_i})$ for $N_{J_i} \in SN_J$, and stop the algorithm. All the output of algorithm are shown in Table 1, and Fig. 8 shows all the subnets of Petri net model in Fig. 6.

By applying the operations proposed above to the results $T_{gen}$, $T_{sink}$, $gen(t)$, and $PROV(N_{J_i})$ of algorithm ≪**Searching Activation Transduction Components**≫, we can draw corresponding arcs between subnets and transitions of $T_{sink}$ and $T_{gen}$. And finally schematize the connection relations between all the subnets that correspond to activation transduction components in signaling pathways as shown Fig. 9, in which there are 19 subnets $\{N_{J_1}, N_{J_2}, \cdots, N_{J_{18}}, N_{J_{19}}\}$ obtained from elementary T-invariants shown in Fig. 7.

Based on connection relations, each chain consisting of enzymic activation processes can be traced from the source subnets on the right side in Fig. 9, e.g. searching from the top subnet $N_{J_{11}}$, the transition $t_7$ in $N_{J_{11}}$ fires to provide tokens to the subnet $N_{J_9}$ and $N_{J_{19}}$, and the ligand-receptor complex (CompII for short) is activated. Once the enzyme of CompII is activate, the enzyme Shc in $N_{J_{19}}$ will be activated continuously by the firing of transition $t_{59}$ to provide tokens to the subnet $N_{J_{18}}$. Activated enzyme Shc induces the activations of Shc/Grb2, Sos, Ras, Raf-1, MEK, ERK, RSK and CREB in turn by firing corresponding transitions to provide tokens to the subnets. In this way, the chains of enzymic activation processes from the extracellular stimulus to the DNA nucleus can be obtained.




① • • • ⑯ : the label of a place is written in the place.

Fig. 8. All subnets obtained from elementary T-invariants as shown in Fig. 7 of Petri net model.

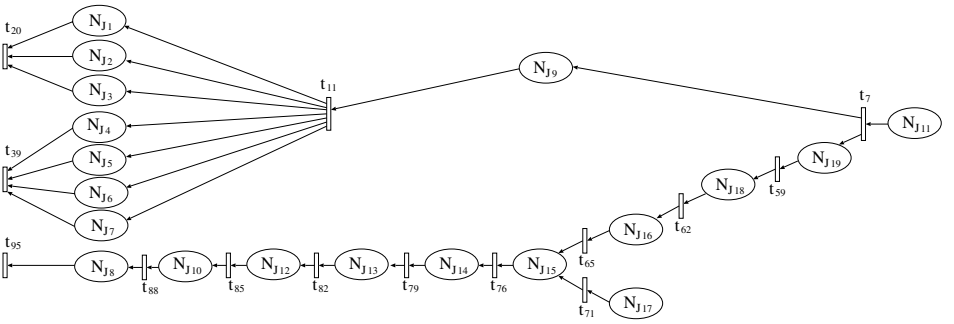



Fig. 9. Connection relations between all the subnets corresponding to activation transduction components in TPO signaling pathways.

## 7. Concluding Remarks

We have proposed a methodology to model and analyze signaling pathways by using Petri net. We have first given a modeling method based on Petri net by taking notice on the molecular interactions and mechanisms. Then we have introduced a new notion "activation transduction component" in order to describe an enzymic activation process of reactions in signaling pathways and shows its correspondence to a so-called elementary T-invariant in the Petri net model. Further, we have designed an algorithm to effectively find basic enzymic activation processes by obtaining a series of elementary T-invariants in the Petri net models. The obtained results from the algorithm are used to schematize the connection relations between the subnets that correspond to activation transduction components in signaling pathways. Finally, we have demonstrated how our method is practically used in modeling and analyzing signaling pathway mediated by TPO signaling pathway as an example. The main contributions are that:

(1) signaling pathways can be described consistently with the Petri net models of molecular interactions in Fig. 2, which enables biologists to intuitively understand the intrinsic structure and features of signaling pathways;
(2) the key enzymic activation processes in signaling pathways can be explicitly expressed by the graph obtained from our algorithm as shown in Fig. 9, which gives us a new insight into the architecture of signaling pathways to grasp their structural and behavioral properties.

In the future works, we will aim to: (i) improve our algorithm to analyze Petri net models including inhibitor arcs; and (ii) develop our current method further to analyze more complicated models and investigate the related properties.

## Acknowledgments

We are grateful to the anonymous reviewers for their valuable hints and suggestions. This research was partially supported by Grant-in-Aid for Scientific Research

on Priority Areas "Systems Genomics" from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

## Appendix: Abbreviations

CREB: cAMP response element-binding protein
CrkL: Crk (CT10-regulated kinase)-like protein
ERK: extracellular signal-regulated kinase
Grb2: growth factor receptor binding protein 2
hSIE: human sis-inducible element
Jak: Janus kinase
MEK: MAPK/ERK kinase
Mpl: myeloproliferative leukemia protein
Raf-1: v-raf-1 murine leukemia viral oncogene homolog 1
Ras: v-Ha-ras Harvey rat sarcoma viral oncogene homolog
RSK: 90-kDa ribosomal S6 kinase
S5RE: STAT5 responsive element
Shc: Src-homology collagen protein
Sos: Son of sevenless
STAT: signal transducers and activators of transcription

## References

1. Reddy VN, Liebman MN, Mavrovouniotis ML, Qualitative analysis of biochemical reaction systems, *Comput Biol Med* **26**(1):9–24, 1996.
2. Reddy VN, Mavrovouniotis ML, Liebman MN, Petri net representations in metabolic pathways, *Proc Int Conf Intell Syst Mol Biol* **1**:328–336, 1993.
3. Heiner M, Koch I, Will J, Model validation of biological pathways using Petri nets, demonstrated for apoptosis, *Biosystems* **75**(1–3):15–28, 2004.
4. Peccoud J, Stochastic Petri nets for genetic networks, *Medicine Sciences* **14**(8–9): 991–993, 1998.
5. Matsuno H, Fujita S, Doi A, Nagasaki M, Miyano S, Towards biopathway modeling and simulation, *Proc 24th ICATPN*, Lecture Notes in Computer Science **2679**:3–22, 2003.
6. Matsuno H, Tanaka Y, Aoshima H, Doi A, Matsui M, Miyano S, Biopathways representation and simulation on hybrid functional Petri net, *In Silico Biol* **3**(3):389–404, 2003.
7. Genrich HJ, Küffner R, Voss K, Executable Petri net models for the analysis of metabolic pathways, *International Journal on Software Tools for Technology Transfer* **3**(4):394–404, 2001.
8. Lee DY, Zimmer R, Lee SY, Park S, Colored Petri net modeling and simulation of signal transduction pathways, *Metabolic Engineering* **8**(2):112–122, 2006.
9. Voss K, Monika H, Koch I, Steady state analysis of metabolic pathways using Petri nets, *In Silico Biol* **3**(3):367–387, 2003.
10. Hofestädt R, Thelen S, Quantitative modeling of biochemical networks, *In Silico Biol* **1**(1):39–53, 1998.

11. Zevedei-Oancea I, Schuster S, Topological analysis of metabolic networks based on Petri net theory, *In Silico Biol* **3**(3):323–345, 2003.
12. Schuster S, Pfeiffer T, Moldenhauer F, Koch I, Dandekar T, Structural Analysis of Metabolic Networks: Elementary Flux Mode, Analogy to Petri Nets, and Application to Mycoplasma Pneumoniae, *German Conference on Bioinformatics 2000*, 115–120, 2000.
13. Schuster S, Dandekar T, Fell DA, Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering, *Trends Biotechnol* **17**(2):53–60, 1999.
14. Schuster S, Fell DA, Dandekar T, A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks, *Nature Biotechnol* **18**(3):326–332, 2000.
15. Peterson J, *Petri net theory and the modeling of systems*, Englewood Cliffs, NJ, Prectice-Hall, 1981.
16. Hofestädt R, A Petri net application to model metabolic processes, *Syst Anal Mod Simul* **16**:113–122, 1994.
17. Takai-Igarashi T, Mizoguchi R, Cell signaling networks ontology, *In Silico Biol* **4**(1):81–87, 2004.
18. Ge QW, Tanida T, Onaga K, Construction of a T-base and design of a periodic firing sequence of a Petri net, *Proc 8th Mathematical Programming Symposium, Japan*, 51–57, 1987.
19. Araki T, Tanida T, Watanabe T, Onaga K, A linear programming-based algorithm for successively enumerating elementary invariants of Petri nets, *Technical Report* of IEICE, 125–132 (CAS89–89), 1989.
20. Avis D, Fukuda K, A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra, *Discrete Comput Geom* **8**:295–313, 1992.
21. Martinez J, Silva M, A simple and fast algorithm to obtain all invariants of a generalized Petri nets, *Proc Second European Workshop on Application and Theory of Petri Nets, Informatik-Fachberichte 52*, Springer-Verlag, pp. 301–310, 1982.
22. Takano K, Taoka S, Yamauchi M, Watanabe T, Experimental evaluation of two algorithms for computing Petri net invariants, *IEICE Trans on Fundamentals* **E84-A**(11), 2871–2880, 2001.
23. Ge QW, Fukunaga T, Nakata M, On generating elementary T-invariants of Petri nets by linear programming, *Proc ISCAS2005*, pp. 168–171, 2005 (available online: *http://genome.ib.sci.yamaguchi-u.ac.jp/~pnp/papers/gqw-iscas2005.pdf*).
24. Geddis AE, Linden HM, Kaushansky K, Thrombopoietin: A pan-hematopoietic cytokine, *Cytokine Growth Factor Rev* **13**:61–73, 2002.
25. Schulze H, Ballmaier M, Welte K, Germeshausen M, Thrombopoietin induces the generation of distinct Stat1, Stat3, Stat5a and Stat5b homo- and heterodimeric complexes with different kinetics in human platelets, *Experimental Hematology* **28**:294–304, 2000.
26. Alexander WS, Thrombopoietin and the c-Mpl receptor: Insights from gene targeting, *IJBCB* **31**:1027–1035, 1999.

**Chen Li** received the B.S. from Zhejiang University, the People's Republic of China, in 2000 and M.E. from Graduate School of Education, Yamaguchi University, Japan, in 2004. She is currently a Ph.D. candidate in Graduate School of Science and Engineering, Yamaguchi University, Japan.

Her research interest includes Petri nets, program net theory and application of net theory to bioinformatics. She is a member of the Institute of Electronics, Information, and Communication Engineers (IEICE) and Japanese Society for Bioinformatics (JSBi).

**Shunichi Suzuki** received the B.S. and M.S. degrees in Information Science from Yamaguchi University, Japan, in 2004 and 2006, respectively. He has been with TIS System Service Inc. since April, 2006.

His research interest includes modeling and analyzing signaling pathways with Petri nets. He is a member of Japanese Society for Bioinformatics (JSBi).

**Qi-Wei Ge** received the B.E. from Fudan University, the People's Republic of China, in 1983, M.E. and Ph.D. from Hiroshima University, Japan, in 1987 and 1991, respectively. He was with Fujitsu Ten Limited from 1991 to 1993. He was an Associate Professor at Yamaguchi University, Japan, from 1993 to 2004. Since April of 2004, he has been a Professor at Yamaguchi University, Japan.

His research interest includes Petri nets, program net theory and combinatorics. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE), the Institute of Information Processing Society of Japan (IPSJ) and the Institute of Electrical and Electronics Engineers (IEEE).

**Mitsuru Nakata** received the B.E., M.E., and Ph.D. from Fukui University, Japan, in 1992, 1994, and 1998, respectively. He was a Lecturer at Yamaguchi University from 1998 to 2004. Since April of 2004, he has been an Associate Professor at Yamaguchi University, Japan.

His research interest includes database system, semi-structured database, text processing and program net theory. He is a member of the Institute of Electronics, Information, and Communication Engineers (IEICE), the Institute of Information Processing Society of Japan (IPSJ), the Institute of Electrical and Electronics Engineers (IEEE), and Japan Society for Software Science and Technology.

**Hiroshi Matsuno** received the B.E. and M.E. in Electronics from Yamaguchi University in 1982 and 1984, respectively. He received the Ph.D. from Kyushu University in 1994. From 1984 to 1995, he worked at Yamaguchi Junior College and Oshima National College of Maritime Technology. In 1995, Dr. Matsuno joined Faculty of Science, Yamaguchi University, and he has been a Professor since May of 2005 at Yamaguchi University, Japan.

His current interest includes systems biology and wireless LAN communications. He is a member of the Institute of Electronics, Information, and Communication Engineers (IEICE), the Institute of Electrical and Electronics Engineers (IEEE), the Institute of Information Processing Society of Japan (IPSJ), and Japanese Society for Bioinformatics (JSBi).

**Satoru Miyano**, is a Professor of Human Genome Center, Institute of Medical Science, University of Tokyo. He received the B.S., M.S., and Ph.D. all in Mathematics from Kyushu University, Japan, in 1977, 1979, and 1984, respectively. His research group is developing computational methods for inferring gene networks from microarray gene expression data and other biological data, e.g. protein–protein interactions, and promoter sequences. The group also developed a software tool Cell Illustrator for modeling and simulation of various biological systems.

Currently, his research group is intensively working for developing the gene network of human endothelial cell by knocking down hundreds of genes. With these technical achievements, his research direction is now heading toward a creation of Systems Pharmacology. He received IBM Science Award and Sakai Special Award both in 1994. He is on the Editorial Board of Bioinformatics, Journal of Bioinformatics and Computational Biology, Theoretical Computer Science, and the Chief Editor of Genome Informatics. He is a member of Japanese Society for Bioinformatics (JSBi) and in the board of directors of International Society for Computational Biology (2006).