Learn Pathway Simulation in 10 pages

ver. Mar-1-2016

Kazuhisa Ichikawa

True Cell Simulations Inc.

http://engl.tc-simulations.com/

True Cell Simulations Inc. Tokyo Japan

Contents

Introduction

This eBooklet is written for those who are interested in, or are planning to use, pathway simulation. The intended audience includes students and professional researchers that have expertise in experimental cell biology, pharmacology, or medical science. I have tried to make this eBooklet as concise and compact as possible so that readers may understand basics of pathway simulation after reading the 10-pages eBooklet.

Combined use of experimental and *in silico* simulation has been steadily increasing over the past decade. It is presumed that several thousands of such papers were published in 2015. Although this is small fraction of all molecular cell biology papers in the literature, rapid growth of these collaborative studies is expected to continue. The history of simulation research in molecular cell biology can be traced back over 50 years[1]. Physicists and mathematicians interested in biological processes began running simulation as early as 1965. In recent years, experimental molecular cell biologists have begun using simulations to inform their work as well.

Getting started with pathway simulation may be difficult for some students and experimental researchers. This is no reflection on their abilities as scientists, but that methods and techniques used in simulation are quite different from those used in experimental research. This eBooklet is not intended to fill this gap. Instead, it is intended to introduce pathway simulation beginning with pathway diagrams commonly used both experimental and simulation work[2].

Introducing pathway simulation without mathematics is impossible. Although every attempt was made to present the relevant mathematics as briefly and simply as possible, those who are not familiar with simulation should carefully read through the equations presented.

Pathway simulation is the basis for, but is not equal to cell simulation. Cell simulation includes all mechanisms and phenomena in a cell including protein-protein interactions, diffusions, membrane potentials, mechanical events, structural changes of organelles, morphological change of a cell, etc. First, one must study pathway simulations, before one can run cell simulation. Then, one can find a path to cell simulation.

---

1 Goodwin B.C., Adv.Enzyme Reg., 1965, 425.

2 I wrote pathway diagram as a starting point. But it might be a goal of the research for experimental researchers. Phenomenon is a starting point, and pathway diagram is written as a result of their studies by combining roles of proteins under investigation. In the simulation, on the contrary, phenomenon is reproduced by a computer starting at the pathway diagram.

Chapter 1: Pathway Simulation

A pathway is a set of transmission cascades of activation states, which is more properly described simply by 'states', of proteins and other biological molecules. Since states and their regulation are mechanisms behind a phenomenon, a pathway is an expression of mechanisms behind a phenomenon[3]. Since a pathway is an intelligible expression of a mechanism, it is often shown as a pictorial summary of experimental research, which is quite helpful for readers (Figure 1 for an example[4]).

Figure 1 shows a pathway for a membrane protein MT1-MMP, which plays an important role in metastatic invasion of cancer cells. MT1-MMP is an enzyme that degrades extracellular proteins such as the extracellular matrix (ECM) and basement membranes. Removal of the ECM by MT1-MMP creates an opening in the extracellular space into which cells can migrate during the initial steps of invasion.

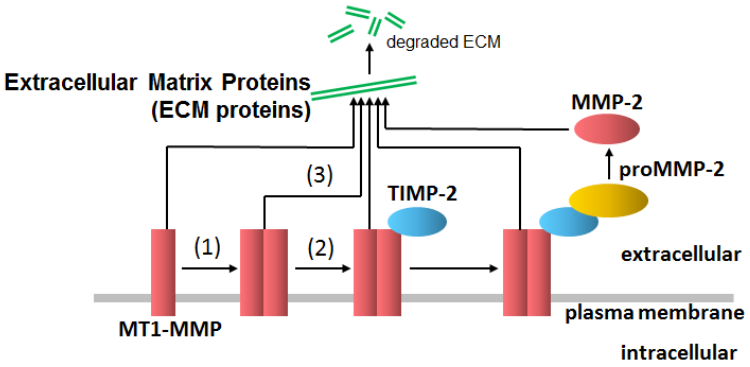

Figure 1 Pathway describing ECM degradation by MT1-MMP.

MT1-MMP forms a homodimer that can be inhibited by an intrinsic extracellular protein called TIMP-2. Inhibition of MT1-MMP by TIMP-2 prevents degradation of the ECM[5]. In addition, TIMP-2 binds a precursor of MMP-2 (ProMMP-2), another ECM-degrading protein, which is activated by MT1-MMP in the complex generating active MMP-2. Figure 1 illustrates ECM degradation by MT1-MMP with arrows indicating either change in the states of proteins or their action. For example, arrow (1) shows the formation of homodimer of MT1-MMP, and arrow (2) shows the binding of TIMP-2 to MT1-MMP homodimer forming MT1-MMP:MT1-MMP:TIMP-2. Arrow (3) is different from arrows (1) and (2), showing degradation of the ECM by MT1-MMP homodimer acting on it. Thus arrow (3) shows action of a protein instead of its sate change.

Pathways regulating MT1-MMP and its function in degrading ECM are well described in Figure 1, which is a summary of many experiments. Hence, identifying the pathway was a goal of the research. Conversely, the pathway is the starting point for building simulations. Why is pathway a starting point for simulation? Computers use equations to simulate a process, and the pathway shown in Figure 1, for instance, can be converted into a set of equations.

---

3 We can say also that a pathway shows functions of proteins.

4 Seiki, M., Cancer Lett., 2003, 194; Hoshino, D., et al., PLoS Comp.Biol., 2012, 8.

5 In the example shown in Figure 1, MT1-MMP possesses half of activity to degrade ECM, because one MT1-MMP in a homodimer is bound with TIMP-2, and another MT1-MMP is free to degrade ECM.

The next step is to show is how we can convert a pathway into a set of equations. An arrow (1) in Figure 1 shows us a chemical reaction, in which two monomeric MT1-MMPs bind together forming a homodimer of MT1-MMP. This is written in the form of chemical reaction as follows:

$$\text{MT1-MMP} + \text{MT1-MMP} \rightleftharpoons \text{MT1-MMP:MT1-MMP} \qquad 1)$$

The left side of 1) shows us the binding of two MT1-MMPs, and right side shows us the formation of homodimer MT1-MMP:MT1-MMP as a result of the reaction. Here we use ":" to indicate a complex formed by the binding of the two monomers. There are other symbols that can be used to indicate interactions between molecules forming a complex, such as ".". A rigid nomenclature rile does not exist. The reaction shown in 1) is reversible as the complex can be formed or dissociate.

Thus, the pathway in Figure 1 can be converted into a set of chemical reactions. The chemical reactions can be converted into a set of equations, which in this case are differential equations. This conversion is shown in detail in Chapter 3. Thus, if chemical reactions are derived from given pathway, we can run a simulation for the pathway. Hence, the pathway is the starting point for a simulation. The conversions are summarized in Figure 2, showing the steps from pathway to a simulation. This conversion is the essence for running simulation. The following chapters will discuss how these conversions are made.

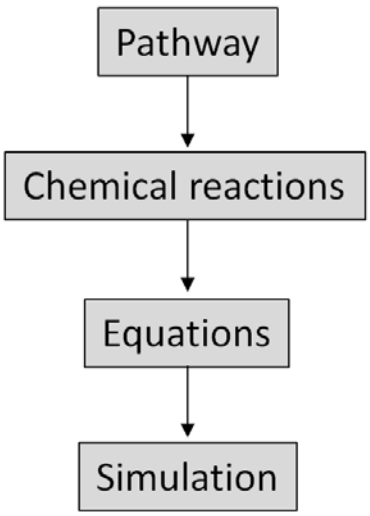

Figure 2 Pathway is converted to a set of chemical reactions, and chemical reactions are converted to a set of differential equations. Equations are calculated by a computer. Thus, we can run a simulation beginning with a pathway.

It should be noted that not all simulation studies follows the conversion shown in Figure 2. It is true that a pathway is the starting point in almost all simulation studies. However, a pathway is not converted to chemical reactions in some simulation studies, or converted chemical reactions are not converted to differential equations. Instead they are described by a set of rules, which is seen in particle movement simulations run by a stochastic simulator. These methods of simulation are beyond the scope of this eBooklet, and will be described elsewhere.

Chapter 2: From Pathway to Chemical Reactions

A pathway is a simplified expression of protein-protein interactions, which are then translated into a set of chemical reactions. Although there are many types of chemical reactions, there are only a few basic chemical reactions used, each discriminated by the number of reactants. Basic chemical reactions are illustrated here. All other pathway chemical reactions are combination of these.

The first basic chemical reaction is the first order reaction, in which a single reactant is converted autonomously into other species.

$$A \xrightarrow{k_1} B \qquad 2)$$

In this example, molecule A is converted into B with the conversion rate defined by the rate constant k_1. A dissociation reaction, in which a single reactant dissociates into two molecules, is also a first order reaction.

$$A \xrightarrow{k_1} B + C$$

If A is a homodimer, it dissociates into two Bs.

The next equation is the second order reaction, in which there are two reactants.

$$X + Y \xrightarrow{k_2} Z \qquad 3)$$

In this example, molecules X and Y bind together producing Z. The second order reaction frequently appears in pathways. A product can be described X:Y, X.Y, or any nomenclature instead of Z. If reactants are two Xs, the product is X:X, a homodimer of X. Reaction 1) in Chapter 1 is a second order reaction. A ligand binding to a receptor is also a second order reaction.

The first and second order reactions are enough to describe all types of reactions. However, it is better to describe an equilibrium reaction as a basic reaction.

$$X + Y \underset{k_b}{\overset{k_f}{\rightleftharpoons}} Z \qquad 4)$$

This is the combined reaction of the binding of X and Y producing Z and its dissociation into X and Y. Reaction 4) can also be described as follows:

$$X + Y \xrightarrow{k_f} Z \xrightarrow{k_b} X + Y$$

The first and second order reactions are drawn consecutively. Obviously, the expression shown in reaction 4) is more appropriate and will be used in this eBooklet.

There are only three basic reactions, reactions 2), 3), and 4). All pathways are combination of these. For example, a Michaelis-Menten enzymatic reaction is a combination of 4) and 2). Although reactions 2), 3), and 4) are well known, the combination of these can result in unexpected phenomena.

Here, the conversion of the pathway shown in Figure 1 into chemical reactions is shown. Conversion of path (1) to a chemical reaction was already described in reaction 1). Path (2) is also converted into a second order reaction as follows:

$$\text{MT1-MMP:MT1-MMP} + \text{TIMP-2} \rightleftharpoons \text{MT1-MMP:MT1-MMP:TIMP-2} \qquad 5)$$

Next, path (3) is converted into chemical reactions. However, this conversion is different from those in paths (1) and (2), because path (3) is not a direct expression of a chemical reaction. It shows a function of the MT1-MMP dimer in the degradation of the ECM. Therefore, path (3) should be converted into a catalytic reaction, in which the MT1-MMP dimer and the ECM are an enzyme and a substrate, respectively. Hence, path (3) is converted into chemical reactions as follows:

$$\begin{aligned} \text{MT1-MMP:MT1-MMP} + \text{ECM} &\rightleftharpoons \text{MT1-MMP:MT1-MMP:ECM} \\ &\longrightarrow \text{MT1-MMP:MT1-MMP} + \text{ECM}_d \end{aligned} \qquad 6)$$

ECM_d indicates degraded ECM. As shown in this example, paths that are not directly converted into reactions are often included in a pathway diagram, because researchers often want to show the relationship and/or functions of proteins rather than simple association between them. Thus, in this case, there is no direct relation between proteins in terms of chemical reactions. When one tries to convert pathways into chemical reactions, these scenarios must be considered. Knowledge of molecular cell biology and the relevant body of research are a great resource for these conversions[6].

[6] In addition, you may find a hint in the path. Careful investigation of the relation between the upstream and downstream proteins can lead to constructing a set of proper chemical reactions. In this case, an arrow shown in the pathway might be converted into many reactions with many arrows. Even diffusion, transportation on microtubules, and membrane potential, which are not converted into chemical reactions, are shown as a part of pathway diagram in some cases. When one realize these, modeling method other than chemical reactions should be employed. This will be described in separate eBooklet.

Chapter 3: From Chemical Reactions to Differential Equations

Differential equation can be difficult to understand and manipulate, but the concepts needed for understanding pathway simulation are not difficult. In this eBooklet, differential equations are used in reference to time derivatives for concentrations of molecular species. For instance, a differential equation for the concentration of molecule A is written as follows:

$$\frac{dA}{dt} = \cdots$$

The right side, shown by "$\cdots$", will be replaced by an equation representing the reactions for A. The left side on the equation is dA/dt, which indicates the rate of the change in A, that is, how much the concentration of A changes, over a period of time, for instance after 1 s. When dA/dt is large, the concentration of A changes rapidly. Conversely, the concentration of A changes slowly when it is small, and when the concentration reaches 0 there is no further change. When the solution of the equation is positive or negative, concentration increases or decreases, respectively. Thus dA/dt shows us the rate of change and its direction (increase of decrease). This is the base concept of differential equations.

The differential equation 7) below is an example.

$$\frac{dA}{dt} = k \cdot A \qquad 7)$$

Since A is concentration, $A \geq 0$. If k>0, the right side will be 0 or positive, and never yields a negative value. This suggests that A increases in many cases. Considering that the concentration of A appears on the both sides of the equation, the rate of the change in A (left side) is proportional to its own concentration (right side). Thus, if the concentration of A is large, the rate of the increase in A is also large and increases explosively. Thus, we can calculate the change in the concentration for a molecule shown in the left side using equations shown in the right side. Since the calculated change is for the present time, the concentration at the next time interval can be estimated by an appropriate approximation method. This is what a simulation program will do.

We have finished the preparation for converting the basic reaction 2) into differential equations as follows:

$$\frac{dA}{dt} = -k_1 \cdot A$$
$$\frac{dB}{dt} = k_1 \cdot A \qquad 8)$$

Since there are two species A and B in 2), there are two differential equations for A and B in equation 8). If we convert reaction 3), we get three differential equations as follows:

$$\frac{dX}{dt} = -k_2 \cdot X \cdot Y$$
$$\frac{dY}{dt} = -k_2 \cdot X \cdot Y \qquad 9)$$
$$\frac{dZ}{dt} = k_2 \cdot X \cdot Y$$

Thus, chemical reactions can be converted into differential equations. Conversely, differential equation can be converted into chemical reaction, which is not shown in this eBooklet.

In the last part of this chapter, the following Michaelis-Menten enzymatic reaction will be converted into differential equations.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E + P \qquad 10)$$

The resulting differential equations are shown below.

For E,

$$\frac{dE}{dt} = -k_1 \cdot E \cdot S + k_{-1} \cdot ES + k_2 \cdot ES \qquad 11),$$

For S,

$$\frac{dS}{dt} = -k_1 \cdot E \cdot S + k_{-1} \cdot ES \qquad 12),$$

For ES,

$$\frac{dES}{dt} = k_1 \cdot E \cdot S - k_{-1} \cdot ES - k_2 \cdot ES \qquad 13),$$

For P,

$$\frac{dP}{dt} = k_2 \cdot ES \qquad 14).$$

All terms on the right side are combined together with respect to the molecules on the left side. For example, the first, the second, and the third terms in right side of differential equation 11) describe the rate of changes in E by the binding of S, by the backward dissociation from ES complex, and by the forward dissociation from an ES complex generating E and P. Since all of these reactions change the concentration of E, three terms are combined together in differential equation 11). If the reaction increase or decrease E, the sign is positive and negative, respectively.

Thus chemical reaction can be converted into differential equations as shown above. Differential equations are mathematical representations of pathway, for which we can run simulation as shown in Chapter 1. Thus, conversion of a pathway to chemical reactions enables us to simulate it.

Chapter 4: From Differential Equations to Simulation

The purpose of pathway simulation is to determine the time-dependent change, i.e. the time course, of a protein concentration. If the activity of MT1-MMP, an ECM-degrading protein shown in Chapter 1, is persistent, we can conclude that MT1-MMP will be a threat promoting invasion and metastasis. If the activity is temporal, it might not be a threat. Thus it is important to know the time-dependent change in the protein concentration.

How can we know the time-dependent change in the concentration of a protein? As discussed in Chapter 3, differential equations derived from chemical reactions describe the rate of the time-dependent change in the concentrations. By using the rate of the change A_0' at the present t_0, we can obtain the concentration of A at time $t_0+\Delta t$, which is a future time relative to t_0. This is the simplest method to obtain the concentration at $t_0+\Delta t$ ($A_{0+\Delta t}$ in Figure 3). In this method, a small change in the concentration is calculated by an equation $\Delta t \cdot A_0$' assuming that the rate of the change A_0' does not change during Δt. This method is exactly the same as that used when estimating how far a car will travel when moving at a constant speed for a given duration of time. If we want to know how far we can get in 30 min at a rate of 80 km/h, we calculate by using an equation 80 km/h × 0.5 h = 40 km. If we perform this estimation repetitively by using the car speed at the time of the estimation, we will be able to estimate traveling distance over a long duration of time. The same method is used in pathway simulation.

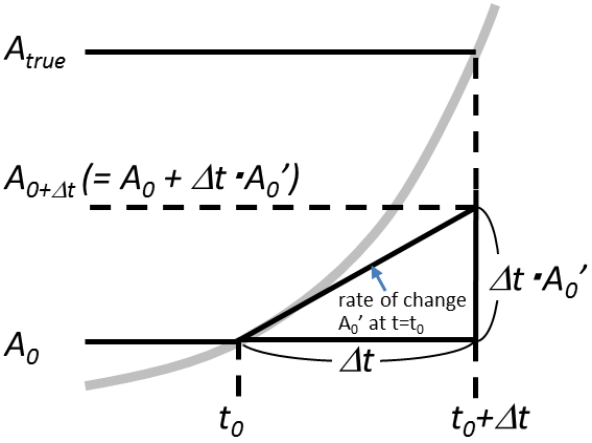

Figure 3 A method approximating the value of A at $t_0+\Delta t$ starting from its value at t_0.

However, the difference between estimated and true concentrations, $A_{0+\Delta t}$ and A_{true}, represents an unavoidable error inherent to simulations. There are many methods to reduce error. One of the simplest methods is to reduce the time step Δt. For example, using the first order reaction shown in 2) in the form of the associated differential equation 8), A can be mathematically integrated as follows:

$$A = A_0 \cdot \exp(-k_1 \cdot t) \qquad 15)$$

If you have a scientific calculator, the value of A at selected time t is easily calculated. Results are shown as real time data (analytic solutions) in Table 1 using initial conditions with $A = 1\mu M$ at $t = 0$ and $k_1 = 5\ /s$. Two simulation methods are shown for comparing error.

Table 1 Comparison of errors between analytic solution and two simulation methods

t (s)	analytic solution	Euler Method Δt=0.1 s			Euler Method Δt=0.05 s			Runge-Kutta Method (A-Cell) Δt=0.1 s	
	A (μM)	dA/dt (μM/s)	A (μM)	error (%)	dA/dt (μM/s)	A (μM)	error (%)	A (μM)	error (%)
0	**1.000E+00**	-5.000E+00	**1.000E+00**	**0**	-5.000E+00	**1.000E+00**	**0**	**1.000E+00**	**0.000E+00**
0.05	7.788E-01				-3.750E+00	7.500E-01			
0.1	**6.065E-01**	-2.500E+00	**5.000E-01**	**-18**	-2.813E+00	**5.625E-01**	**-7**	**6.068E-01**	**4.441E-02**
0.15	4.724E-01				-2.109E+00	4.219E-01			
0.2	**3.679E-01**	-1.250E+00	**2.500E-01**	**-32**	-1.582E+00	**3.164E-01**	**-14**	**3.682E-01**	**8.714E-02**
0.25	2.865E-01				-1.187E+00	2.373E-01			
0.3	**2.231E-01**	-6.250E-01	**1.250E-01**	**-44**	-8.899E-01	**1.780E-01**	**-20**	**2.234E-01**	**1.209E-01**
0.35	1.738E-01				-6.674E-01	1.335E-01			
0.4	**1.353E-01**	-3.125E-01	**6.250E-02**	**-54**	-5.006E-01	**1.001E-01**	**-26**	**1.355E-01**	**1.217E-01**
0.45	1.054E-01				-3.754E-01	7.508E-02			
0.5	**8.208E-02**	-1.563E-01	**3.125E-02**	**-62**	-2.816E-01	**5.631E-02**	**-31**	**8.225E-02**	**2.010E-01**
0.55	6.393E-02				-2.112E-01	4.224E-02			
0.6	**4.979E-02**	-7.813E-02	**1.563E-02**	**-69**	-1.584E-01	**3.168E-02**	**-36**	**4.991E-02**	**2.469E-01**
0.65	3.877E-02				-1.188E-01	2.376E-02			
0.7	**3.020E-02**	-3.906E-02	**7.813E-03**	**-74**	-8.909E-02	**1.782E-02**	**-41**	**3.028E-02**	**2.736E-01**
0.75	2.352E-02				-6.682E-02	1.336E-02			
0.8	**1.832E-02**	-1.953E-02	**3.906E-03**	**-79**	-5.011E-02	**1.002E-02**	**-45**	**1.837E-02**	**3.132E-01**
0.85	1.426E-02				-3.758E-02	7.517E-03			
0.9	**1.111E-02**	-9.766E-03	**1.953E-03**	**-82**	-2.819E-02	**5.638E-03**	**-49**	**1.115E-02**	**3.691E-01**
0.95	8.652E-03				-2.114E-02	4.228E-03			
1	**6.738E-03**	-4.883E-03	**9.766E-04**	**-86**	-1.586E-02	**3.171E-03**	**-53**	**6.765E-03**	**4.015E-01**

The column *dA/dt* in the 'Euler method' shows values calculated by equation 8). Concentrations of *A* at subsequent times were calculated using the following equation.

$$A_{t+\Delta t} = A_t + (dA/dt)_t \cdot \Delta t = A_t - k_1 \cdot A_t \cdot \Delta t \qquad 16)$$

There are two sets of simulation by Δt of 0.1 s and 0.05 s in the 'Euler method'. By comparing values for these two conditions at $t=0.5$ s, smaller Δt ($\Delta t=0.05$ s) gave a better approximation to analytical value with a smaller error (-31%). However, "Runge-Kutta method"[7] gave a much better approximation than the 'Euler method', even at a larger time interval Δt of 0.1 s (0.201%). Although the error is a problem in simulations, it can be reduced reasonably in pathway simulations. An important conclusion in this Chapter is that simulations are performed if differential equations can be derived.

In Chapters 1 to 4, we have connected pathways to simulations as follows: pathway → chemical reactions → differential equations → simulation. You can run a simulation if you go through these conversions starting with a pathway. Since simulations are run using equations, the exact definitions of mechanisms and parameters are required, which might be feasible in experimental biology.

7 Runge-Kutta method is frequently used in simulations because it gives better approximation. A-Cell employs Runge-Kutta method.

Conclusions

Hopefully, you now have a basic understanding of how simulations can be run by starting with pathway. One can practice converting a pathway in a paper to chemical reactions, which will bring you nearly to the point of running a pathway simulation.

However, there are still some important details required that have not yet been discussed. For instance, obtaining a simulation program that can be run on your computer. In addition, the conversion of a pathway to differential equations can be time consuming and fraught with opportunities for making mistakes. As discussed in Chapter 4, the conversion from chemical reactions to differential equations follows definite rules, and can be done by a computer. The program A-Cell was developed for the purpose of this conversion for the first time as a result of personal experience in making mistakes performing conversions. A-Cell performs conversions from chemical reactions → differential equations → simulation automatically. Thus, A-Cell generates a simulation program for a specific pathway simulation. Therefore, you can focus your mental resources on the conversion of a pathway to chemical reactions with analyses of the simulation results generated by A-Cell.

It should be noted that the conversion of a pathway → chemical reactions and subsequent analyses of simulation results do not follow definite rules. Thus, these steps are not replaced by a computer, but can be the most interesting in simulation studies. You can predict a novel mechanism or phenomenon starting with an incomplete pathway, adding your own theoretical mechanisms into your simulation studies. Alternatively, you can determine an essential mechanism by analyzing complex simulation results that are not easily understood. It may then be possible to confirm your prediction through experimentations. It is always better to use a computer by rule-based conversions, and to focus your intellectual resources on creative activities.

I focused on pathway simulation in this eBooklet. You know that protein-protein interactions described in a pathway are not the only mechanisms within cells. There are many other mechanisms such as diffusion, transportation, mechanical and osmotic change, membrane potential, etc. However, starting with pathway simulation is the best way to begin simulation studies in molecular cell biology as this foundation is necessary for further complex studies.

Appendix: Chemical Reactions for MT1-MMP Pathway

An example of a set of chemical reactions for the pathway shown in Figure 1 is shown below. For simplicity, M14, T2, pM2, and M2 are used instead of MT1-MMP, TIMP-2, proMMP-2, and MMP-2, respectively. ECM degradation by MMP-2 is neglected.

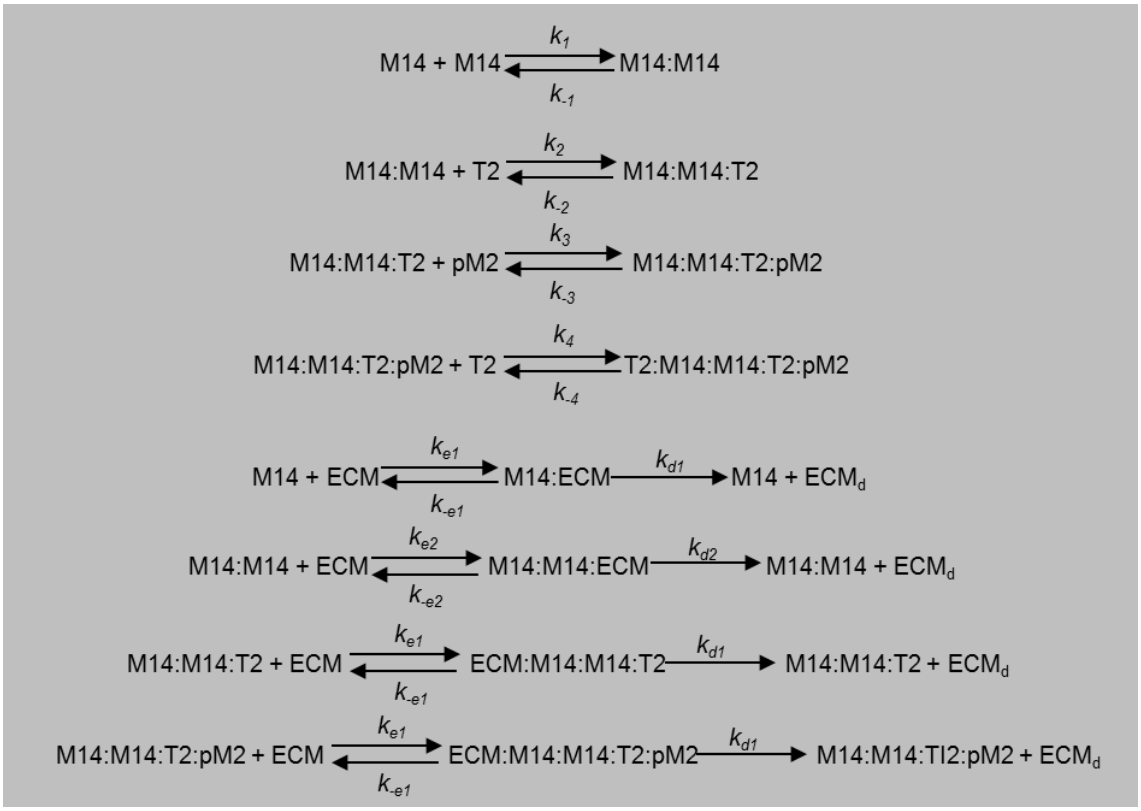

Figure 4 Chemical reactions for the regulation of MT1-MMP and ECM degradation for the pathway shown in Figure 1.

If you have practiced drawing your own chemical reactions for the pathway, compare these with yours. Drawing chemical reactions starting with a pathway is a valuable skill in running simulations. Figure 4 is shown for illustrative purposes only as there are many ways to draw chemical reactions. It is important to remember that any pathway can be converted to chemical retains as long as there is extensive knowledge of molecular cell biology.

There is a problem in Figure 4 in that the relationships between reactions are not obvious. This is a result of writing reactions in parallel. In contrast, the same reactions can be drawn as shown in Figure 5 using A-Cell. This clearly shows relationships and the structure of reactions. In fact, reactions shown in Figure 5 are drawn nearly identical to the pathway shown in Figure1.

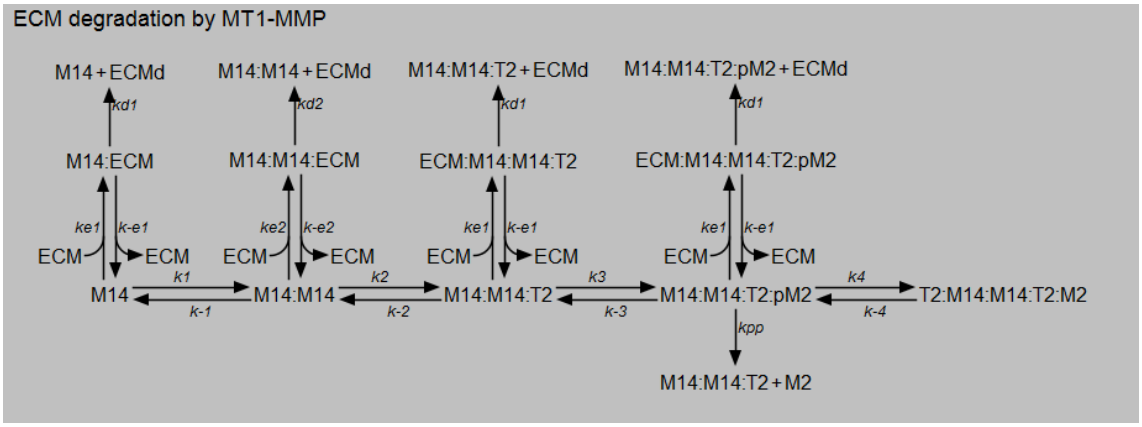

Figure 5 Structured chemical reactions for Figure 1

A-Cell does allow you to draw chemical reactions as shown in Figure 4, but Figure 5 makes it much easier to grasp the overall relationships of each reaction. When analyzing simulation results, it is much easier to find a cause of a result if chemical reactions were drawn by structured format as shown in Figure 5. It should be noted that A-Cell gives the same differential equations for Figures 4 and 5; hence, we get exactly the same simulation results.

Chemical reactions shown in Figures 4 and 5 are drawn by the A-Cell Free edition. When you try to simulate these reactions using the Free edition, refer http://engl.tc-simulations.com/models/ for concentrations and rate constants. Although the model shown in this Web site is scientifically correct, a simpler model is shown in this eBooklet for the purpose of explaining the importance of the structured chemical reactions. Take time to compare the difference in the simulation results between the correct chemical reactions shown in the Web site and those shown in Figures 4 or 5.