

# Metabolic modeling of Rosmarinic acid biosynthetic pathway

Shanthy Sundaram\*, Ashutosh Tripathi, Deepak K Gupta

Centre for Biotechnology, University of Allahabad, Allahabad, U.P, India; Shanthy Sundaram – Email: shanthy\_s@rediffmail.com; \*Corresponding author

Received July 20, 2010; accepted August 20, 2010; published October 06, 2010

## Abstract:

Rosmarinic acid (RA) is an ester of caffeic acid and 3, 4-dihydroxyphenyllactic acid. It is commonly found in *Coleus blumei*, *Salvia officinalis*, *Melissa officinalis* and *Rosmarinus officinalis*. The biosynthesis of RA starts with precursor molecules L-phenylalanine and L-tyrosine. Simulation of RA biosynthetic pathway was done using Gepasi Software, includes the reaction kinetics of each step of the pathway and different integration methods such as Euler's method. Optimization of the significant parameters responsible for RA biosynthesis was carried out. As the goal of the work was to increase the productivity of i.e. to maximize the concentration of the RA, the final concentration of RA ( $[RA]_t$ ) was selected as an objective function and selected initial concentration of the Caffeoyl-3'-4'-hydroxyphenyllactic acid (3'C4HPLA) as parameter constraint and varied its initial concentration as:  $0 \leq [3'C4HPLA]_i \leq 0.025$ . Several optimization methods such as Simulated annealing, Evolutionary algorithms and Genetic algorithms were used to optimize the objective function. After optimization the final concentration of RA was slightly higher (4.566132e-002 mM) than before optimization (4.047119e-002 mM). On the basis of results obtained, it is clear that 4-hydroxyphenyllactic acid and 3'C4HPLA play major role in the high productivity of the RA.

**Keywords:** Metabolic modeling, Rosmarinic acid, Simulation, Simulated annealing.

## Background:

A model is a representation of some observable natural phenomenon. To model a particular phenomenon on the basis of some observations, first a conceptual part is developed, then (but not always) a mathematical part and finally (also not always) an experimental part. The process of making the model work is called simulation. Simulation can be carried out for at least two distinct purposes: one might be interested in using it to test the theory that the model is based upon, or, if the theory seems to be satisfactory, to predict situations that might occur (be observed) in 'reality'. Simulation can be carried out in biological systems also. One can simulate any biological molecule (protein) or any biological process such as any metabolic pathway that may be for synthesis or degradation of particular metabolite.

## Modeling Kinetics

Traditionally, kinetics has been defined in biochemistry in terms of enzyme steady-state kinetics. This corresponds to a detailed study of the local properties of the individual enzymes. However, one can go further and create kinetic models of whole pathways. Such models are composed of coupled ordinary differential (for time courses) or algebraic (for steady states) equations. These equations are non-linear and most often without analytical solution. This means that they can only be studied through numerical algorithms, such as the Newton method for solving non-linear equations and numerical integrators. Kinetic modeling of metabolic pathways may be carried out by computer software that simulates the behavior of a real pathway. Simulation then resembles a true experiment: one sets the initial concentration of the metabolites and the software then produces the time evolution and/or the steady state of these concentrations. There is, however, an additional amount of information to be supplied in the case of simulation: the differential equations describing the kinetics of the pathway and values for all the parameters involved in these equations. [1].

## Biochemical dynamics

Biochemical dynamics addresses questions such as 'How do reactions in closed system reach equilibrium?', 'In open systems, do reactions approach a steady state?', 'In which conditions do metabolite concentrations oscillate?' or 'Can some simultaneous reactions display complex behavior?' The main philosophical difference between biochemical dynamics and traditional enzyme kinetics is then reflected on the latter putting a great emphasis on the determination of mechanisms and estimation of kinetic parameters while the former is primarily concerned with temporal behavior, or trajectories. For this reason, it is frequently possible to ignore mechanistic arguments in biochemical dynamics and use phenomenological descriptions of rates of single enzymes, simplifying the lower level step significantly [2].

## Optimization

Optimization problems are concerned with locating optima (maxima or minima) of the function. Finding a maximum of a function  $f(x)$  is equivalent to finding the minimum of the function  $-f(x)$ . The problem can be stated in the general term as follows:

Given a real valued scalar function  $f(x)$  of  $n$  variables,  $x = (x_1 \dots x_n)$ , problem is to find the optimum of function  $f(x)$  such that  $g_i(x) \geq 0$  with  $i = 1 \dots m$  (inequality constraint) and  $h_j(x) = 0$  with  $j = 1 \dots m$  (equality constraint).

In general, the objective function  $f(x)$  and the constraints  $g_i(x)$  and  $h_j(x)$  are non linear, although frequently the only constraints are linear boundaries of the form  $a_i < x_i < -b_i$  (these actually translate into two separate constraints:  $x_i - a_i > 0$  and  $-b_i - x_i > 0$ ), where  $a_i$  and  $b_i$  are often positive constants [3].

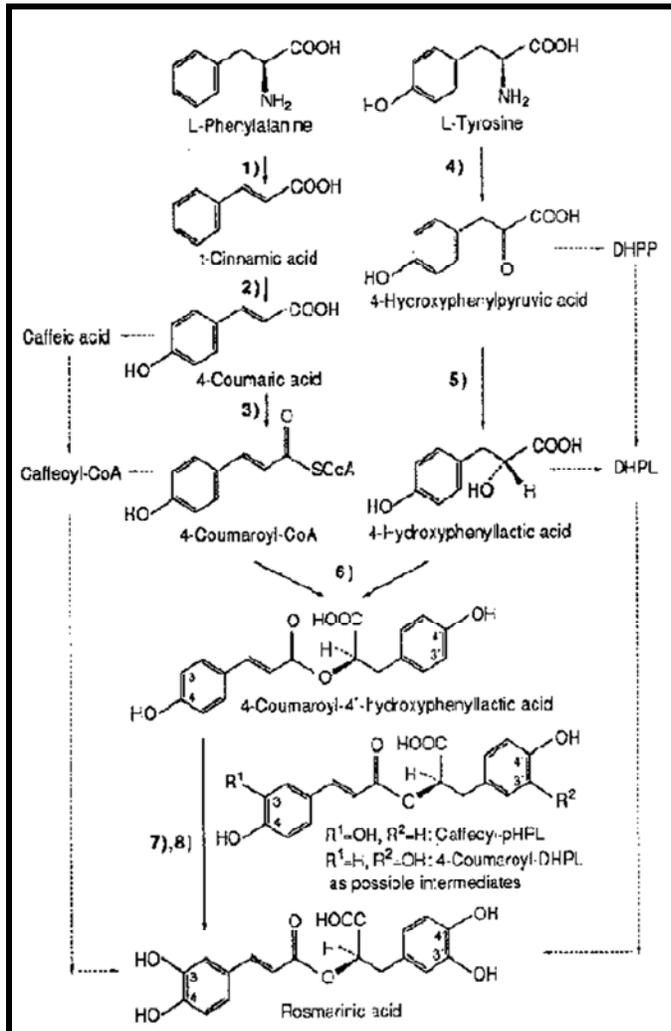


Figure 1: Biosynthetic pathway of Rosmarinic acid (M.Petersan *et al.* 1993) [8]

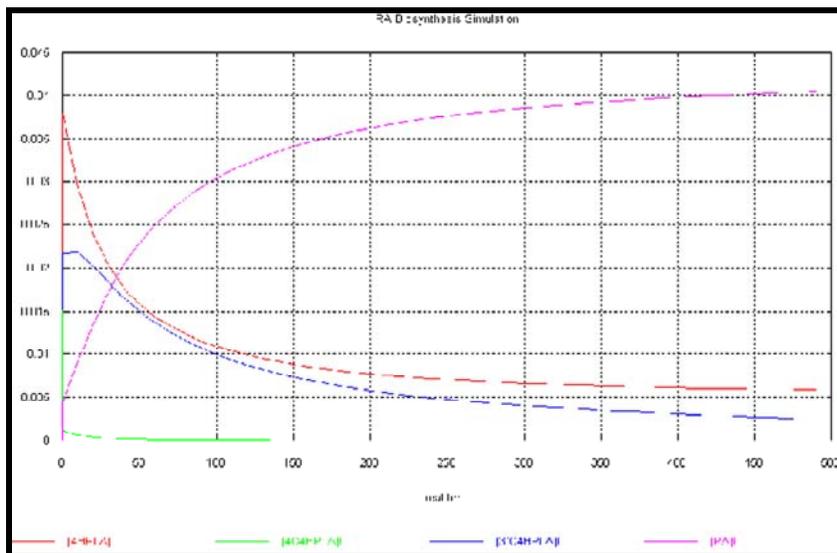


Figure 2: Change in concentration of the metabolites with time

**Methodology:****Pathway of the Rosmarinic acid biosynthesis**

Chemical synthesis of Rosmarinic acid was long sought after and was finally achieved in 1991 by Albrecht [8] (Figure 1). Since then a number of chemical syntheses of Rosmarinic acid and derivatives, e.g. the methylester, different stereoisomers or the less hydroxylated isorinic acid [6] have been described.

**Metabolic simulation software**

**Gepasi 3.3** is a Microsoft Windows program intended for the simulation of the kinetics of systems of chemical and biochemical reactions. Gepasi simulates the steady-state and time-course behaviour of reactions in several compartments of different volumes [2].

**Modeling of the Rosmarinic acid biosynthesis pathway**

To run the simulation of the biosynthetic pathway of Rosmarinic acid on Gepasi, the primary task is to input the pathway into the Gepasi. This process is known as defining the model. (Table 1 see supplementary material). Gepasi takes the reaction in the form of ODEs as mentioned in (Table 2 see supplementary material). In the Metabolites field, the initial concentration of all metabolites in the pathway is defined as in (Table 3 see supplementary material). In this simulation, we were interested in the study of the time course behaviour of the metabolite concentration and the increase of the Rosmarinic acid productivity. The main objective of the simulation was to test the relative change in the concentrations of the metabolites during the course of the simulation. For this purpose, we defined 50 sample points for each iteration and end time was set at 500. Metabolites for the analysis of time course behaviour were set in the Time Course menu. This menu provides the interactive plot facility to examine the relative change in the concentration of the metabolites with time. Optimization parameters such as objective function and parameters-constraints were selected in the Optimization tab. The method of optimization must be specified in the Methods option. The process of the optimization involves following steps: (1) Set initial values for the adjustable parameters. (2) Evaluate the objective function by simulation. (3) Finish if stopping criterion satisfied. (4) Generate new guess for the adjustable parameters. (5) Go back to step 2.

First, the simulation was run without assigning optimization parameters. The need of this step was to evaluate the objective function in the simulation. Once the objective function is determined, the optimization is carried out. Objective function was set into the "Objective function" tab in the optimization page of Gepasi. After this, parameters were included in the optimization simulation which may be responsible for the maximization or minimization of the objective function. We used Genetic Algorithm (Stochastic method) method for optimization at 10 population and 200 iterations.

**Discussion:**

Our goal was to study the dynamics (time course behaviour) of the simulation. During the course of simulation, initial concentration of the metabolites changes with time. Gepasi evaluates these concentrations by integrating the ordinary differential equations (Table 4 see supplementary material). At  $t = 0$ , the initial concentration of Rosmarinic acid was zero, but, as the reaction proceeds, the concentration of the Rosmarinic acid increases with time. Simultaneously the concentration of the other metabolites decreases with time as they had been consumed during the course of the simulation. We could not get the steady state of the simulation. Integration of ODEs states that concentration of only few metabolites shows remarkable change which can be observed by figure 2 and table 5. On the basis of the results, we can conclude that certain parameters such as 4HPLA and 3'-C-4HPLA show remarkable change. So

to maximize the production of the Rosmarinic acid, one should adjust the concentration of the 4HPLA and 3'-C-4HPLA.

In the optimization step, we tried to optimize (maximize) the transient concentration (concentration at time  $t$ ) of Rosmarinic acid  $[RA]_t$  as an objective function. We selected initial concentration of the Caffeoyl-3'-4'-hydroxyphenyllactic acid (3'C4HPLA) as parameter constraint and varied its concentration between 0 and 0.025 such that:  $0 \leq [3'C4HPLA]_i \leq 0.025$

We used Genetic algorithm methodology to optimize the objection function that is  $[RA]_t$ , and carried out 200 iterations of optimization. For results (see supplementary material)

In 200 iterations, 2048 simulations were carried out in 31 sec. at the rate of 66.0323 simulations /sec. The best concentration of Rosmarinic acid at time  $t$  was found as  $[RA]_t = 0.006323031(6.32031e-002)$  when the initial concentration of Caffeoyl-3'-4'-hydroxyphenyllactic acid (3'C4HPLA) was 0.025.

There was a slight increase in the productivity of the Rosmarinic acid after the optimization step (Figure 2).

Before optimization: 4.047119e-002 mM; After optimization: 4.566132e-002 mM; Change in concentration: 0.519013e-002 mM

**Conclusion:**

If one desires to increase the productivity of any biochemical product in the pathway, a long process of biochemical analysis and researches needs to be carried out. This takes much amount of money and human effort and there was no guarantee of positive result. But, the concept of metabolic modeling has solved the problem up to much extent. Computer evaluates the pathway in the mathematical language by using ordinary differential equations and many numerical integration methods. On the basis of simulation of the pathway, approximation can be done about the pathway and productivity can be predicted. Optimization of the metabolic pathway determines the significant parameters, on which one can optimize (maximize or minimize) the yield as per requirement. Here, we used Rosmarinic acid and tried to optimize its biosynthetic pathway to obtain higher yield. On the basis of results obtained, it is clear that 4-hydroxyllactic acid and Caffeoyl-3'-4'-hydroxyphenyllactic acid play major role in the high productivity of the Rosmarinic acid. On the basis of these approximations, wet lab experiments can be carried out and high yield of the product can be obtained.

Thus, *in silico* analysis of metabolic pathways provides an opportunity to find out the step, which affects the end product, consequently making it easy to design our wet lab experiments with respect to time and economic gains.

**Acknowledgement**

We thank to The Coordinator, Centre for Bioinformatics, University of Allahabad for providing us necessary infrastructure facilities.

**References:**

- [1] J Snoep *et al.*, *The Biochemist*. February 25 (1999)
- [2] P Mendes, *Comput Appl Biosci*. 9(5): 563 (1993)
- [3] P Mendes and DB Kell, *BIOINFORMATICS*. 14 (10): 869 (1998)
- [4] M Petersen and MSJ Simmonds, *Phytochemistry*. 62: 121 (2003)
- [5] J William Zinsmeister *et al.*, *Journal of Paleontology*. 63 (6): 731 (1989)
- [6] Michiyo Matsuno *et al.*, *Science*. 325 (5948): 1688 (2009)
- [7] M Petersen *et al.* *Planta*. 189: 10 (1993)
- [8] AR Zinsmeister, *Journal of the National Cancer Institute*. 83 (23): 1734 (1991)

Edited by P. Kanguane

Citation: Sundaram *et al.* Bioinformatics 5(4): 168-172 (2010)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

### Supplementary material:

**Genetic algorithm methodology to optimize the objection function that is [RA]t, carried out 200 iterations of optimization. Results are as follows:**

Optimization:

Maximize [RA]t

Method: Genetic algorithm

Generations = 200  
 Iterations = 200  
 Simulations = 2048  
 Time = 31 s  
 Speed = 66.0323 simulation/s

Best solution:

[RA]t = 0.00632031

[3C4HPLA]i = 0.025

**Table 1: Metabolites and their initial concentrations (in mM)**

Metabolites	Initial Concentration in mM (Approx)
L-Phenylalanine (L-Phe)	1.394000e-002
t- Cinnamic acid (t-CA)	1.250000e-002
4- Coumaric acid (4CouA)	1.385000e-002
4-Coumaroyl- CoA (4CouCoA)	1.538000e-002
L-tyrosine (L-Tyr)	1.519000e-002
4-Hydroxyphenylpyruvic acid (4HPA)	1.527000e-002
4-Hydroxyphenyllactic acid (4HPLA)	1.538000e-002
4-Coumaroyl-4'-hydroxyphenyllactic acid (4C4HPLA)	1.542000e-002
Caffeoyl-4'hydroxyphenyllactic acid (3C4HPLA)	1.526000e-002
Caffeoyl-3'-4'hydroxyphenyllactic acid (3'C4HPLA)	1.526000e-002
Rosmarinic acid (RA)	0.0

**Table 2: Results of integration (after 5.00e+002 s)**

[ L	Phe]	initial = 1.394000e	002 mM, final = 5.714014e	152 mM
[ t	CA]	initial = 1.250000e	002 mM, final = 4.960081e	151 mM
[ 4CouA]		initial = 1.385000e	002 mM, final = 7.150533e	149 mM
[ CAH]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM
[ 4CL]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM
[4CouCoA]		initial = 1.538000e	005 mM, final = 3.342895e	004 mM
[ LTyr]		initial = 1.519000e	002 mM, final = 2.373387e	141 mM
[ TAT]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM
[ 4HPA]		initial = 1.527000e	002 mM, final = 1.595839e	138 mM
[ HPPR]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM
[ 4HPLA]		initial = 1.538000e	002 mM, final = 5.868909e	003 mM
[4C4HPLA]		initial = 1.542000e	002 mM, final = 1.535336e	006 mM
[ RA]		<b>initial = 0.000000e+000 mM, final = 4.047119e</b>	<b>002 mM</b>	
[3C4HPLA]		initial = 1.526000e	002 mM, final = 2.483589e	003 mM
[3'C4HPLA]		initial = 1.526000e	002 mM, final = 2.483589e	003 mM
[ RAS]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM
[ P450]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM
[ PAL]		initial = 1.000000e	002 mM, final = 1.000000e	002 mM

**Table 3: Change in the concentration of the metabolites with time.**

Real time	[4HPLA]t	[3C4HPLA]t	[RA]t	J(R9)
<b>1.60E-0</b>	<b>1.54E-02</b>	<b>1.53E-02</b>	<b>0.00E+00</b>	<b>2.33E-04</b>
3.10E-02	3.80E-02	2.17E-02	4.39E-03	4.69E-04
1.00E+01	2.94E-02	2.17E-02	9.18E-03	4.72E-04
2.00E+01	2.38E-02	2.02E-02	1.36E-02	4.07E-04
3.00E+01	2.02E-02	1.83E-02	1.73E-02	3.36E-04
4.00E+01	1.77E-02	1.66E-02	2.04E-02	2.75E-04
5.00E+01	1.58E-02	1.50E-02	2.29E-02	2.26E-04
6.00E+01	1.44E-02	1.37E-02	2.49E-02	1.88E-04
7.00E+01	1.32E-02	1.26E-02	2.67E-02	1.58E-04
8.00E+01	1.23E-02	1.16E-02	2.81E-02	1.34E-04
9.00E+01	1.16E-02	1.07E-02	2.94E-02	1.15E-04
1.00E+02	1.09E-02	9.98E-03	3.04E-02	9.97E-05
1.10E+02	1.04E-02	9.33E-03	3.14E-02	8.70E-05
1.20E+02	9.91E-03	8.75E-03	3.22E-02	7.65E-05
1.30E+02	9.51E-03	8.23E-03	3.29E-02	6.78E-05
1.40E+02	9.15E-03	7.77E-03	3.35E-02	6.04E-05
1.50E+02	8.84E-03	7.36E-03	3.41E-02	5.41E-05
1.60E+02	8.57E-03	6.98E-03	3.46E-02	4.87E-05
1.70E+02	8.32E-03	6.64E-03	3.51E-02	4.41E-05
1.80E+02	8.10E-03	6.33E-03	3.55E-02	4.01E-05
1.90E+02	7.90E-03	6.05E-03	3.59E-02	3.66E-05
2.00E+02	7.73E-03	5.79E-03	3.62E-02	3.35E-05
2.10E+02	7.56E-03	5.55E-03	3.66E-02	3.08E-05
2.20E+02	7.42E-03	5.33E-03	3.68E-02	2.84E-05
2.30E+02	7.28E-03	5.12E-03	3.71E-02	2.62E-05
2.40E+02	7.16E-03	4.93E-03	3.74E-02	2.43E-05
2.50E+02	7.05E-03	4.75E-03	3.76E-02	2.26E-05
2.60E+02	6.95E-03	4.59E-03	3.78E-02	2.10E-05
2.70E+02	6.85E-03	4.43E-03	3.80E-02	1.96E-05
2.80E+02	6.77E-03	4.28E-03	3.82E-02	1.83E-05
2.90E+02	6.69E-03	4.15E-03	3.84E-02	1.72E-05
3.00E+02	6.61E-03	4.02E-03	3.86E-02	1.61E-05
3.10E+02	6.54E-03	3.89E-03	3.87E-02	1.52E-05
3.20E+02	6.48E-03	3.78E-03	3.89E-02	1.43E-05
3.30E+02	6.42E-03	3.67E-03	3.90E-02	1.35E-05
3.40E+02	6.37E-03	3.57E-03	3.91E-02	1.27E-05
3.50E+02	6.32E-03	3.47E-03	3.93E-02	1.20E-05
3.60E+02	6.27E-03	3.38E-03	3.94E-02	1.14E-05
3.70E+02	6.22E-03	3.29E-03	3.95E-02	1.08E-05
3.80E+02	6.18E-03	3.20E-03	3.96E-02	1.03E-05
3.90E+02	6.14E-03	3.12E-03	3.97E-02	9.75E-06
4.00E+02	6.11E-03	3.05E-03	3.98E-02	9.27E-06
4.10E+02	6.07E-03	2.97E-03	3.99E-02	8.83E-06
4.20E+02	6.04E-03	2.90E-03	4.00E-02	8.42E-06
4.30E+02	6.01E-03	2.83E-03	4.00E-02	8.03E-06
4.40E+02	5.98E-03	2.77E-03	4.01E-02	7.67E-06
4.50E+02	5.96E-03	2.71E-03	4.02E-02	7.33E-06
4.60E+02	5.93E-03	2.65E-03	4.03E-02	7.01E-06
4.70E+02	5.91E-03	2.59E-03	4.03E-02	6.71E-06
4.80E+02	5.89E-03	2.54E-03	4.04E-02	6.43E-06
<b>4.90E+02</b>	<b>5.87E-03</b>	<b>2.48E-03</b>	<b>4.05E-02</b>	<b>6.17E-06</b>