Commentary for PARSEDAE

Conversion of regulator-reaction equations into mathematical equations

A table of contents

1 Abstract

In order to convert the regulator-reaction equations into mathematical models, CADLIVE classifies the reactions into three layers: gene, protein, and metabolic layers, and divides the conversion process into two stages, the first conversion (ordinary transcription and translation equations = TT, conventional mass action = CMA, general mass action = GMA, simplified Michaelis-Menten equations = MM), and the second conversion (differential and algebraic equations = DAEs, S-system). From the standpoint of mathematical conversions, the applied mathematical conversion strongly depends on the layer that the regulator-reaction equations belong to. At the first stage, both gene and protein networks employ TT and CMA, whereas the metabolic network uses GMA or MM. In the gene layer, since various molecules such as proteins, amino acids, nucleic acids, and RNAs act in concert for transcription and translation, it is difficult to mathematically describe such reactions based on their concrete molecular mechanism. The use of TT is a rational choice for taking in gene regulations within a cell. In the conversion of metabolic networks into GMA or MM, the concentrations of enzyme-metabolite complexes are cancelled compared with those of metabolites, because the former are far less than the latter. By contrast, in the protein layer, many proteins function in a complex or modified form, thus it is not practical to cancel the concentrations of the active complexes or modified ones. Thus, the use of CMA and TT describes the protein signal transduction pathways.

The problem for CMA is that it is often so stiff that it is hard to solve them numerically, because the rate constants of biochemical reactions and some molecular concentrations show huge differences in the order of their values. CMA also generates lots of kinetic parameters. At the second stage, using the Two-Phase Partition method (TPP) converts CMA into DAEs in order to overcome those problems of CMA. TPP substitutes algebraic equations for stiff differential equations. Applying TPP to conversion of the gene-protein layer (CMA with TT) reduces not only the stiffness but also decreases the number of kinetic parameters. As the other conversion, the differential equations of TT, CMA, GMA, and MM are converted into S-system at the second stage. The use of S-system enables one to carry out the sensitivity and stability analysis in symbolic form. S-system is also a standard formula available for simulating dynamic behaviors without insisting on the detailed molecular mechanism.

The parsedae module converts the regulator-reaction equations (sanac file), which are constructed by the CADLIVE editor, into the mathematical equations (checkdae file) including TT, CMA, GMA, S-system, and MM. The checkdae file employs the indexes of species' names so that one can understand the model or edit it manually. One is allowed to edit the checkdae file directly according to the instruction. The checkdae file is written in the text format.

1.1 Input data

The parsedae module inputs the "sanac" file (described later) that mainly consists of two kinds of data. One is the species (molecules) with their associated attributes. The other is the regulator-reaction equation that determines the interaction among the species.

1.2 Mathematical formulas

The following mathematical formulas have been employed.

CMA（Conventional Mass Action）is applied to the protein layer, focusing on the formation of complexes and modified proteins.

TT (ordinary transcription-translation equations) is applied to the gene layer.

GMA（General Mass Action）is applied to the metabolic layer, neglecting the detailed mechanisms including the formation of enzyme-metabolite complexes

MM (simplified Michaelis-Menten equations) is applied to the metabolic layer, canceling the concentrations of enzyme-metabolite complexes. MM is obtained by simplifying usual Michaelis-Menten equations.

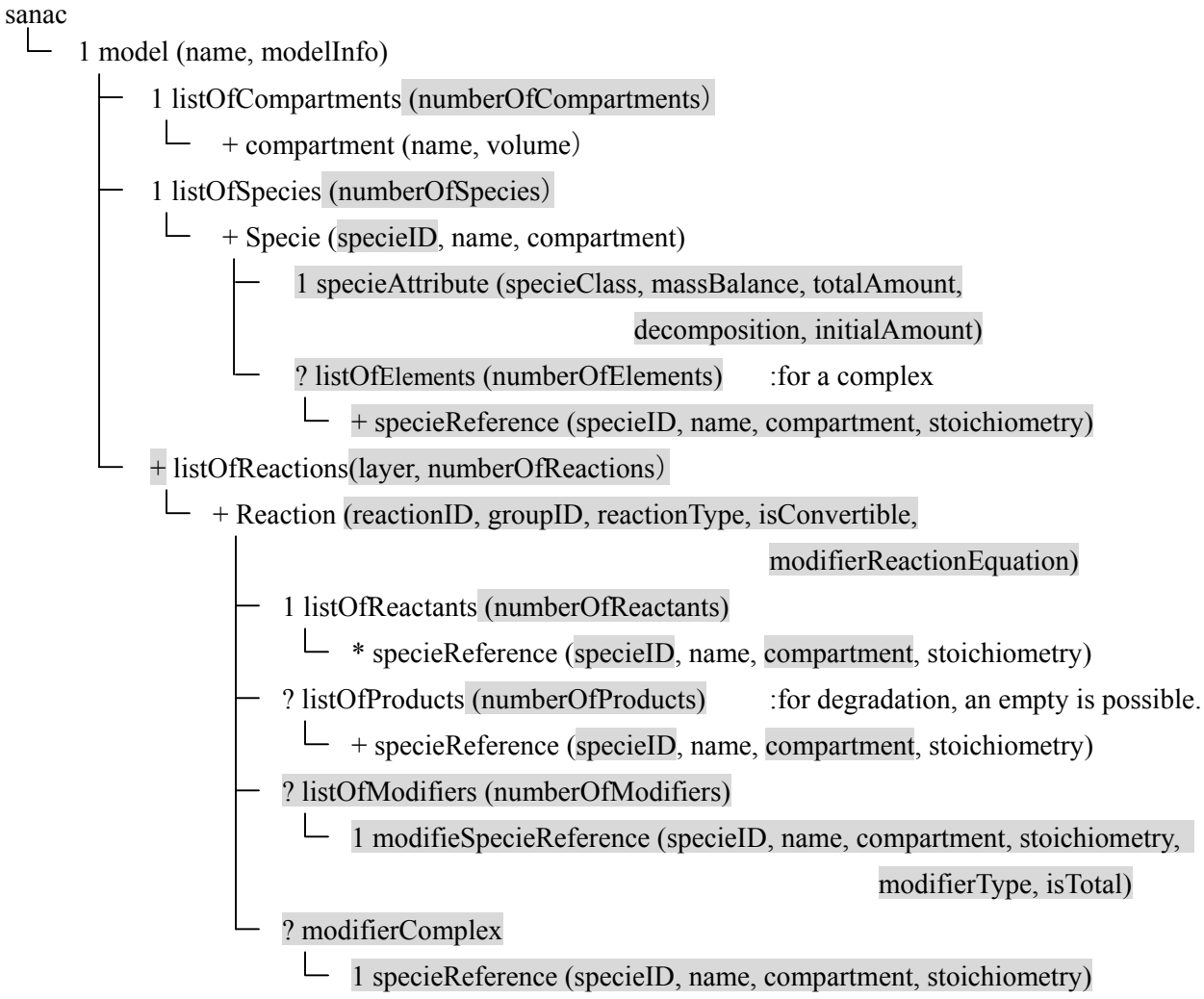
DAEs (differential and algebraic equations) are applied to the gene-protein layer using the two-phase partition (TPP) method.

S-system

2 Input data

2.1 Structure of input data

Systems Biology Markup Language (SBML) is one of the most advanced markup language that describes biochemical networks at concrete molecular interaction level. We extend SBML level 2 to establish the sanac format that includes all the information necessary for mathematical conversion and dynamic simulation. Although the original sanac file contains various data including the coordinates of the species, which are used for drawing a biochemical network map by the CADLIVE GUI editor, we omit the data that are not necessary for mathematical conversion and dynamic simulation.



*The net regions are the extension from SBML level 2.

Legend:

1: an element that appears once. *: an element that repeats more than zero.

+: an element that repeats more than one. ?: an element that appears once or zero.

t: an element that has text in content. (): provided as attributes

The major extensions are as follows:

- The attribute of "specieID" refers from the lists of reaction equations or species.
- The element of <listOfElements> shows the components of the complexes, which are defined as "specieReference".
- The attribute of "layer" is added to the element of <listOfReactions>, whereby a biochemical network can be divided into three layers, metabolic (metabolic), protein signal transduction (protein), and gene regulatory networks (gene).
- The attribute of "modifierReactionEquation" is added to the element of <reaction>.
- The elements of <listOfModifiers> and <modifierComplex> are added to the element of <Reaction>, which are necessary for mathematical conversion. The modifierComplex corresponds to

substrate-enzyme complex in ordinal enzyme reactions.

- The attribute of "reactionType" is added to distinguish regulator-reaction equations, which are indispensable for mathematical conversion.
- The attribute of "modelInfo" is added to comment the model.

2.2 Species naming rules

The following characters are allowed to name species

Initial character

・upper and lower alphabets	[a-z]｜[A-Z]
・number	[0-9]
・left parentheses	(

Second characters

・upper and lower alphabets	[a-z]｜[A-Z]
・underbar	_
・number	[0-9]
・left and right parentheses	( )
・colon	:
・hyphen	_
・period	.

The period may be used for indicating a decimal point.

2.3 Complex

A complex is defined by listing its components as specieReference in <listOfElements>. The use of "stoichiometry" of "specieReference" is able to determine the ratio of components. Not only a monomer/modified but also a complex is allowed to list as the components of the complex, thus users are able to define a higher complex by using complex, modified, or monomer. It is required to keep the naming rules. The employed species are required not to show reference circulation and to be defined in the model.

The element of <listOfElements> is used:

1: to assign the species decomposition to the respective differential equations,

2: to generate the mass balance equations and the differential equations for the total amounts of the species whose class is protein or modified in the TPP conversion,

3: to calculate the total concentrations of promoters and enhancers in the "gene1" reaction.

2.4 Species Attribute

The species is defined by three indispensable attributes:

- specieID :The integer unique to the combination of species' name with its compartment
- name :User can name the species arbitrarily.
- compartment :Location of the species

The other attributes are provided as shown in Table 1.

Table 1 Attributes of species (molecules)

Attribute	Content	Value
specieID	the integer unique to the combination of species' name with its compartment	
Name	name of species	named by user
Compartment	location of species	see text
SpecieClass	kinds of species	DNA_gene
		DNA_promoter
		DNA_enhancer
		DNA_others
		RNA
		protein
		metabolite
		environmental_factor
		ion_signal
		complex
		modifier_complex
		modified
		small_molecule
		text_option
		others
massBalance	total mass balance equation is required.	on
		off
totalAmount	total concentration	constant
		variable
decomposition	spontaneous degradation in vivo	on
		off
initialAmount	initial concentration of species	real value

The attribute of "specieClass" indicates the kinds of species, such as DNAs, RNAs, proteins, complexes, metabolites, promoter, and enhancer, and "compartment" shows which compartment the species is located. This attribute of "specieClass" classifies the molecules by their chemical features. We refer a few key values here. Binding multiple molecules without any stoichiometric change generates "complex", whose components are listed in <listOfElements> in the sanac file. In this system, the complex is named by joining its associated elements with a colon (:), *e.g.*, PER:dCLK. "modified" indicates the molecules that are chemically modified through modification processes such as phosphorylation and acetylation, which are accompanied with stoichiometric changes. The value of "modified" is named by joining the modifying molecule to the modified

one using a hyphen (-), *e.g.*, a phosphorylated PER is expressed as PER-P. For example, the molecules of dCLK:Enhancer, s32:DnaK, and PER-P:dCLK are "complex", PER-P is "modified". The value of "modifierComplex" means the active complex that consists of an enzyme and substrates, which is automatically generated when an enzyme reaction is defined.

The attribute of "totalAmount" determines whether the total concentration for the species is a constant or variable. When "totalAmount" of protein is "variable", the protein is synthesized and decomposed according to the transcription and translation equations. The total concentration of the species with "totalAmount = constant" is named by adding "T" to the head of species' name, whereas species' name without "T" indicates the free molecule. The total concentration of the protein with "totalAmount = constant" is constant, because the synthesis and degradation of the protein are assumed not to occur.

In converting CMA to DAEs, the attribute of "massBalance" determines whether total mass balance equations are made only for the species with "specieClass = monomer, modified, promoter, or enhancer", not for the complex. The TPP applies the quasi-steady state or rapid equilibrium approximation to the differential equations for the species with "specieClass = complex". For the species with "massBalance = on", the mass balance equations are generated that sums all the complexes containing the species as components. The attribute of "decomposition" indicates if the species is degraded *in vivo* spontaneously.

2.5 Definition of regulator-reaction equations

A model has multiple elements of <listOfReactions>, which has two indispensable attributes of "layer". The attribute of "layer" has three values, "metabolic", "protein", and "gene", which indicate where reactions occur. The metabolic layer indicates metabolic networks, the protein layer signal transduction pathways through protein interactions, and the gene layer transcription and translation. The mathematical formulas are basically determined by the layer that reaction belongs to.

- layer :determining the layer that reactions occur, which consists of "metabolic", "protein", and "gene".
- numberOfReactions :the number of reactions

The attribute of "isConvertible" in the element of <reaction> determines if the regulator-reaction equation can be converted into mathematical equations. The parser handles all the mechanistic models including "transcription" and "translation". The reactions with "isConvertible = off" are excluded from the mathematical model. Consequently, the Simulator converts all the reactions with "isConvertible = on".

At the first stage that generates ordinary differential equations, such as TT, CMA, GMA, and MM, the attributes of "layer" in the elements of <listOfReactions>, "reactionType" and "groupID" of the elements of <Reactions>, and "modifierType" in the element of <listOfModifier> produce the differential equations for mathematical models. The attribute of "layer", which has three values, "metabolic", "protein", and "gene" to indicate where reactions occur, determines the basic mathematical formula. The metabolic layer employs GAM or MM, the protein layer indicating signal transduction pathways through protein interactions adopts

CMA, and the gene layer regarding the regulations of transcription and translation uses TT. The attribute of "reactionType", which consists of the various kinds of regulator-reaction equations shown in Table 2, determines which differential equation is employed to each type of reaction. The element of <listOfModifier> presents concrete regulator information, where the attribute of "modifierType" indicates the type of modifiers, "enzyme", "activator", or "inhibitor". The value of "enzyme" indicates general enzyme reactions, whereas "activator" or "inhibitor" is used for "reactionType = transcription". In transcription, an activator binds to an enhancer to enhance the transcription of a gene, whereas an inhibitor binds to a promoter to repress the transcription of a gene. The attribute of "groupID" is required to group a series of reactions, where the multiple regulators with the same groupID act on the identical reaction. They are often used for the "gene1" transcription. For example, when multiple transcription regulation factors (TRFs) act on the transcription of a gene, these regulator-reaction equations with the same "groupID" generate one transcription regulation equation with the multiple TRFs.

In order to express the spontaneous degradation of proteins and mRNAs *in vivo*, the attribute of "decomposition" in the element of <specieAttribute> is employed. Generally, proteins and RNAs constitutively degrade *in vivo* due to some factors including proteases ("decomposition = on"), whereas DNAs do not degrade ("decomposition = off"). When the species consists of multiple components, the attribute of "decomposition" has to be searched recursively, while checking the decomposition of each component is on or off.

In the second stage, the attributes of "specieClass" "massBalance", and "totalAmount" in the element of <specieAttribute> play intrinsic roles in converting the ordinary differential equations into DAEs. The attribute of "specieClass" indicates the kinds of species, such as DNAs, RNAs, proteins, complex, and metabolites. This attribute of "specieClass" determines which components are required to make mass balance equations for each "monomer", "modified", "enhancer", or "promoter" species. Binding multiple molecules without any stoichiometric change generates "complex", whose components are listed in <listOfElements>. The attribute of "massBalance" determines whether total mass balance equations are made only for the species with "specieClass = monomer, modified, enhancer, or promoter", not for the complex. The TPP applies the quasi-steady state or rapid equilibrium approximation to the differential equations for the species with "specieClass = complex". For the species with "massBalance = on", the mass balance equations are generated that sums all the complexes containing the species as components.

In order to combine the gene-protein layers with the metabolic layer, the attribute of "isTotal" is presented to determine if a modifier acts as free species or as the total amount of the species. It is a critical attribute when the concentration of the enzyme is defined as a time-dependent variable in the metabolic layer. With "isTotal = true" for an enzyme, the total concentration of the enzyme is employed for GMA or MM. Conversely, with "isTotal = false" for an enzyme, the free concentration of the enzyme is used for the metabolic layers. This rule enables one to combine the three layers, *i.e.*, metabolic and gene regulatory networks.

Table 2 Values of the attribute of reactionType

Value
binding
binding_with_stoichiometric_changes
homo_association_or_modification
homo_association_or_modification_with_stoichiometric_changes
elimination
elimination_with_stoichiometric_changes
reversible_conversion
irreversible_conversion
reversible_conversion_regarding_multicomponent
irreversible_conversion_regarding_multicomponent
transport
option_transport
transcription
translation
degradation

*Explanatory note

In this system, one regulator-reaction equation is restricted to have one or zero modifier. In order to assign multiple modifiers for a reaction, the attribute of "groupID" is defined. The regulators with the same "groupID" act on the same reaction. Originally, "groupID" has been defined to visualize the reaction that multiple regulators act on the identical reaction.

The element of <reaction> consists of the elements of <listOfReactants>, <listOfProducts>, <listOfModifiers>, and <modifierComplex>.

- listOfReactants :defines reactant(s)
- listOfProducts :defines product(s), For "reactionType = degradation", an empty is possible.
- listOfModifiers :define modifier.
- modifierComplex :define the complex containing a modifier

The elements of <listOfReactants>, <listOfProducts>, <modifierComplex>, and <listOfElements> have to be referred from <specieReferences>, which consists of four attributes, "specieID", "name", "compartment", and "stoichiometry".

- specieID :
- name :
- compartment :cytoplasm, nucleoplasm, membrane, environment, …..
- stoichiometry :defined by real values. The default value is one.

The element of <listOfModifiers> must be referred from <modifierSpecieReference>, which is the element

of <specieReference> with the attributes of "modifierType" and "isTotal".

- modifierType :enzyme, activator, inhibitor
- isTotal :false, true

3 Conversion to mathematical models

3.1 TT (ordinary Transcription and Translation)

In the gene layer, transcription and translation occur in complicated manners, which involve a large number of nucleic acids, amino acids, RNAs and proteins. It is not practical to describe all possible reactions, and it is difficult to convert the transcription or translation (gene layer) into CMA based on their detailed molecular mechanism. Thus, we apply a general type of differential equations suitable for transcription and translation equations. For transcription, we conveniently divide two types: gene0 and gene1. For gene0, no transcription regulation factor involves transcription; for gene1, activators or inhibitors regulate the transcription.

3.1.1 Transcription without regulators (gene0)

Assuming that no gene regulation involves transcription, the transcription occurs at a constant rate, which is provided by:

Regulator-reaction equations: $gene(A) \rightarrow mRNA(A)$

Differential equations: $\frac{d[mRNA(A)]}{dt} = km[gene(A)]$ *km*: transcription rate constant

3.1.2 Transcription with regulators (gene1)

The activators or inhibitors that bind to enhancers or promoters regulate transcription. For example, the regulators, *W*, *X*, *Y*, and *Z* bind to regulatory regions, Pro1, Pro2, Enh1, and Enh2, to inhibit or activate the transcription, as follows:

$W : Pro1 \; -|> gene(A) \rightarrow mRNA(A)$
$X : Pro2$ -II $gene(A) \rightarrow mRNA(A)$
$Y : Enh1$ -II $gene(A) \rightarrow mRNA(A)$
$Z : Enh2$ -|> $gene(A) \rightarrow mRNA(A)$

Since these four reaction equations have the same "groupID", we convert them into the following differential equations.

$$\frac{d[mRNA(A)]}{dt} = k_m \cdot \left(\frac{[W : Pro1]}{[TPro1]} + \frac{[Z : Enh2]}{[TEnh2]} \right) \times \left(1 - \frac{[X : Pro2]}{[TPro2]} \right) \times \left(1 - \frac{[Y : Enh1]}{[TEnh1]} \right) \times [gene(A)]$$

where *km* is the transcription rate constant.

3.1.3 Translation

For example, the protein A is synthesized from *mRNA(A),* which is given by:

$$\frac{d[A]}{dt} = +kp[mRNA(A)],$$

where kp is the translation rate constant.

3.2 CMA（Conventional Mass Action）

Regulator-reaction equations can be divided into binding reactions and conversion reactions.

3.2.1 Binding reactions

For example, a binding reaction is provided by:

$$\alpha_1 A_1 + \alpha_2 A_2 + \ldots + \alpha_n A_n \underset{ks}{\overset{ka}{\rightleftarrows}} \alpha_1 A_1 : \alpha_2 A_2 : \ldots : \alpha_n A_n \quad ,$$

where α_i is the stoichiometry.

$$\frac{d[A_i]}{dt} = \alpha_i \left\{ -ka \prod_{j=1}^{n} [A_j]^{\alpha_j} + kd[\alpha_1 A_1 : \alpha_2 A_2 : \ldots : \alpha_n A_n] \right\} \qquad (\text{i} = 1 \sim \text{n})$$

$$\frac{d[\alpha_1 A_1 : \alpha_2 A_2 : \ldots : \alpha_n A_n]}{dt} = ka \prod_{j=1}^{n} [A_j]^{\alpha_j} - kd[\alpha_1 A_1 : \alpha_2 A_2 : \ldots : \alpha_n A_n],$$

where ka is the binding association rate constant, and kd is the dissociation rate constant. The amount of the corresponding species on the left and right sides has to be the same. The parsedae checks the stoichiometry of reactions. In addition, the parsedae module checks the components of the complex according to the rules. For the binding reaction equations:

$$A_1 + (A_2 : A_3) \underset{ks}{\overset{ka}{\rightleftarrows}} A_1 : (A_2 : A_3) \quad \text{and} \quad A_1 + (A_2 : A_3) \underset{ks}{\overset{ka}{\rightleftarrows}} A_1 : A_2 : A_3$$

are allowed, but the equation:

$$A_1 + A_2 + A_3 \underset{ks}{\overset{ka}{\rightleftarrows}} A_1 : (A_2 : A_3)$$

is not allowed, when the parsedae outputs the error.

3.2.2 Conversion reaction

The reaction:

$$\alpha_1 A_1 : \alpha_2 A_2 : \ldots : \alpha_n A_n \xrightarrow{kx} \beta_1 B_1 + \beta_2 B_2 + \ldots + \beta_m B_m$$

is converted into CMA:

$$\frac{d[B_i]}{dt} = \beta_i \, kx[\alpha_1 A_1 : \alpha_2 A_2 : \ldots : \alpha_n A_n] \quad (\text{i} = 1 \sim \text{n})$$

$$\frac{d[\alpha_1\ A_1 : \alpha_2 A_2 : \ldots : \alpha_n\ A_n]}{dt} = -kx[\alpha_1\ A_1 : \alpha_2 A_2 : \ldots : \alpha_n\ A_n],$$

where kx is the conversion rate constant.

3.2.3 Transport

The species A is transported from one compartment to another compartment (A')

$$\frac{d[A]}{dt} = -ktr[A], \qquad \frac{d[A']}{dt} = +ktr[A]$$

where *ktr* is the transport rate constant. Active transport with an enzyme is processed in the same manner as a conversion with an enzyme.

3.2.4 Spontaneous decomposition

Proteins and RNAs are generally degraded *in vivo*. Thus, we omit the regulator-reaction equations for spontaneous decomposition. Instead, we prepare the attribute of decomposition for species. With "decomposition = on" for a species, the parsedae module adds the spontaneous degradation term to the differential equation for the species, as follows:

$$\frac{d[A]}{dt} = -kpd[A],$$

where *kpd* is the degradation rate constant. For mRNA, *kmd* is defined as the degradation rate constant for mRNAs. Notice that the reaction type of "degradation" is different from the spontaneous decomposition.

When the components have complexes, the parsedae checks the attributes of decomposition recursively, while checking whether "decomposition" is "on" or "off". For example, when the decomposition of A, B, C, A:B, B:2D, (A:B):(2(B:2D)):(3(C:D)) is on, that of D, C:D is off, the differential equations for the respective components include the following terms :

i) d[A]/dt = - *kpd*[1] * [A]
ii) d[B]/dt = - *kpd*[2] * [B]
iii) d[C]/dt = - *kpd*[3] * [C]
iv) d[D]/dt = + 2**kpd*[5] * [B:2D] + 4**kpd*[6] * [(A:B):(2(B:2D)):(3(C:D))]
v) d[A:B]/dt = - *kpd*[4] * [A:B]
vi) d[B:2D]/dt = - *kpd*[5] * [B:2D]
vii) d[C:D]/dt = + 3 * *kpd*[6] * [(A:B):(2(B:2D)):(3(C:D))]
viii) d[(A:B):(2(B:2D)):(3C:D)]/dt = - *kpd*[6] * [(A:B):(2(B:2D)):(3(C:D))]

Since the decomposition for the species of (A:B):(2(B:2D)):(3(C:D)) is on, the decomposition terms are added to the respective differential equation (vii). Next, the decomposition of its components of A:B, B:2D, and C:D is checked. Since the decomposition of A:B, A, and B is on, no source term is added to the equations

(i,ii,v). The decomposition of B:2D is on, the decomposition term is not added to the respective differential equation (iv). However, since the decomposition of D is off, and the decomposition of B:2D releases 2*D, the source term for D is added to the equation (iv). Since C:D is not decomposed despite the decomposition of (A:B):(2(B:2D)):(3(C:D)), the source term for C:D is added to the equation (vii).

3.3 GMA（General Mass Action）

The enzyme reactions in the metabolic layer can be converted into GMA. Different from CMA, this conversion neglects the enzyme complex $E:S_1:S_2$, the differential equation for $E:S_1:S_2$ is not produced. For regulator-reaction equation:

$$E - o\ S_1 + S_2 > P_1 + P_2$$

For GMA,

$$\frac{dS_1}{dt} = -\ kxg_1 E^{f_1} S_1^{\ f_2} S_2^{\ f_3}$$
$$\frac{dS_2}{dt} = -\ kxg_1 E^{f_1} S_1^{\ f_2} S_2^{\ f_3}$$
$$\frac{dP_1}{dt} = +\ kxg_1 E^{f_1} S_1^{\ f_2} S_2^{\ f_3}$$
$$\frac{dP_2}{dt} = +\ kxg_1 E^{f_1} S_1^{\ f_2} S_2^{\ f_3}$$

If necessary, GMA produces the term for spontaneous decomposition as follows:

$$\frac{dS}{dt} = -\ kxgS^{f}\ .$$

3.4 MM（simplified Michaelis-Menten equations）

The metabolic equations can be converted into simplified Michaelis-Menten equations (MM), as follows.

The regulator-reaction equation:

$$E - o\ S_1 + S_2 > P_1 + P_2$$

is converted to :

$$\frac{dS_1}{dt} = -\ Q_1 E \frac{S_1}{Kmich_1 + S_1} \frac{S_2}{Kmich_2 + S_2}$$
$$\frac{dS_2}{dt} = -\ Q_1 E \frac{S_1}{Kmich_1 + S_1} \frac{S_2}{Kmich_2 + S_2}$$
$$\frac{dP_1}{dt} = +\ Q_1 E \frac{S_1}{Kmich_1 + S_1} \frac{S_2}{Kmich_2 + S_2}$$
$$\frac{dP_2}{dt} = +\ Q_1 E \frac{S_1}{Kmich_1 + S_1} \frac{S_2}{Kmich_2 + S_2}$$

If necessary, this conversion also produces the term for spontaneous decomposition as follows,

$$\frac{dS}{dt} = -Q\frac{S}{Kmich + S}.$$

3.5 DAEs

In order to solve the stiffness in CMA expressing a gene regulatory network, the two-phase partition method (TPP) has been proposed that divide biochemical reactions into two phases, binding and reaction phases, by applying rapid equilibrium approximation or quasi-steady state approximation to complex formation, which greatly reduces the number of kinetic parameters and remarkably accelerates the calculation speed. TPP is a powerful method that is indispensable for mathematical simulation of gene regulatory networks.

3.5.1 Steady-state approximation I

The two-phase partition method begins with CMA and TT.

(1) The differential equation for the complexes that have been generated through a binding reaction is set as zero, resulting in an algebraic equation. The generated equations are simplified. For example, when a + b + c + d = 0 and b – d = 0 is generated, the former is replaced by a + c = 0.

(2) For the species that do not involve binding reactions and those whose attribute of "massBalance" is off, their differential equations do not change.

(3) For the monomers that involve binding reactions and whose attribute of "massBalance" are on, the concentrations of the monomer and its derivative complexes are all added, producing the differential equations for their total concentration. For the species whose attribute of "massBalance" is off, they are not. For example, for the specie regarding the monomer A, the equations are made as follows:

dy[TA]/dt = differential equations for monomer A
+ Σ (differential equations for the complex that contains A as a component * the number of A that the complex contains)

(4) The mass balance equations for the species whose "specieClass" is "monomer" is provided by:

[TA] = [A] + Σ ([complex that contains A as a component] * the number of A that the complex contains)

(5) After simplifying the DAEs, the differential equations whose right hand sides become zero are omitted, and the species are registered as constant values.

The species are classified as follows:

A) Constant

The species whose attribute of "totalAmount" is constant, and whose "massBalance" is off.

The species whose differential equation is zero (dy/dt=0).

B) Variables of algebraic equations

The species whose class is "complex".

The species whose class is "monomer" and whose "massBalance" is on.

C) Variables of differential equations

The total concentration of the species with the attribute of monomer and modified. (excludes the species whose differential equation is zero (dy/dt=0).)

3.5.2 Steady-state approximation Ⅱ

The algebraic equations at the steady state are further simplified by setting *kpd* = 0.

3.5.3 Rapid equilibrium approximation

The algebraic equations at the steady state are further simplified by setting *kpd*, *kx* << *ka, kd*. In addition, the association constant of *Kb* is substituted for *ka/kd*. For example, the regulator-reaction equation:

mA + nB <-> mA:nB

is converted to the algebraic equation:

[mA:nB] = *Kb*^(m*n)*[A]^m*[B]^n.

3.5.4 Complexes in TPP

We illustrate in detail how TPP derives the differential equations for the total amount of specific species and their mass balance equations. The regulator-reaction equation:

A + 2B <—> A:2B

is converted into the CMA:

d[A]/dt = - *ka*[A][B]^2 + *kd*[A:2B] ①

d[B]/dt = - 2*ka*[A][B]^2 + 2*kd*[A:2B] ②

d[A:2B]/dt = + *ka*[A][B]^2 – *kd*[A:2B] ③.

We apply TPP to the CMA by setting the equation ③ as zero, producing the differential equations for the total concentrations (TA, TB) for the components of A and B. In order to make the differential equation for TB, we recursively search all the components of the complexes to find the species that contain B as a component, and add them to the differential equation. They are multiplied by the number of B contained in the complexes

d[TB]/dt = the right hand side of ②＋ the right hand side of ③ * 2

= Σ(the right hand sides of the CMA for the species that contains B as a component * the number of B that the species contains)

In the same manner, the mass balance equation for TB is produced as follows:

[TB] = [B] + [A:2B]*2

= Σ(the concentrations of the species that contain B as a component * the number of B that the species contains)

3.6 S-system

Ordinary differential equations are divided into the positive terms and negative terms:

$$\frac{dX_i}{dt} = V_i^+ - V_i^- , \qquad (i = 1,....,n,....n+m),$$

where $\mathbf{V_i^+}$ is the sum of positive terms, and $\mathbf{V_i^-}$ is the sum of negative terms. Generally, S-system is given by:

$$\frac{dX_i}{dt} = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} ,$$

where $X_1 \sim X_n$ are the dependent variables, whose values vary with time, and $X_{n+1} \sim X_{n+m}$ are the independent variables, whose values are fixed as constants.

3.7 Connection between layers

When DAEs express the gene-protein layers, and GMA or MM expresses the metabolic layer, dependent variables for the regulators that appear as modifiers in the metabolic layer correspond to dependent variables for the gene-protein layers. Table shows how to connect the corresponding variables between the gene-protein layer and the metabolic layer. The combination of the values of the attributes of "isToal", "massBalance" and the method for conversion determines the variables that appear as a modifier in the metabolic layer.

Table Connection between layers

Conversion method for Gene-Protein layer	massBalance	isTotal	The variable of E that appears as a modifier in the metabolic layer
CMA	on	true	Total concentration of enzyme: T_E.cyt
		false	y[E.cyt]
	off	true	Total concentration of enzyme T_E.cyt
		false	y[E.cyt]
DAEs (TPP)	on	true	y[TE.cyt]
		false	x[E.cyt]
	off	true	Total concentration of enzyme T_E.cyt
		false	y[E.cyt]

The total concentration of the enzyme E is the total amount of the monomer E and the complexes that contain the monomer E as a component. For example, when TE consist of E, E:A:B, E:C, the total concentration for the enzyme E is given by:

T_E.cyt = y[E.cyt] + y[E:A:B.cyt] + y[E:C.cyt]

4. Selection of conversion methods

Selection of conversion methods is carried out for the two systems: gene-protein layer and metabolic layer.

4.1 Gene-protein layer

For the gene-protein layers, users select four types of conversion methods: CMA, TPP_STEADYSTATE_1, TPP_STEADYSTATE_2, and TPP_RAPID, which are suitable for converting the regulator-reaction equations that generate various complexes.

The gene-protein layer can be converted into GMA or MM, but it is not practically useful. Thus, such conversions are error.

4.2 Metabolic layer

For the metabolic layer, GMA and MM are generally employed. It is possible to convert the metabolic layer by using CMA or TPP, but it is not useful practically.

4.3 Connection between the layers

Since the different conversion methods are applied to the gene-protein and metabolic layers, the differential equations for the species common to both layers have to be combined to make one differential equation, which enables us to simulate the three layers together.