Abstract

Carminic acid is the product after a long metabolic process from CoA. Thus, because mathematical modeling can maximize the end product yield by a quantified design, it is essential for us to quantify the metabolic pathway design. This model is aimed to analyze the biosystem and illustrate the synthesis of enzyme and the rate of enzyme catalyzed reaction by derivative functions. Since the rate of enzyme catalyzed reaction is determined by the rate of synthesis of enzymes, we maximize our yield by controlling the rate of synthesis of enzymes. In addition, as biokinetic system depends on the organism itself, the analysis of the growth of microorganism allows us to deepen the analysis of the whole biokinetic system and make the model more realistic. This will facilitate our teammates to design the pathway.

Construction of Model

In our biosystem, only the rate of expression of enzymes ia able to control the yield of the end product. However, the rate of growth of the enzyme is irrelevant to the control of the end product yield, because there is no factor that will boost or delay the growth rate of the S. cerevisiae in our biosystem. Thus, the yield of the end product (carminic acid) will not be affected. Being a eukaryotic organism, S. cerevisiae has non-linear gene expression, and the expression of any gene is random. Nevertheless, as the total number of gene is constant, the possibility a gene will be expressed is direct proportional to the copy number of the gene (number of promotor in the gene). Since this model considered the rate of synthesis of certain enzyme after certain gene is expressed, with the increasing length of the gene, the rate of expression of the enzyme decreases. Set the number of S. cerevisiae is ‘K’, we can exploit the function below to illustrate the relative synthesis rate of five enzymes: OKS, ZhuI, ZhuJ, monooxygenase, and DCUGT. ‘N’ represents copy number, and ‘L’ stands for gene length.

$$\frac{d[OKS]}{dt} = \frac{N_{oks}}{L_{oks}}W$$
$$\frac{d[ZhuI]}{dt} = \frac{N_{ZhuI}}{L_{ZhuI}}W$$
$$\frac{d[ZhuJ]}{dt} = \frac{N_{ZhuJ}}{L_{ZhuJ}}W$$
$$\frac{d[mono]}{dt} = \frac{N_{mono}}{L_{mono}}W$$
$$\frac{d[DCUGT]}{dt} = \frac{N_{DCUGT}}{L_{DCUGT}}W$$

As OKS, ZhuI, ZhuJ, monooxygenase, and DCUGT are enzymes, the reactions in the S. cerevisiae are standard enzyme catalyzed reaction. Therefore, we can use Michaelis-Menten equation to illustrate the rate of synthesis of the intermediate products in the pathway:

$$\frac{d[Octaketide]}{dt} = \frac{K_{cat(oks)}[COA][OKS]}{K_{m(oks)} + [COA]}$$

$$\frac{d[FA]}{dt} = \frac{K_{cat(ZhuI)}[Octaketide][ZhuI]}{K_{m(ZhuI)} + [Octaketide]}$$

$$\frac{d[FK]}{dt} = \frac{K_{cat(ZhuJ)}[FA][ZhuJ]}{K_{m(ZhuJ)} + [FA]}$$

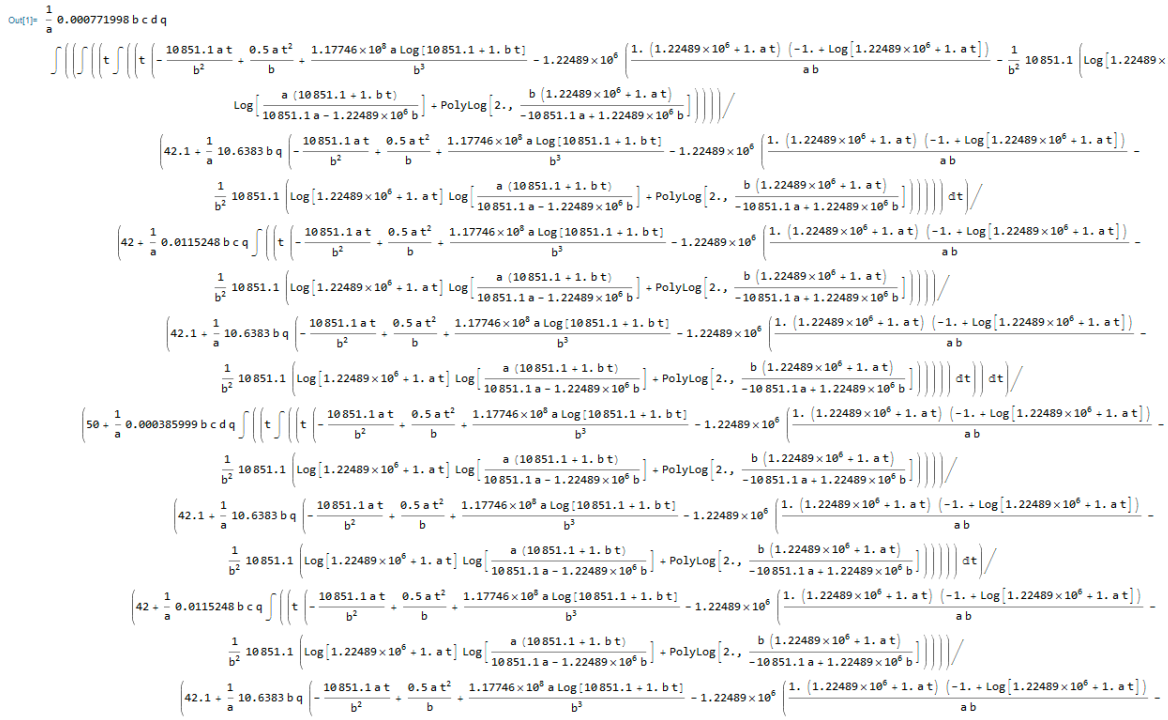
$$\frac{d[KA]}{dt} = \frac{K_{cat(mono)}[FK][mono]}{K_{m(mono)} + [FK]}$$

$$\frac{d[CA]}{dt} = \frac{V_{\max(DCUGT)}][DCUGT]}{K_{m(KA)} + [KA]}$$

Table shows the parameter of the equations

	N	L	K_{cat}	K_m	V_{max}
OKS	a	1212	0.094	95	-
ZhuI	b	510	1	2	-
ZhuJ	c	771	1.3	42.1	-
Monooxygenase	d	1254	42	42	-
DcUGT	E	1548	-	50	2

Solve the function above and input the parameters of the equation, we get the expression of [CA]:



Wolfram Mathematica could not directly express the equation of [CA], it is exceptionally hard to even stimulate system of the

Model Analysis

Advantage: the model is clear as it shows the essence of the mechanism.
Disadvantages: the model only shows the copy number of each enzyme when the yield of carminic acid reaches the plateau, whereas the consideration of the biokinetic system as a whole is missing; it is unable to show the yield of carminic acid when the optimum gene copy number is reached.

Improvements

Considered that S. cerevisiae is eukaryotic organism, its chromosomes will condense when it is replicating, and the gene will not be expressed. Thus, this factor should be considered when constructing the model. The natural growth rate of S. cerevisiae can be illustrated by the function below:

$\frac{dY(t)}{dt}$=K*Y(t)*(1-Y(t))

Solve the equation:

Y(t)=$\frac{e^{K*t}}{e^{K*t}+C[1]}$

As every S. cerevisiae has nearly the same length of time from the start of replication to the end of replication, we can take 'T' as the time for the start of mitosis to the end of cytokinesis. Set at t_0 there are 'k' numbers of S. cerevisiae stop expressing and begin mitosis. Therefore, after 'T', there are 'k' more S. cerevisiae. This means that at t_0, the rate of replicating S. cerevisiae equals to the rate of increase of S. cerevisiae at time t_{0+T}. By this approach, we can exert this idea to illustrate the relationship between replicating S. cerevisiae and time. Use W(t) to express the rate of S. cerevisiae that initiates replication, condenses the chromosome, and stops expressing at time 't':

W(t)'=Y(t+T)'=$-\frac{K*e^{2*K*(t+T)}}{(e^{K*(t+T)}+C[1])^2}+\frac{K*e^{K*(t+T)}}{e^{K*(t+T)}+C[1]}$

Solve the indefinite integral, we can get W(t):

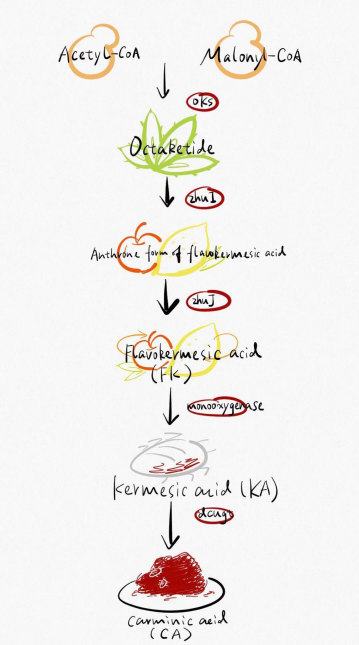
W(t)=$\frac{C[1]}{e^{K(t+T)}+C[1]}$

Thus, the number of S. cerevisiae that are able to express gene can be represent by Y(t)-W(t):

Y(t)-W(t)=$\frac{e^{K*t}}{e^{K*t}+C[1]}-\frac{C[1]}{e^{K(t+T)}+C[1]}$

Being a eukaryotic organism, S. cerevisiae has non-linear gene expression, and the expression of any gene is random. Nevertheless, as the total number of gene is constant, the possibility a gene will be expressed is direct proportional to the copy number of the gene (number of promotor in the gene). Since this model considered the rate of synthesis of certain enzyme after certain gene is expressed, with the increasing length of the gene, the rate of expression of the enzyme decreases (rate if gene expression keeps constant). S. cerevisiae needs 5 kinds of enzyme within the reaction: OKS, ZhuI, ZhuJ, monooxygenase, and DCUGT. The functions of the rate of expression of each enzyme is listed below:

$$\frac{d[OKS]}{dt} = \frac{N_{oks}}{L_{oks}}\left(\frac{e^{K*t}}{e^{K*t}+C[1]} - \frac{C[1]}{e^{K(t+T)}+C[1]}\right)$$
$$\frac{d[ZhuI]}{dt} = \frac{N_{ZhuI}}{L_{ZhuI}}\left(\frac{e^{K*t}}{e^{K*t}+C[1]} - \frac{C[1]}{e^{K(t+T)}+C[1]}\right)$$
$$\frac{d[ZhuJ]}{dt} = \frac{N_{ZhuJ}}{L_{ZhuJ}}\left(\frac{e^{K*t}}{e^{K*t}+C[1]} - \frac{C[1]}{e^{K(t+T)}+C[1]}\right)$$
$$\frac{d[mono]}{dt} = \frac{N_{mono}}{L_{mono}}\left(\frac{e^{K*t}}{e^{K*t}+C[1]} - \frac{C[1]}{e^{K(t+T)}+C[1]}\right)$$
$$\frac{d[DCUGT]}{dt} = \frac{N_{DCUGT}}{L_{DCUGT}}\left(\frac{e^{K*t}}{e^{K*t}+C[1]} - \frac{C[1]}{e^{K(t+T)}+C[1]}\right)$$



Solve the equation:

$$\frac{d[Enzyme]}{dt} = \frac{N*K*t + L*K*C[2] + N*log(e^{K*t}+C[1]) + N*log(e^{K*(t+T)}+C[1]}{L*K}$$

As OKS, ZhuI, ZhuJ, monooxygenase, and DCUGT are enzymes, the reactions in the S. cerevisiae are standard enzyme catalyzed reaction. Therefore, we can use Michaelis-Menten equation to illustrate the rate of synthesis of the intermediate products in the pathway:

$$\frac{d[Octaketide]}{dt} = \frac{K_{cat(oks)}[COA][OKS]}{K_{m(oks)}+[COA]}$$

$$\frac{d[FA]}{dt} = \frac{K_{cat(ZhuI)}[Octaketide][ZhuI]}{K_{m(ZhuI)}+[Octaketide]}$$

$$\frac{d[FK]}{dt} = \frac{K_{cat(ZhuJ)}[FA][ZhuJ]}{K_{m(ZhuJ)}+[FA]}$$

$$\frac{d[KA]}{dt} = \frac{K_{cat(mono)}[FK][mono]}{K_{m(mono)}+[FK]}$$

$$\frac{d[CA]}{dt} = \frac{K_{cat(DCUGT)}[KA][DCUGT]}{K_{m(KA)}+[KA]}$$

As K_{cat}, K_m keep constant, solving the derivative function above will get the function between [CA] and time, and those coefficient of gene copy number, N_{oks},N_{ZhuI},N_{ZhuJ},N_{mono},N_{DCUGT}, can control the function between [CA] and time. Therefore, the objective function g(N_{oks},N_{ZhuI},N_{ZhuJ}, N_{mono},N_{DCUGT}) measuring the production rate of CA reaches its maximum point when

$$\frac{\partial[CA]}{\partial t} = 0$$

$$\frac{\partial[CA]}{\partial N_{oks}} = 0$$

$$\frac{\partial[CA]}{\partial N_{ZhuI}} = 0$$

$$\frac{\partial[CA]}{\partial N_{ZhuJ}} = 0$$

$$\frac{\partial[CA]}{\partial N_{mono}} = 0$$

$$\frac{\partial[CA]}{\partial N_{DCUGT}} = 0$$

At this time, by solving the six equations above, we would know that at which sets of gene copy numbers would maximize the production of CA. Furthermore, our model establish a quantitative relationship between time and the mount of production, which enables us to gain understanding of the metabolic pattern.

Analysis of Improvements

Advantages: the model accurately illustrates the dynamic process of the biosystem and considers the growth rate of microorganisms, rate of expressing and synthesizing enzymes, and rate of enzyme catalyzed reactions. The model converts the complicated biosystem into explicate planning problem.

Disadvantages: the complexity of biosystem indicates the complexity of solving the biology model. Indeed, when we tried to solve the rate of formation of [FA], the convoluted function made it impossible for us, a group of high school students, to use ploylog in mathematica to express the rate of formation of [FA], because we have limited resources to simulate this highly complicated system.