

Modeling the Folate Pathway in *Escherichia coli*

by

Carmen Lopez

Department of Mathematics  
Duke University  
Durham, NC

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Anita Layton, Supervisor

\_\_\_\_\_  
Fred Nijhout, Advisor

\_\_\_\_\_  
Mainak Patel

A thesis submitted for honors  
in the Department of Mathematics in Trinity College of  
Duke University

2013

ABSTRACT

Modeling the Folate Pathway in *Escherichia coli*

by

Carmen Lopez

Department of Mathematics  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Anita Layton, Supervisor

\_\_\_\_\_  
Fred Nijhout, Advisor

\_\_\_\_\_  
Mainak Patel

An abstract of a thesis submitted for honors  
in the Department of Mathematics in Trinity College of  
Duke University

2013

Copyright by  
Carmen Lopez  
2013

## **Abstract**

Folates are a class of metabolites that are essential to living cells across all life. Because these molecules are especially important in aiding cell division, folates are targets for antibiotics, such as trimethoprim. I aim to model the pathways of folate production and interconversion in *Escherichia coli*. I began my work by assessing a previously published mathematical model of the *E. coli* folate pathways, published by Kwon and Rabinowitz, which explored the effects of trimethoprim on the system. I explored the model in depth and found avenues for improvement. My model produced enhancements in accuracy and laid the foundation for future work. In the future, this model can be expanded further and be used to model a variety of other experimental conditions under the effect of trimethoprim.

# Contents

Abstract.....	iv
List of Tables.....	vii
List of Figures.....	viii
Acknowledgements.....	ix
1. Preliminary Work.....	10
1.1 Introduction.....	10
1.2 Mathematical Model Setup.....	12
1.3 Previous Studies on the System.....	14
1.4 Ambiguities in a Previous Mathematical Model.....	15
1.5 Improving the Model.....	20
2. Revised Model.....	22
2.1 Methods.....	25
2.2 Steady State under Control.....	26
2.3 Model under Trimethoprim.....	27
2.3 Future work.....	31
2.4 Conclusion.....	32
Appendices.....	33
A.1 The Kwon Model.....	33
A.1.1 Model Ordinary Differential Equations.....	33
A.1.2 Kinetic Constants.....	34
A.1.3 Kwon Model Graphs.....	35
B.1 Revised Model.....	37
B.1.1 Model Ordinary Differential Equations.....	37

B.1.2 Enzyme Kinetic Equations.....	38
B.1.3 Kinetic Constants .....	38

## List of Tables

Table 1: Abbreviations used throughout the thesis .....	14
Table 2: Full Names of the Enzymes in the Model.....	25
Table 3: Parameter Values for Control, Steady State Simulation.....	26
Table 4: Change in Parameter Values over time under drug conditions.....	28

## List of Figures

Figure 1: Structure of folate (PteGlu <sub>n</sub> ).....	10
Figure 2: Structure of DHF (H <sub>2</sub> PteGlu <sub>n</sub> ) [3].....	11
Figure 3: Structure of THF (H <sub>4</sub> PteGlu <sub>n</sub> ) [3].....	11
Figure 4: Functional groups of folates [5].....	12
Figure 5: The effect of trimethoprim on the growth of <i>E. coli</i> .....	15
Figure 6: Model schematic from Kwon 2008. [3].....	16
Figure 7: Kwon model accounting for the discrepancy in $k_1$ in DHF.....	18
Figure 8: Kwon model accounting for the discrepancy in $k_1$ in DHF <sub>2</sub> .....	19
Figure 9: Kwon model accounting for the discrepancy in $k_1$ in DHF <sub>3</sub> .....	19
Figure 10: Revised Model .....	23
Figure 11: Adjusted Kwon model to demonstrate the differences. ....	24
Figure 12: Adjusted revised model to demonstrate the differences. ....	24
Figure 13: Metabolite Concentration Data from Kwon Model.....	28
Figure 14: Change in Parameter Values over time in log scale. ....	29
Figure 15: Change in GCS and SHMT values .....	29
Figure 16: Change in MTHFR1 values .....	30
Figure 17: Change in MTHFR2 values .....	31
Figure 18: Kwon model accounting for the discrepancy in $k_1$ in (C <sub>1</sub> )-THF <sub>1</sub> . ....	35
Figure 19: Kwon model accounting for the discrepancy in $k_1$ in (C <sub>1</sub> )-THF <sub>2</sub> . ....	35
Figure 20: Kwon model accounting for the discrepancy in $k_1$ in (C <sub>1</sub> )-THF <sub>3</sub> . ....	36

## **Acknowledgements**

First, I would like to thank both of my advisors, Prof. Anita Layton and Prof. Fred Nijhout, without whom this project would not have been possible. I would also like to thank the members of the Nijhout Lab for allowing me to work with them. Also, I would like to thank Sophia Kwon for generously providing her data and for corresponding with us to improve the model further. I would also like to thank Inder Jalli for being a great mentor throughout this whole process. I have learned so much by working with him.

I would especially like to thank my parents. They have been supportive of everything that I have done and have always been there for me when I needed them to. I would also like to thank my sister. She's my best friend and has always been there to support me. I would like to thank my entire my family for always supporting me and encouraging me in everything I do. I would not be who I am without them, and I owe everything I have been able to accomplish to them.

# 1. Preliminary Work

## 1.1 Introduction

Folates are a class of metabolites that are essential to living cells across all life. Folates themselves are not biologically active. [1] It is their derivatives, such as tetrahydrofolate (THF) and dihydrofolate (DHF) (shown below) that are necessary for cell function. Each derivative can have an independent number of glutamations. The structure in the brackets in Figures 1-3 is the glutamate. This is an important distinction because derivatives and glutamations are dissimilar.

Folates are necessary to synthesize, repair, and methylate DNA. [2] For example, THF acts as a substrate and coenzyme in transporting one-carbon units for amino acid and nucleic acid metabolism. Since growth is obviously related to cancer and its treatment, folate related pathways are an important target of study. A common cancer treatment that is no longer employed regularly, but is still a topic of study, is an antifolate drug known as methotrexate. Methotrexate competitively inhibits dihydrofolate reductase (DHFR), an enzyme that catalyzes the conversion of DHF to the active THF. The drug also inhibits the synthesis of DNA, RNA, and proteins.

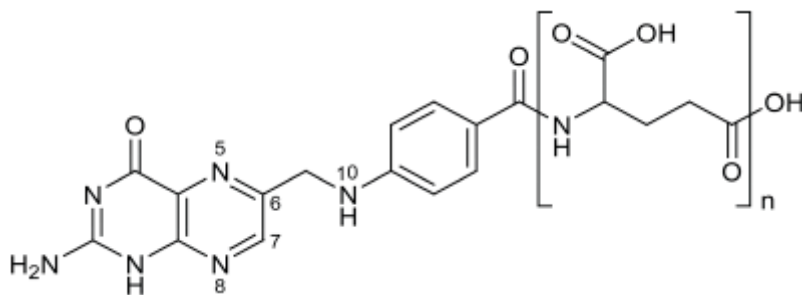


Figure 1: Structure of folate (PteGlu<sub>n</sub>).

Folate can be reduced at C7 and C8 to form DHF (H<sub>2</sub>PteGlu<sub>n</sub>) or at N5, C6, C7, and N8 to form THF (H<sub>4</sub>PteGlu<sub>n</sub>) [3]

Because folates are critical for the cell division process, these molecules are targets for antibiotics as well. I aimed to model the pathways of folate production and interconversion in *Escherichia coli* under treatment with an antibiotic that acts on the folate pathways. I began my work by assessing and replicating a previously published mathematical model of the *E. coli* folate pathways, published by Kwon and Rabinowitz [3]. The Kwon model was used to explore the effects of trimethoprim, an antibiotic that inhibits the enzyme dihydrofolate reductase, DHFR, in the folate pathway of *E. coli*. DHFR is responsible for the reduction of DHF to THF. Figures 2 and 3 below demonstrate the differences in the structures of DHF and THF.

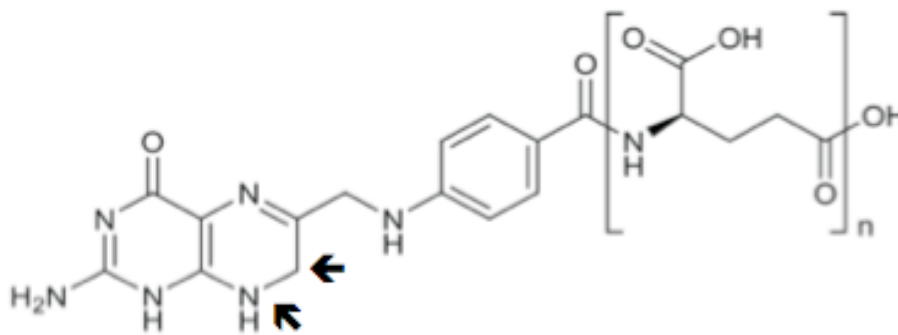


Figure 2: Structure of DHF (H<sub>2</sub>PteGlu<sub>n</sub>) [3]

The arrows point to where folate was reduced in order to create DHF.

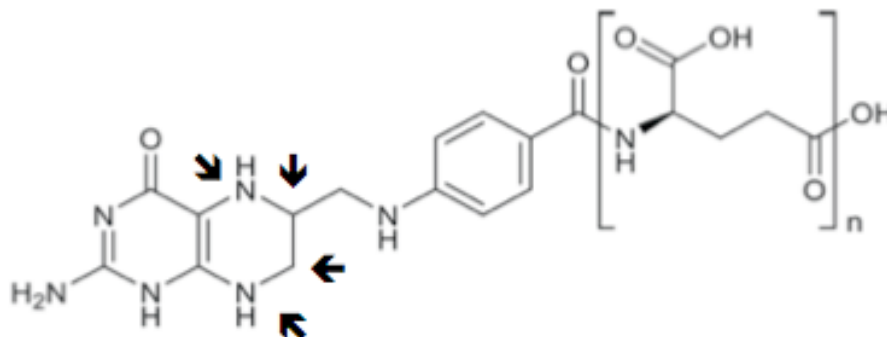


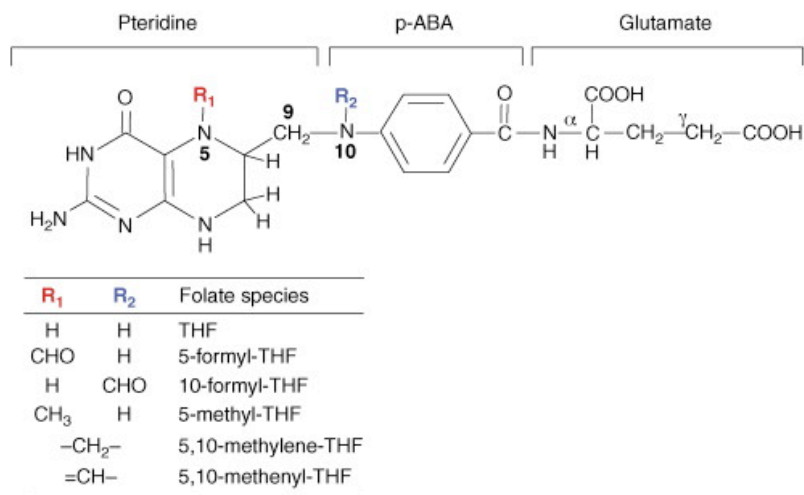
Figure 3: Structure of THF (H<sub>4</sub>PteGlu<sub>n</sub>) [3]

The arrows point to where folate was reduced in order to create THF.

I explored the model in depth and identified aspects for improvement and expansion. I also used another mathematical model, written by Nijhout et al. [4] Both models answered different questions, and but were highly informative as a guide for expansion.

## 1.2 Mathematical Model Setup

Folate metabolites are usually converted into different derivatives with various levels of glutamation by enzymes. In the Kwon model, a very important enzyme was folylpolyglutamate synthase, or FPGS. There are several folate derivatives, defined by the type of functional groups on the folate, such as methyl groups, methylene groups, etc. Several examples are shown below in Figure 4. In addition, each folate derivative can have multiple copies of a glutamate functional group. FPGS attaches glutamates to folate derivatives. Therefore, these folate derivatives may have anywhere between zero and eight glutamates. The Kwon model specifically considers the first three glutamations. Using all of this information I attempted to create an improved model of this system.



TRENDS in Plant Science

Figure 4: Functional groups of folates [5]

Each enzymatic conversion in the model is represented by a different deterministic equation. Some conversions are non-enzymatic. I will now describe a few example equations that have been used to model these kinds of systems. The easiest to use (and therefore information low), variant of an equation to describe enzyme kinetics is the Michaelis Menten equation. When kinetic data for an enzyme target is not richly described in literature, we use this equation:

$$v = \frac{d[P]}{dt} = \frac{V_{\max} [S]}{K_m + [S]}$$

where  $K_m$  is the Michaelis constant,  $[S]$  is the concentration of the substrate, and  $V_{\max}$  is the maximum rate achieved by the enzyme. [6] In the case of no data, the  $K_m$  is a parameter that is fitted to the data. In our case previously published data existed for all target enzymes. For well-studied enzymes, more descriptive and accurate kinetic equations are generally available. When better kinetic data has been published, we employ equations that approximate behaviors such as competitive inhibition. Competitive inhibition is represented by this approximation equation:

$$V_0 = \frac{V_{\max} [S]}{K_m^{app} + [S]}$$

$$K_m^{app} = K_m \left( \frac{1 + [I]}{K_i} \right)$$

where

and where  $K_i$  is the inhibitor's dissociation constant, and  $[I]$  is the inhibitor concentration. [6]  $K_i$  and concentration values are found in previously published studies. Please note that these are only examples of the types on enzyme equations that were

employed. See the appendices for the full list of enzyme kinetic equations used in both the Kwon model and my revised model.

**Table 1: Abbreviations used throughout the thesis**

Short Name	Full Name
PteGlu	folate
DHF	7,8-dihydrofolate
THF	tetrahydrofolate
(C <sub>1</sub> )-THF	sum of 5-methyltetrahydrofolate, 5,10-methylenetetrahydrofolate, tetrahydrofolate
FPGS	folylpolyglutamate synthase
5CH3THF	5-methyltetrahydrofolate
5-methyl-THF	
510mTHF	5,10-methylenetetrahydrofolate
5,10-methylene-THF	
GTP	Guanosine-5'-triphosphate
pABA	para-aminobenzoate
pABGlu	para-aminobenzoylglutamate
5, 10-methenyl-THF	5,10-methenyltetrahydrofolate
5-formyl-THF	5-formyltetrahydrofolate
10-formyl-THF	10-formyltetrahydrofolate
MTHFR	5,10-methylenetetrahydrofolate reductase
TS	thymidylate synthase
folM	dihydrofolate reductase
DHFR	
SHMT	serine hydroxymethyltransferase;
GCS	glycine cleavage system
Pte/pAB_G_1	sum of folate and para-aminobenzoylglutamate

### **1.3 Previous Studies on the System**

Kwon collected very high quality enzyme kinetic data for FPGS and folate concentration data for her experiments. *E. coli* K-12 strain NCM3722 was dosed with 4 ug/mL trimethoprim in growth media. Folate concentrations were then measured at six different time points, and the concentration data of each metabolite was found using an LC-MS/MS methodology. The raw ion counts from the LC-MS/MS were converted into micromolar concentrations based on the ion count signals of the internal standard, 5-

methyl-THF<sub>n</sub>. This was accomplished via a series of calculations and cell density approximations and produced the relative concentrations of each species. The cell density from previous growth experiments were used and extrapolated to fit this set of time points. The absolute concentrations were calculated using the optical densities at each time point. The figure below demonstrates how treatment with trimethoprim affects the growth of *E. coli*. Trimethoprim was added at the time point labeled TM.

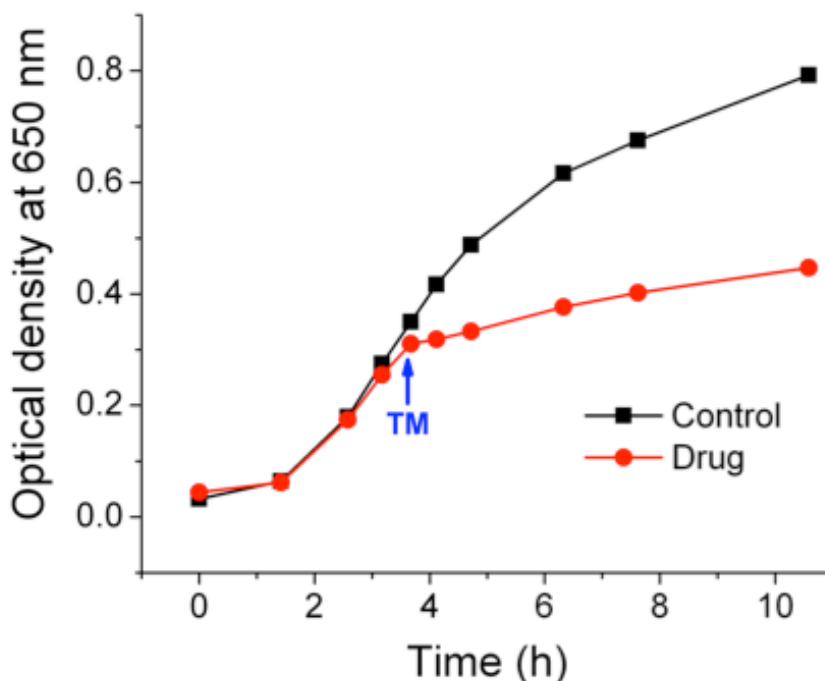


Figure 5: The effect of trimethoprim on the growth of *E. coli*.

4 ug/mL of trimethoprim was added at the time point labeled TM. [3]

#### 1.4 Ambiguities in a Previous Mathematical Model

I began by deconstructing and replicating a previously published model of the folate network, the Kwon model, and identifying aspects for improvement. [3] The Kwon model focused on the conversions between DHF<sub>n</sub> and (C1)-THF<sub>n</sub>, the addition of glutamates by FPGS, and how the system behaved under treatment with trimethoprim. The source term (see Figure 6, below) creates DHF<sub>1</sub>, the monoglutamate species. DHF<sub>n</sub> is

capable of being catabolized or anabolized. In the model, this is represented as  $k_5$  and is the flux out of each DHF species. Also, each  $\text{DHF}_n$  can be converted into its respective  $(\text{C}_1)\text{-THF}_n$  and vice versa since this reaction is reversible. FPGS is the enzyme that adds glutamates to  $(\text{C}_1)\text{-THF}_1$  and  $(\text{C}_1)\text{-THF}_2$ . The other sink in the model, represented by  $k_6$ , is the flux from  $(\text{C}_1)\text{-THF}_3$  to  $(\text{C}_1)\text{-THF}_4$ . This glutamate addition is performed by FPGS and is not considered in the model. The model schematic is shown below in Figure 6, and the exact model equations are given in the appendix.

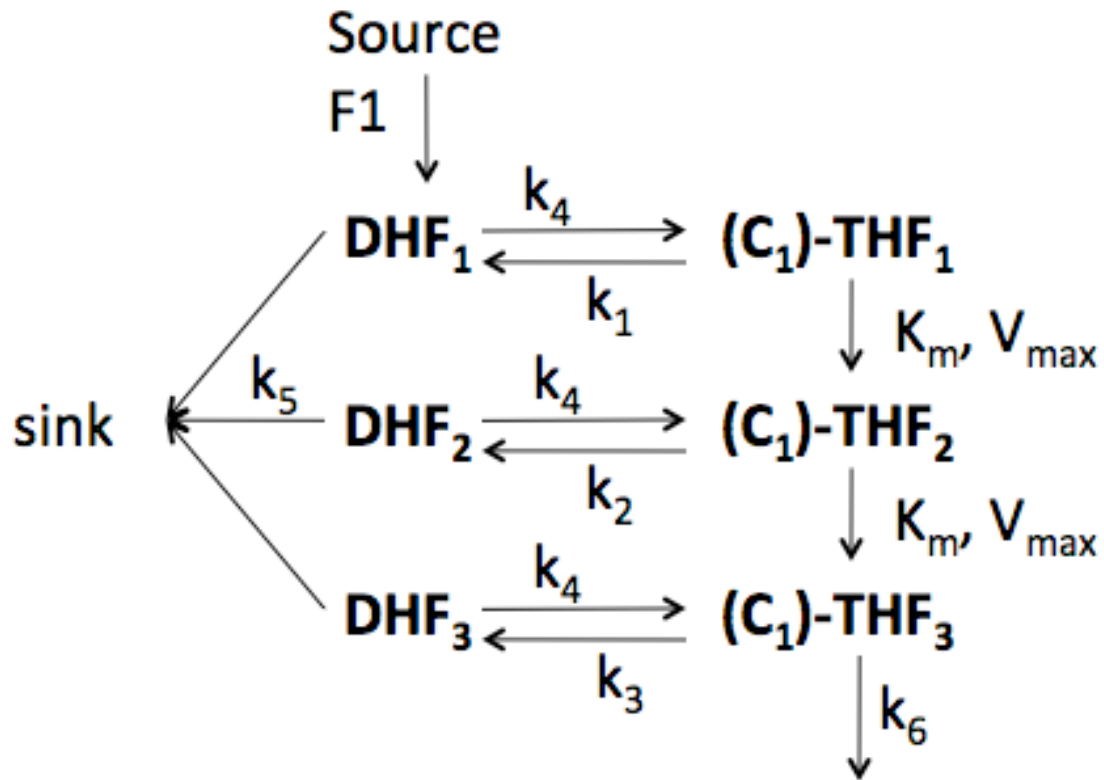


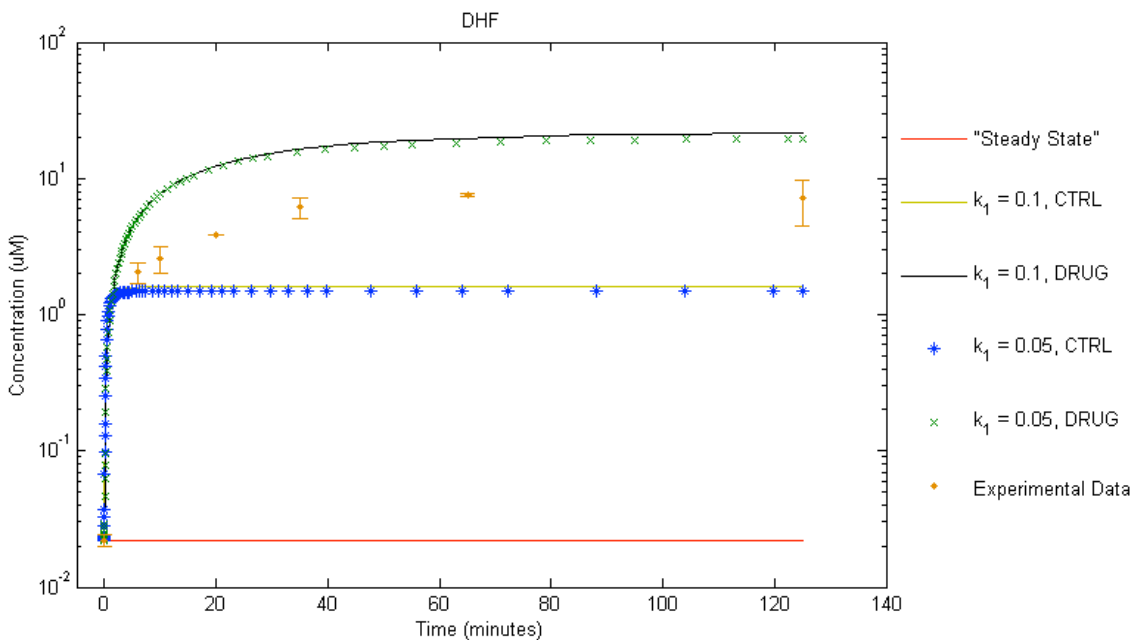
Figure 6: Model schematic from Kwon 2008. [3]

The subscripts represent the number of glutamations.  $(\text{C}_1)\text{-THF}_n$  represents the sum of all measured reduced folates containing  $n$  glutamates, which are  $\text{THF}_n$ , 5,10-methylene- $\text{THF}_n$ , and 5-methyl- $\text{THF}_n$ . F<sub>1</sub> is the source term into DHF<sub>1</sub>.  $k_5$  represents the outward flux of DHF<sub>n</sub> into the sink, in other words all conversions away from DHF<sub>n</sub> that do not lead to  $(\text{C}_1)\text{-THF}_n$ .  $k_6$  represents the flux into the tetraglutamates, in other words all conversions away from  $(\text{C}_1)\text{-THF}_3$  that do not lead to DHF<sub>3</sub>. [3]

I began my study of the system by examining the procedures and methods behind the Kwon model. As I examined their explanation of the model, it became clear that there was an inconsistency in one of the Kwon model kinetic constants,  $k_1$ , which is the conversion of  $(C_1)$ -THF<sub>1</sub> into DHF<sub>1</sub>. In their mathematical model code, which was written in the program R,  $k_1$  had the value of 0.05. However, in the published paper,  $k_1$  was equal to 0.1. The correct value of  $k_1$  is unclear. Because of this discrepancy, I performed simulations with both values of  $k_1$  under control and drug conditions, in order to see if this small error has a significant effect on simulation outcomes. The change makes little difference in the time course evolution of both control and drug conditions, which is evident in the figures below. Therefore, either value could be equally correct. After performing simulations of the Kwon model, it was clear that the model matched the data very well qualitatively. However, the model did not fit the experimental data quantitatively, or within 10%, which is the threshold I used in my mathematical modeling.

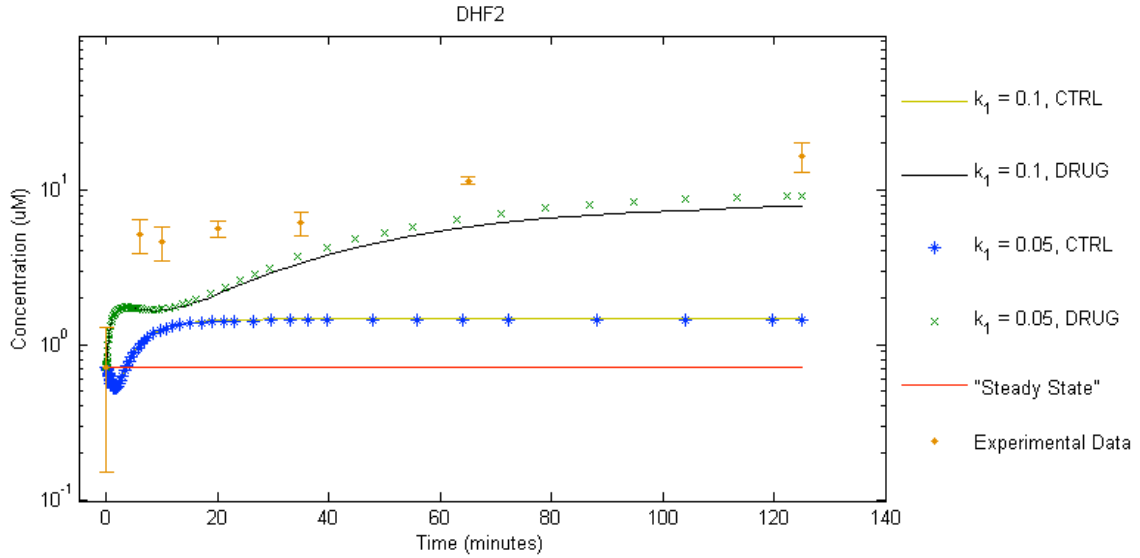
The figures below demonstrate several points. In the figures, the dots with error bars are the experimental folate concentration data for the drug condition, and the line is the experimental steady state, assuming that the folate concentration at time point zero is the steady state. First, the figures show how varying  $k_1$  to take into account the discrepancy affects the behavior of the concentrations. The line with stars represents the steady state simulation for the different  $k_1$  values. The line with X's also represents the varying  $k_1$  values under drug conditions. In general, the differences in  $k_1$  values are not significant and do not have a large effect on Kwon's conclusions. Also, the steady state simulation that is provided is not truly a constant steady state. The graphs also show that the model for the drug data does not fit the experimental values quantitatively. The line with "X" markers are what the model produced for folate concentrations under

drug conditions. At times the model predictions are orders of magnitude away from the experimental drug data. All of this is shown below in Figures 7-9, and additional graphs are in the appendix, which show the progression for each metabolite concentration modeled by Kwon et al.

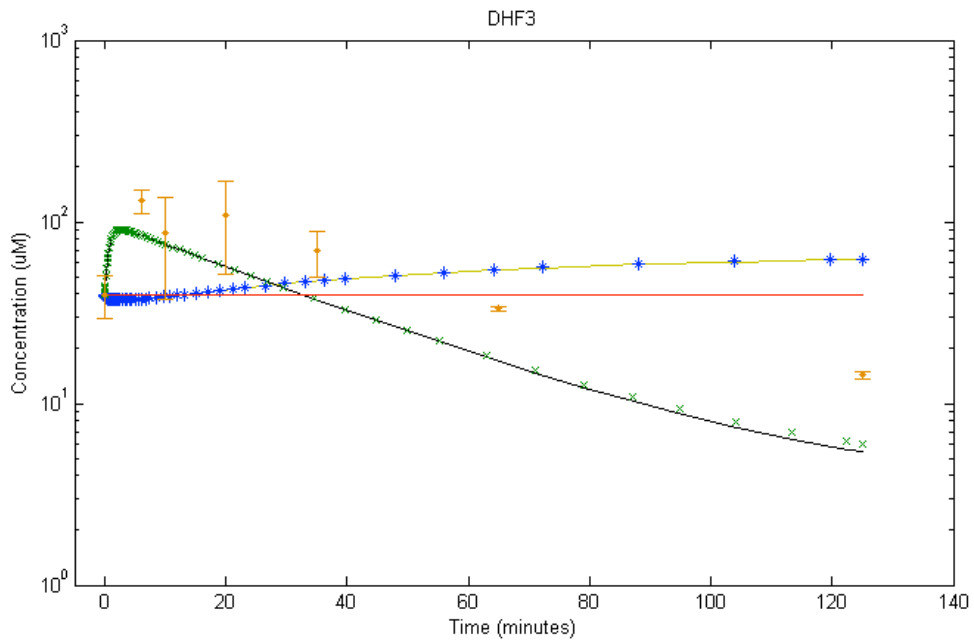


**Figure 7: Kwon model accounting for the discrepancy in  $k_1$  in DHF.**

The unadorned line labeled "Steady State" is the proposed steady state based on the experimental folate concentration data, and the orange dots are the experimental data with error bars. The lines labeled CTRL and DRUG are the values produced by the Kwon Model. By looking at the graph it is clear to see the differences between the model CTRL lines and the "Steady State" and the differences between the model DRUG lines and the experimental data. Changing  $k_1$  appears to make little difference. Keep in mind that these graphs are in log scale when looking at the errors.



**Figure 8: Kwon model accounting for the discrepancy in  $k_1$  in DHF2.**



**Figure 9: Kwon model accounting for the discrepancy in  $k_1$  in DHF3.**

Two of the Kwon model parameters were fitted to match the experimental data qualitatively. These parameters were the  $V_{max}$  of FPGS and  $k_s$ , the sink out of DHF. I attempted to vary  $V_{max}$  but found that both doubling and halving the  $V_{max}$  was not

helpful in getting any closer to a steady state approximation. The parameters of the Kwon model were only fit to the experimental drug data and not to the steady state.

### **1.5 Improving the Model**

I explored several possible methods for expanding and improving upon the model. The first was to expand the number of folates. The Kwon model simulates  $(C_1)$ -THF, the sum of reduced folates, which is the sum of 5-methyl-THF, THF, and 5,10-methylene-THF, instead of exploring each metabolite separately. I divided up  $(C_1)$ -THF into its separate components, 5-methyl-THF, THF, and 5,10-methylene-THF, so that I can model each metabolite individually. In order to do this, I focused only on the monoglutamates. Therefore, I temporarily removed the extra glutamations in order to expand the number of folates. The model had to be simplified before it could be built upon again. In the future, the polyglutamates will be reinserted back into the model.

Then I reassessed the sources and sinks. The source represented the de novo synthesis of DHF from GTP, pABA, and glutamate. GTP is Guanosine-5'-triphosphate. pABA is para-aminobenzoic acid. Glutamates were discussed in the introduction. All are essential intermediates in the role of folate production. I researched the pathway by which DHF is constructed and decided not to include the concentrations of the intermediate metabolites, GTP, pABA, and glutamate, into the model. This is because these metabolites would provide little value to understanding the system. Therefore, the source is a simple linear mass input term, F1, as in the Kwon model. The sink represented the catabolism of DHF to para-aminobenzoylglutamate, pABGlu<sub>n</sub>, and oxidation of DHF to folate, PteGlu<sub>n</sub>. The sink is represented by  $k_5$ , which is, in other words, all conversions away from DHF<sub>n</sub> that do not lead to  $(C_1)$ -THF<sub>n</sub>. It is not known how these reactions occur. It has been postulated that the catabolism could occur enzymatically or by spontaneous cleavage and that the oxidation may be a nonspecific

chemical reaction. [7-9] Because of this lack of information, there will be one sink term encompassing both of these reactions, and it will be a fitted parameter, similar to the Kwon model.

The next modeling targets were the variety of experimental conditions. The Kwon model explored the system under normal growth conditions and then also under treatment with trimethoprim. I modeled the same conditions: the control and the perturbation under trimethoprim. A note on perturbation: often in biological studies, perturbation refers to a gene knockout or some sort of permanent change to the system being studied. In our case, perturbation means the introduction of a foreign substance, in this case trimethoprim, into the growth mixture, which severely cripples a part of the pathway. In our case, the functional element being disabled is still in the system, just at a vastly reduced capacity.

In order to understand how Kwon et al built their model, I verified the calculations to find the kinetic constants. In doing so I re-evaluated some of the constant values, such as  $k_1$ , the conversion of  $(C_1)$ -THF<sub>1</sub> to DHF<sub>1</sub>, mentioned earlier. Kwon et al used a steady state assumption, leading to an assumption that the conversion rate of each DHF<sub>n</sub> to THF<sub>n</sub> is the same as the conversion rate of THF<sub>n</sub> to DHF<sub>n</sub>. This is a valid assumption when attempting to ascertain the steady states.

My next task was to determine how Kwon et al calculated the fluxes of the glutamated species and to see if there was a way I could exploit this data further in my model. Kwon equation was used:

$$f_n = X_n^T \cdot k_n$$

where  $f_n$  is the flux through each polyglutamated species,  $X_n^T$  is the total pool of folate species with  $n$  glutamates, and  $k_n$  is the labeling data fit to the equation:

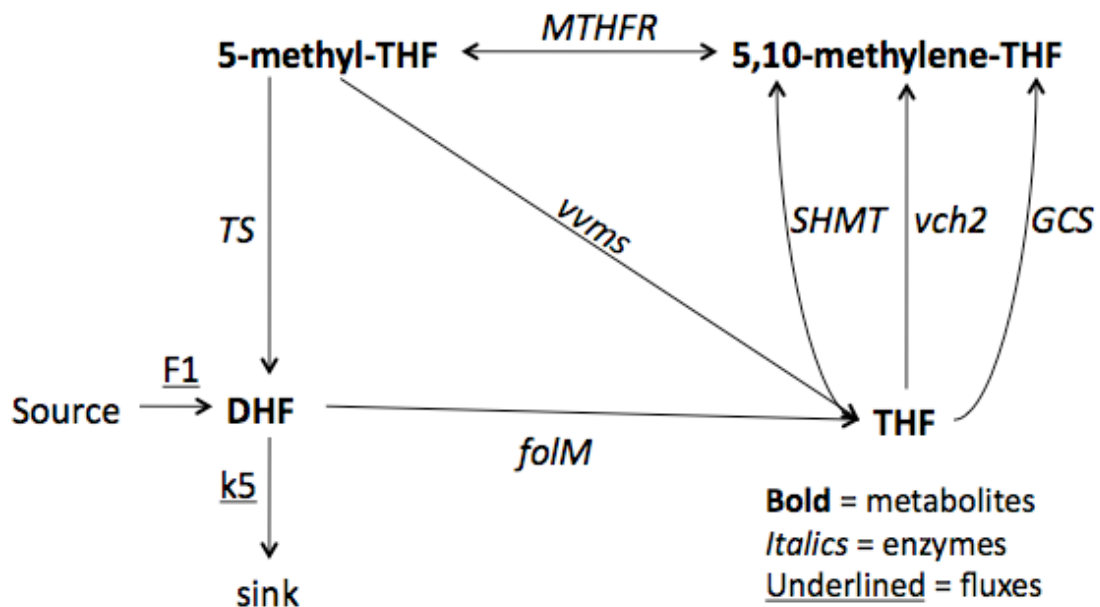
$$X_n^U = X_n^T \cdot \left[ \left( \frac{k'_{n-1}}{k'_{n-1} - k_n} \right) \cdot e^{-k_n t} - \left( \frac{k_n}{k'_{n-1} - k_n} \right) \cdot e^{-k'_{n-1} t} \right]$$

$X_n^U$  is the unlabeled form of a folate species with  $n$  glutamates in its tail, and  $k'_{n-1}$  is a single exponential fit to the data for  $X_{n-1}$ , the precursor to  $X_n$ . [3] In order to do this I used curve fitting to an exponential to determine  $k'_{n-1}$  using the sum of least squares method, which then allowed me to determine  $k_n$  by fitting to the above equation by again using the sum of least squares. To find  $X_n^T$ , I summed up all the concentrations of the metabolites of a given glutamation. Then, I was able to determine the flux using the equation above. Kwon et al used the whole 120 minute time scale to calculate one flux value. My initial goal was to divide this flux data into two or more parts within the time scale. Doing this would have allowed me to deduce different flux values within a time period. I was able to subdivide the flux into the two parts, however there were not enough data points to make the flux estimates useful or reliable.

## 2. Revised Model

My revised model consisted of additional folates and parameters. I chose to sacrifice the glutamations in order to expand the sum (C<sub>1</sub>)-THF into its individual metabolites. This allowed me to explore the behavior of each metabolite separately. The source and the sink were both located at DHF<sub>1</sub>. The sink from the Kwon model represented the catabolism of DHF to para-aminobenzoylglutamate, pABGlu<sub>n</sub>, and oxidation of DHF to folate, PteGlu<sub>n</sub>. Not much is known about these two reactions. Therefore, the flux of the sink was parameter fitted to the data. Also, several of my original metabolites could not be modeled. These metabolites include 5, 10-methenyl-THF, because no concentration data existed for it under trimethoprim, and 5-formyl-THF and 10-formyl-THF, because these are not detectable when only glutamated once.

Since the polyglutamations are a significant aspect of this system, these additional folates will be modeled in the future. Figure 10 displays the model I used, and Table 2 gives the full names of each enzyme. Figures 11 and 12 are present to make the comparison between the Kwon model and my revised model easier. Because of these various changes the equations are different from the Kwon model's. The new model ODEs described by this diagram are presented in the appendix. Instead of using mostly linear relationships, like the Kwon model, I used enzymatic kinetic equations, such as Michaelis Menten type mentioned earlier. The exact kinetic equations are also listed in the appendix.



**Figure 10: Revised Model**

The enzyme *TS*, thymidylate synthase, plays a critical role in DNA synthesis.

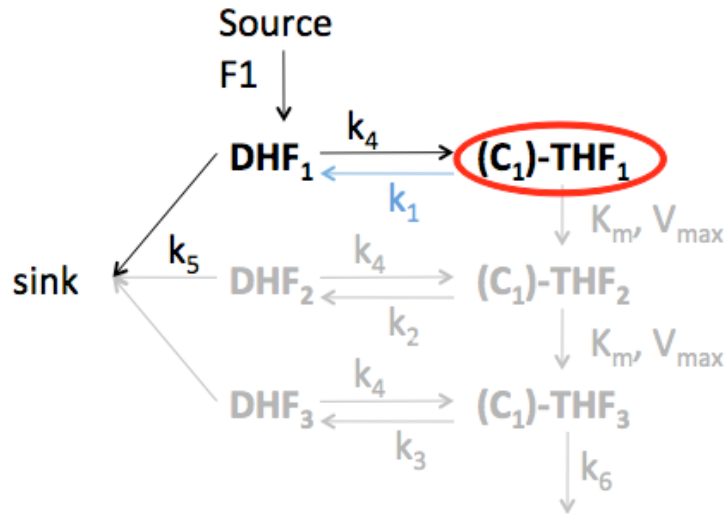


Figure 11: Adjusted Kwon model to demonstrate the differences.

The polyglutamations were grayed out to make it clear that I will not be considering those in my model. Recall that (C<sub>1</sub>)-THF<sub>1</sub> is a sum of three metabolites. Therefore, the conversion k<sub>1</sub> represents three metabolites being converted into DHF<sub>1</sub>.

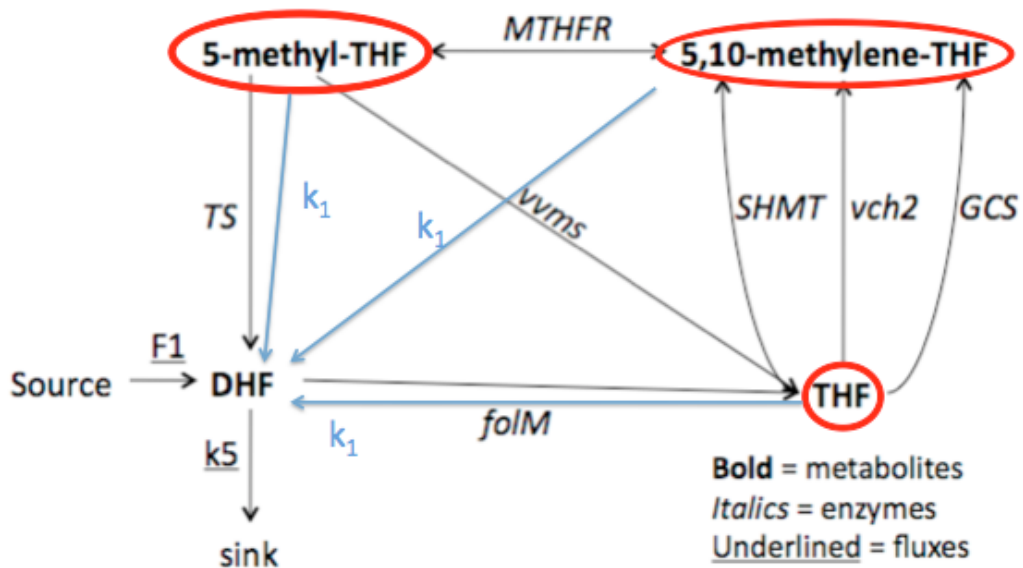


Figure 12: Adjusted revised model to demonstrate the differences.

The revised model only considers the monoglutamates. The sum from the Kwon model, (C<sub>1</sub>)-THF<sub>1</sub>, was split into the three circled metabolites in the revised model. The conversions for k<sub>1</sub> have been placed on the revised model to demonstrate how much the complicated nature of the system has increased. Many reactions were being represented by the one k<sub>1</sub> value in the Kwon value. These reactions have been separated, which provides the reason for only studying the monoglutamates.

**Table 2: Full Names of the Enzymes in the Model**

Abbreviation	Full Name
TS	thymidylate synthase
MTHFR	5,10-methylenetetrahydrofolate reductase
vvms	combination of methionine synthase and homocysteine transmethylase
foIM	dihydrofolate reductase
SHMT	serine hydroxymethyltransferase
vch2	non-enzymatic reaction
GCS	glycine decarboxylase

## **2.1 Methods**

For my model, I used the metabolite concentration data from the Kwon experiments. With their data, I was able to model the system. This system could not be approached analytically due to the large number of nodes and the complicated nature of the system. As a result, the equations were modeled numerically using a computer. I used the ODE solver, ode15s, in Matlab to solve a series of linked ordinary differential equations describing the change in concentration of each metabolite over time. Ode15s is a stiff multistep solver using the numerical differentiation formula method and employs a variable time step. A stiff equation has a solution that is not stable unless the time step is sufficiently small. By using ode15s, a little accuracy was sacrificed in order to gain speed. However, because the model predicted the folate concentration within 10% of the experimental value, using a higher accuracy solver was not deemed necessary. The simulations were run on an Intel Q8200 @ 2.33GHz using Microsoft Windows 7 and Matlab r2012b. The simulations ran for 7500 seconds, the total time for which folate concentration data was gathered. The enzyme kinetic parameters I used were found in BRENDA [10] or in other various papers noted in the appendix. If no information was available from previous experiments, then the kinetic parameters were found manually by fitting to the data. The kinetic parameters were constant throughout each simulation. The metabolite concentrations did vary throughout the simulation because those were

the target of study described by the ODEs. The enzyme  $V_{\max}$  parameters were fit to the data. The parameter estimations were performed manually running simulations with various parameters under the model. The predictions fit the experimental data. The  $V_{\max}$  and flux values were chosen so that the concentration of each metabolite was at least within 10% of the correct value for both the steady state of the control condition and the drug condition. All concentrations,  $K_m$ 's, and  $K_i$ 's mentioned in this paper are in concentration units micromolar. All fluxes and rate constants are in inverse seconds. All  $V_{\max}$ 's are in micromolar per second.

## 2.2 Steady State under Control

A new steady state was found under the control conditions using my model. The data was matched quantitatively to model predictions guaranteeing that each concentration was within 10% of the experimental steady state value. I refrained from changing the kinetic constants, such as  $K_m$ 's and  $K_i$ 's, because those have been verified by multiple previous studies. The kinetic constants are also global parameters meaning that the parameters do not change under different initial conditions. Therefore, I focused on changing the  $V_{\max}$ 's because these parameters are products of the treatment condition. I changed the flux values in the sink and source only after I found that altering the  $V_{\max}$ 's would not give the desired result. The fluxes had to be changed from the Kwon model because I was only considering the single glutamation species; therefore, the flux out of the single glutamation had to be different than the flux out of the triglutamates. The enzyme  $V_{\max}$  and flux parameter values for the control condition are reported below. The  $K_m$  values used in all simulations are reported in the appendix.

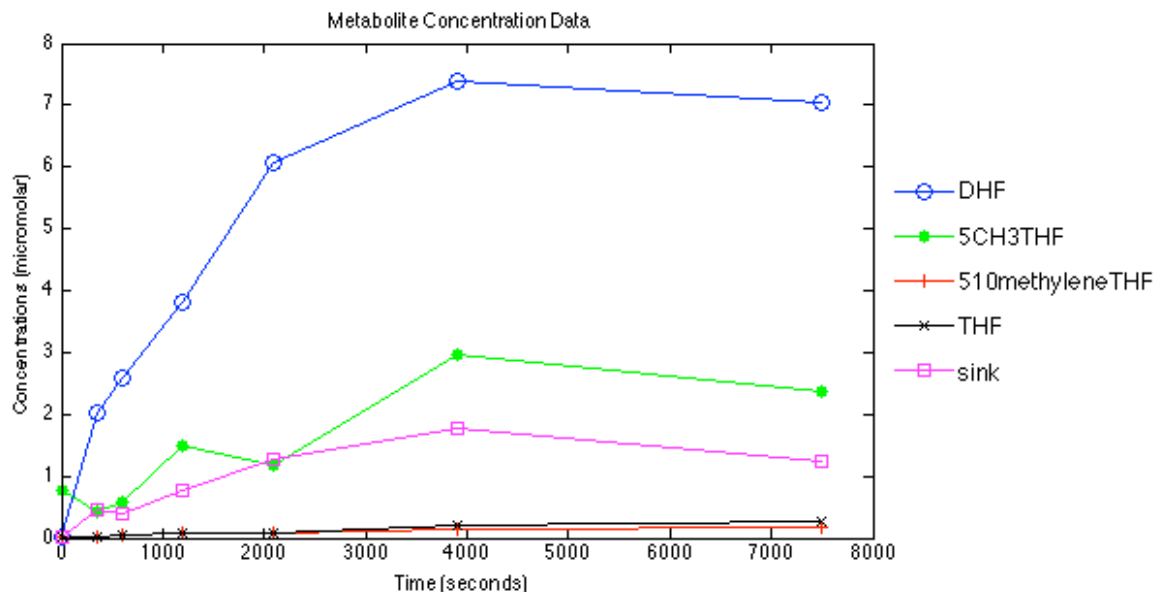
**Table 3: Parameter Values for Control, Steady State Simulation**

Parameters	Value ( $\mu\text{M}$ )
folM_Vmax	50
SHMT_Vmax	10000

GCS_Vmax	15000
MTHFR1_Kcat	220
MTHFR2_Kcat	665
vvms_Vmax	75
TS_Vmax	1800
F1_flux	0.0055
K5	0.25

### **2.3 Model under Trimethoprim**

Next I modeled the metabolite concentration data changes under trimethoprim conditions, shown in Figure 11 below. I initially attempted to find one set of parameters to match the data. However, I discovered that it was not possible to match one set of  $V_{max}$  and flux values to the metabolite concentration time course data, so I concluded that the parameter values must change over time. Therefore, another method had to be used to match the data. I divided the simulation into several steps. Each time point was treated as its own steady state with its own parameter values. Each simulation ran for as long as that section of time lasted before the next time point. By using this method, I found the parameter values at each specific time point and was able to see how the parameter values changed over time. I again focused on changing the  $V_{max}$  parameters first and then changed the flux into the sink, k5, when changing the  $V_{max}$ 's alone did not fit the data. I kept the flux for the source, F1, constant at the same ratio from control to drug that Kwon provided in her work. Therefore, the flux values I used were different than the Kwon model's, but the values were kept at the same ratio in order to maintain consistency. The parameter values for the drug condition are shown below in Table 4. This change in the  $V_{max}$  value can be attributed to the constant up regulation, down regulation, and allosteric effects of the system in order for the *E. coli* to maintain itself.



**Figure 13: Metabolite Concentration Data from Kwon Model**

**Table 4: Change in Parameter Values over time under drug conditions**

Change in parameters over time under Trimethoprim							
Parameters\Time (min)	0	6	10	20	35	65	125
folM_Vmax	50	20.6	20.6	20.6	20.6	20.6	20.6
SHMT_Vmax	10000	50000	35000	30000	30000	20000	20000
GCS_Vmax	15000	50000	35000	30000	30000	20000	20000
MTHFR1_Kcat	220	500	200	200	167	300	291
MTHFR2_Kcat	665	665	665	400	400	400	400
vvms_Vmax	75	100	350	350	375	433	750
TS_Vmax	1800	54600	8500	3500	4000	2500	1750
F1_flux	0.0055	0.00196	0.00196	0.00196	0.00196	0.00196	0.00196
K5	0.25	0.001	0.000762	0.00052	0.000324	0.000266	0.000278

In the above chart, time point 0 represents the control values because the drug was never administered at exactly time point 0, and the effects could not have been measured immediately. Therefore, time point 0 has a different flux value. The graph of these changes for all of the parameters are shown below in Figure 12. Some individual parameter values are displayed in Figures 13-15.

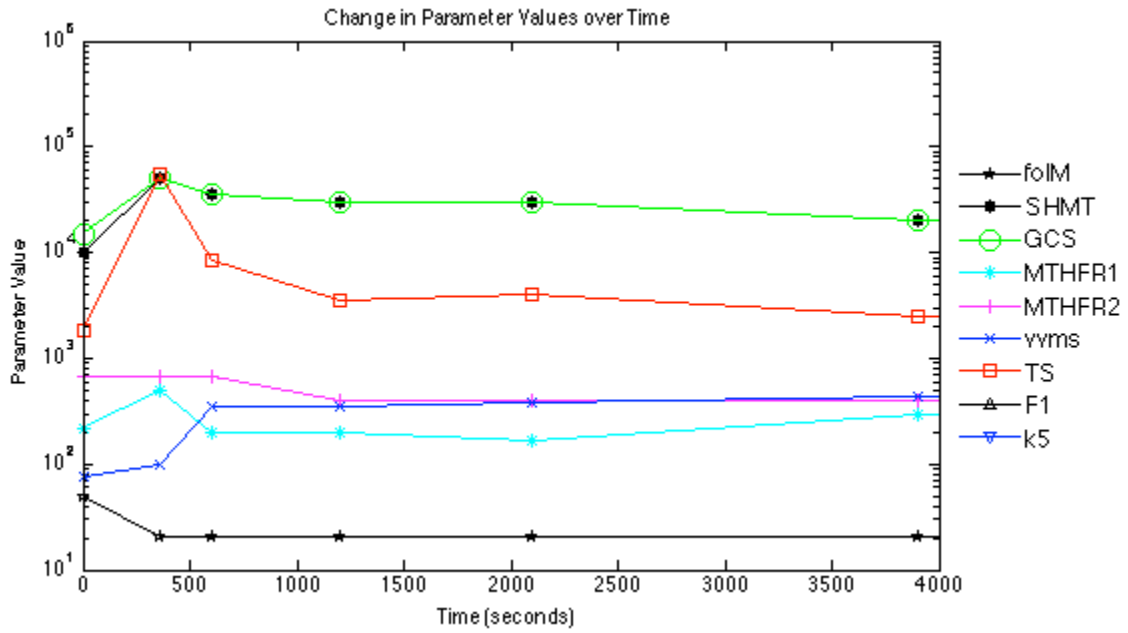


Figure 14: Change in Parameter Values over time in log scale.

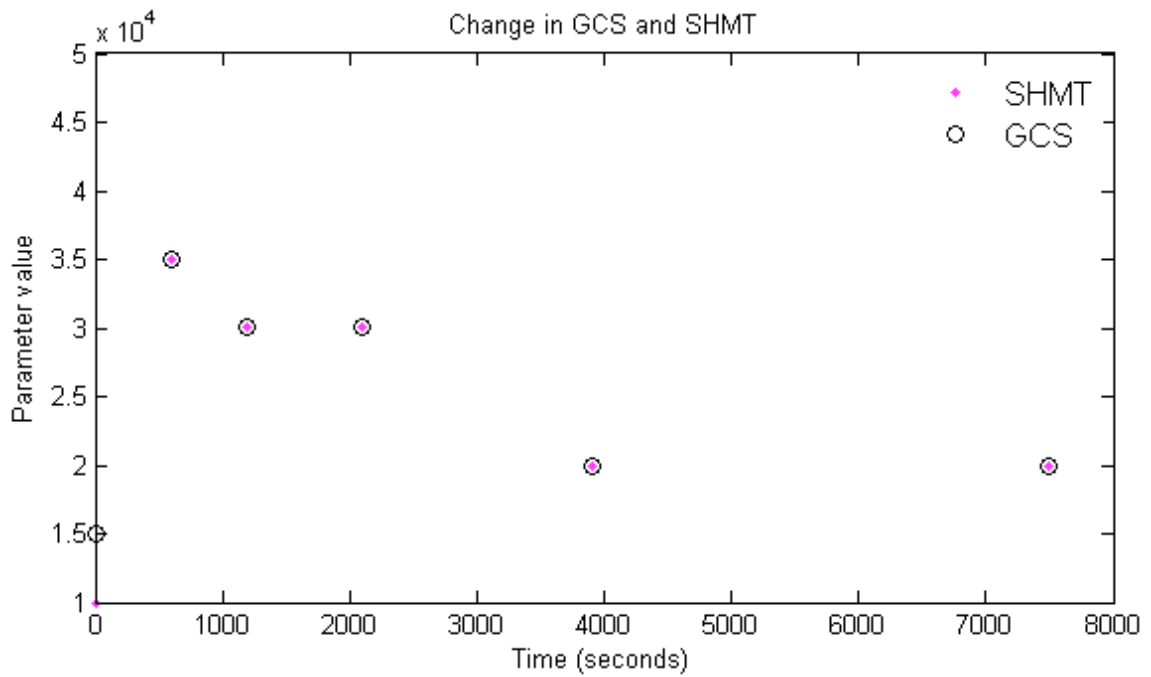
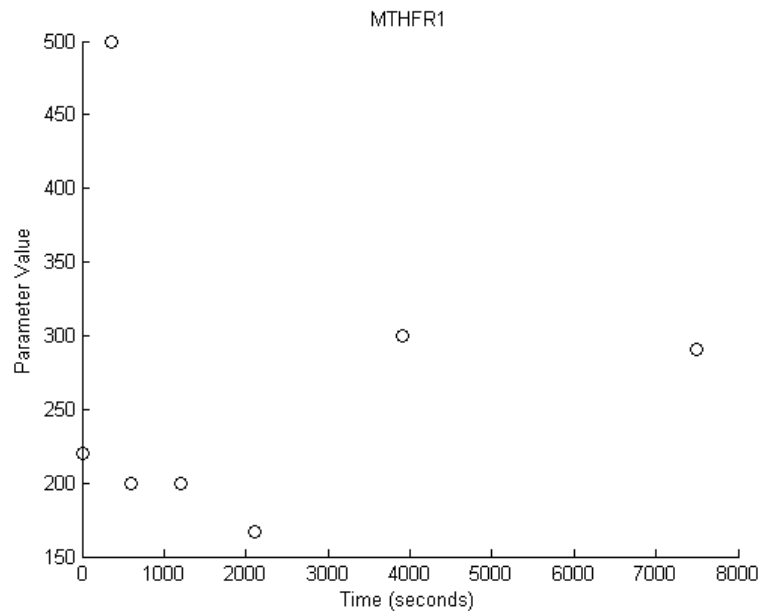


Figure 15: Change in GCS and SHMT values

There is a clear downward trend.

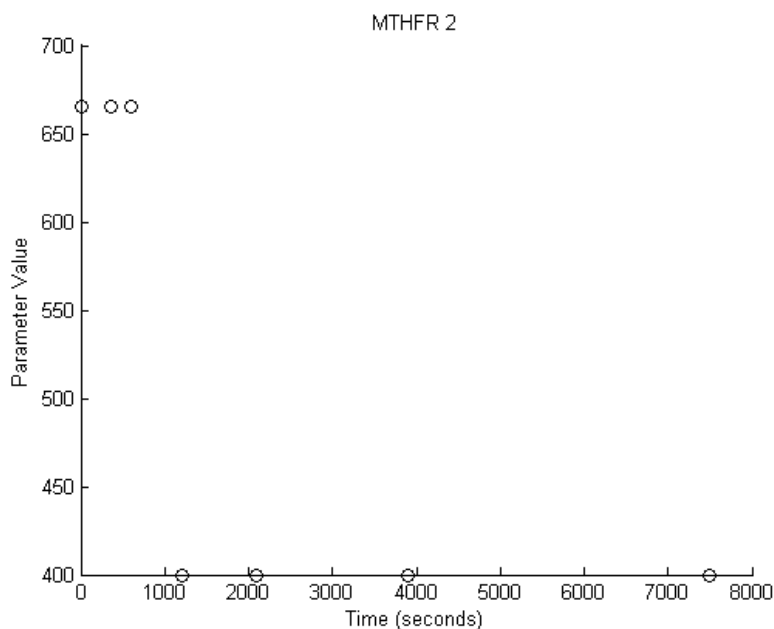


**Figure 16: Change in MTHFR1 values**

**MHFR1 appears to have no real trend.**

There was great variety in the changes in parameter values. Some parameters, such as GCS and SHMT, had a clear downward trend, shown above in Figure 13. The parameters TS and k5 (not shown) had a clear downward trend as well. Others, such as F1 and folM (not shown), remained relatively unchanged. Some, such as MTHFR1 shown above in Figure 14, had no discernible trend across time points. Other parameters show evidence of something more interesting. MTHFR2, shown in Figure 16, provides some evidence for an initial shock and a recovery. What this appears to show is that there is an initial shock from the drug treatment, and after the first two or three time points it recovers and displays an evident pattern. For the first three time points, MTHFR2 maintains a parameter value of 665, but then immediately falls to 400 for the remainder of the time points. This demonstrates that the initial shock of the addition of trimethoprim caused this higher parameter value. Once the system recovers from the initial shock the parameter value changed to what it would be under normal conditions.

While MTHFR2 follows this recovery path, its reverse, MTHFR1 (shown below), has no real pattern. For full disclosure, these trends may change once the polyglutamates are added back into the model.



**Figure 17: Change in MTHFR2 values**

### **2.3 Future work**

My model successfully simulated the single glutamation metabolites in this system and laid the foundation for improvements. The model had to be simplified in order to build upon it again. The next step in modeling this system would be to add the additional polyglutamations back into the model. Kwon has concentration data for up to five levels of glutamation. I was unable to use this data for this current work because doing so requires the  $K_m$ 's for each enzyme at every glutamation. The data on those kinetic constants is sparse. Adding the glutamations will allow 5-formyl-THF<sub>n</sub> and 10-formyl-THF<sub>n</sub> to be included in the model because these folates are negligible at the monoglutamate level but present at higher glutamations. Integrating data for 5, 10-methenyl-THF<sub>n</sub> under trimethoprim is also a priority. A genetic algorithm based

parameter fit will also be performed in order to estimate these parameters. There are too many to fit manually because of the large number of enzymatic reactions once the polyglutamations are added.

Kwon et al also explored a variety of other experimental conditions, including carbon, nitrogen, and phosphate starvation experiments and data under treatment with various supplements such as inosine. In the future the starvation scenarios could be modeled as separate independent states of *E. coli*. Inosine was found to offset the effects of trimethoprim. This rescue from Inosine can be studied by modeling how the Inosine affects the system.

Now that we have ascertained the parameter trends in the drug condition, we can go back and reintegrate the parameter changes back into the model. We can find equations to fit the changes in parameters and use those equations to vary the  $V_{\max}$  over time. Then, I can attempt a single simulation attempt a better prediction of system behavior.

## **2.4 Conclusion**

Folates are a class of metabolites that are essential to living cells across all life. Because folates are especially important in aiding cell division, these molecules are targets for cancer treatments and antibiotics. Therefore, it is imperative to improve our understanding of folates. The revised model is a fundamental part of a total revision that will greatly expand the predictive ability and understanding of the system in a variety of contexts.

# Appendices

## A.1 The Kwon Model

### A.1.1 Model Ordinary Differential Equations

ODEs:

These are the generating changes for how variable changes are re-calculated per time step. Each  $dx/dt$  represents one variable change.

$$d([(C1)-THF2])/dt = 1/unnamed*(ReactionFlux3 - ReactionFlux4 + ReactionFlux7 - ReactionFlux8)$$

$$d([(C1)-THF1])/dt = 1/unnamed*(ReactionFlux1 - ReactionFlux2 - ReactionFlux7)$$

$$d(DHF2)/dt = 1/unnamed*(-ReactionFlux3 + ReactionFlux4 - ReactionFlux11)$$

$$d([(C1)-THF3])/dt = 1/unnamed*(ReactionFlux5 - ReactionFlux6 + ReactionFlux8 - ReactionFlux10)$$

$$d(DHF3)/dt = 1/unnamed*(-ReactionFlux5 + ReactionFlux6 - ReactionFlux13)$$

$$d(DHF1)/dt = 1/unnamed*(-ReactionFlux1 + ReactionFlux2 + ReactionFlux9 - ReactionFlux12)$$

$$d(sink)/dt = 1/unnamed*(ReactionFlux11 + ReactionFlux12 + ReactionFlux13)^*$$

$$d([(C1)-THF4])/dt = 1/unnamed*(ReactionFlux10)$$

Fluxes:

These are the equations that are used in the  $dx/dt$  variable calculations. Please note that some are linear while reaction flux 7 and 8 are a Michaelis Menten style approximation.

$$ReactionFlux1 = K4*DHF1$$

$$ReactionFlux2 = K1*[(C1)-THF1]$$

$$ReactionFlux3 = K4*DHF2$$

$$ReactionFlux4 = K2*[(C1)-THF2]$$

$$ReactionFlux5 = K4*DHF3$$

$$ReactionFlux6 = K3*[(C1)-THF3]$$

$$ReactionFlux7 = ((Vmax*[(C1)-THF1]/Km)/(1+DHF1/KI+DHF2/KI+[(C1)-THF1]/Km+[(C1)-THF2]/Km))$$

---

\* Please note that the sink represents the sum of pABGlu<sub>n</sub> and folate.

$$\text{ReactionFlux8} = ((V_{\max} * [(C1)\text{-THF2}] / K_m) / (1 + \text{DHF1} / K_I + \text{DHF2} / K_I + [(C1)\text{-THF1}] / K_m + [(C1)\text{-THF2}] / K_m))$$

$$\text{ReactionFlux9} = .01667$$

$$\text{ReactionFlux10} = K_6 * [(C1)\text{-THF3}]^+$$

$$\text{ReactionFlux11} = K_5 * \text{DHF2}^{\circ}$$

$$\text{ReactionFlux12} = K_5 * \text{DHF1}$$

$$\text{ReactionFlux13} = K_5 * \text{DHF3}$$

### A.1.2 Kinetic Constants

Parameter Values for control condition:

$$K_I = 3.1$$

$$K_m = 50$$

$$K_4 = 0.0333$$

$$K_1 = 0.000833$$

$$K_2 = 0.01$$

$$K_3 = 0.02$$

$$V_{\max} = 1$$

$$K_6 = 0.0001$$

$$K_5 = 0.0005$$

$$F_1 = 0.0466$$

Parameter Values for drug condition:

$$K_I = 3.1$$

$$K_m = 50$$

$$K_4 = 0.0005$$

$$K_1 = 0.000833$$

$$K_2 = 0.01$$

$$K_3 = 0.02$$

$$V_{\max} = 1$$

$$K_6 = 0.0001$$

$$K_5 = 0.0005$$

$$F_1 = 0.01667$$

---

<sup>+</sup> K<sub>6</sub> represents all conversions away from THF3 that do not go to DHF3.

<sup>°</sup> K<sub>5</sub> represents all conversions away from DHF1, DHF2, and DHF3 that do not go to their respective (C1)-THF.

### A.1.3 Kwon Model Graphs

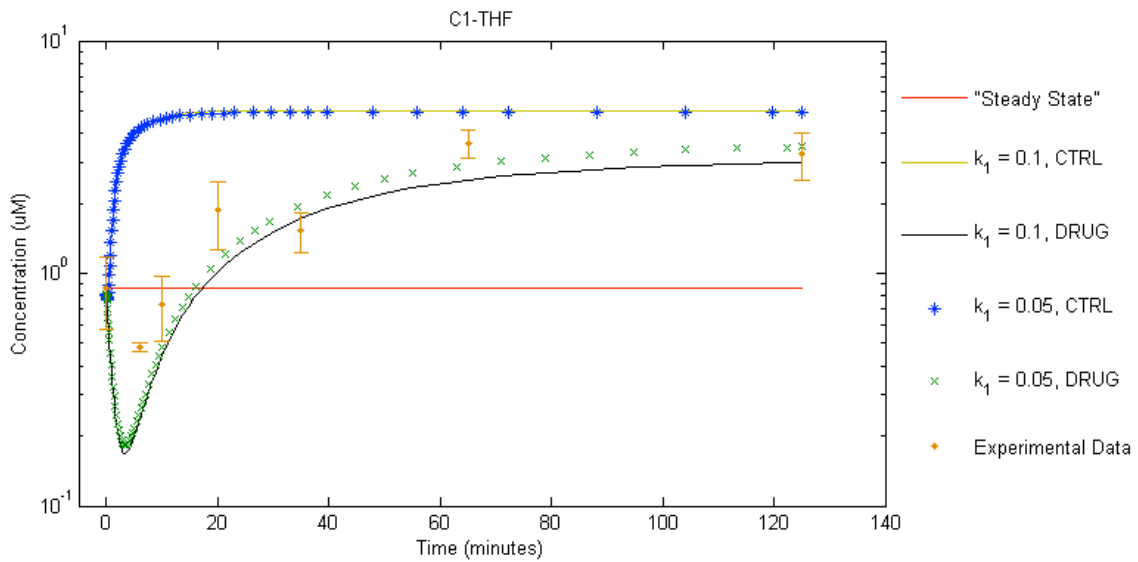


Figure 18: Kwon model accounting for the discrepancy in  $k_1$  in  $(C_1)$ -THF<sub>1</sub>.

By looking at the graph it is clear to see the differences between the model CTRL lines and the "Steady State" and the differences between the model DRUG lines and the experimental data. Changing  $k_1$  appears to make little difference.

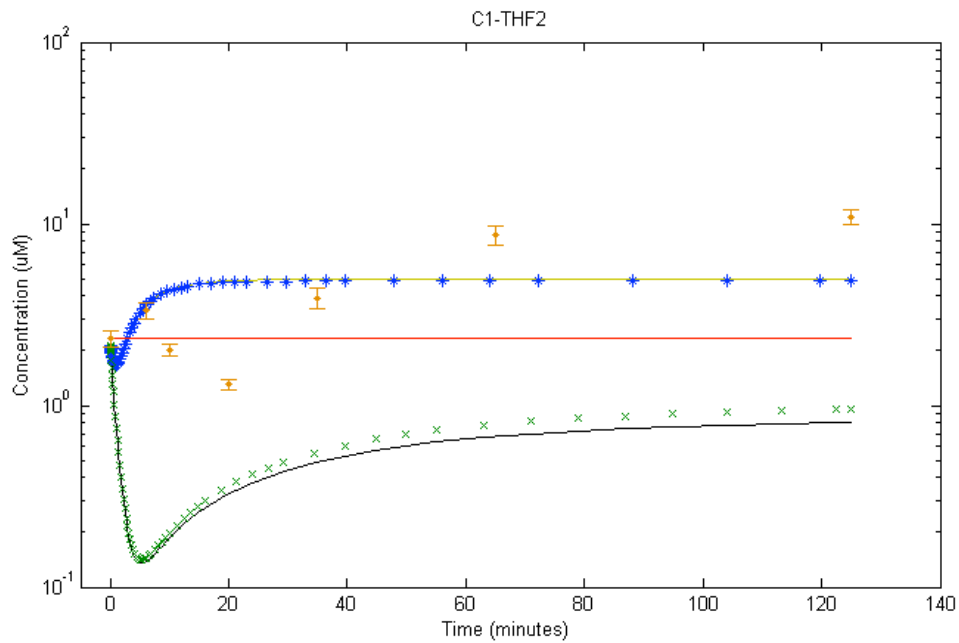


Figure 19: Kwon model accounting for the discrepancy in  $k_1$  in  $(C_1)$ -THF<sub>2</sub>.

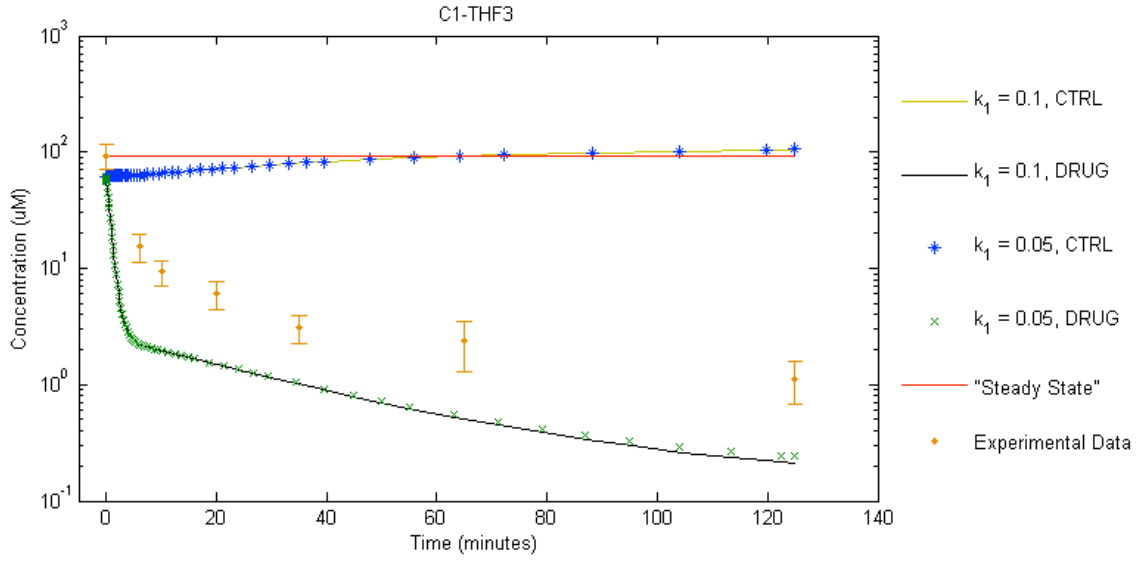


Figure 20: Kwon model accounting for the discrepancy in  $k_1$  in  $(C_1)$ -THF<sub>3</sub>.

## B.1 Revised Model

### B.1.1 Model Ordinary Differential Equations

ODEs:

$$d(\text{DHF\_M})/dt = 1/\text{unnamed}*(\text{ReactionFlux1} - \text{ReactionFlux4} - \text{ReactionFlux7} + \text{ReactionFlux8})$$

$$d([\text{5CH3THF\_M}])/dt = 1/\text{unnamed}*(-\text{ReactionFlux5} - \text{ReactionFlux9} + \text{ReactionFlux10})$$

$$d([\text{510mTHF\_M}])/dt = 1/\text{unnamed}*(-\text{ReactionFlux1} + \text{ReactionFlux2} + \text{ReactionFlux3} + \text{ReactionFlux6} + \text{ReactionFlux9} - \text{ReactionFlux10})$$

$$d(\text{THF\_M})/dt = 1/\text{unnamed}*(-\text{ReactionFlux2} - \text{ReactionFlux3} + \text{ReactionFlux4} + \text{ReactionFlux5} - \text{ReactionFlux6})$$

$$d([\text{Pte/pAB\_G\_1\_M}])/dt = 1/\text{unnamed}*(\text{ReactionFlux7})$$

Fluxes:

$$\text{ReactionFlux1} = (\text{TS\_Vmax}*\text{dUMP\_C}*\text{[510mTHF\_M]})/(\text{TS\_K1}*\text{TS\_K2}*((\text{dUMP\_C}/\text{TS\_K1})+(\text{[510mTHF\_M]}/\text{TS\_K2})+((\text{dUMP\_C}*\text{[510mTHF\_M]})/(\text{TS\_K1}*\text{TS\_K2}))+1))$$

$$\text{ReactionFlux2} = (\text{GCS\_Vmax}*\text{glycine\_C}*\text{THF\_M})/(\text{GCS\_K1}*\text{GCS\_K2}*((\text{glycine\_C}/\text{GCS\_K1})+(\text{THF\_M}/\text{GCS\_K2})+((\text{glycine\_C}*\text{THF\_M})/(\text{GCS\_K1}*\text{GCS\_K2}))+1))$$

$$\text{ReactionFlux3} = (\text{SHMT\_Vmax}*\text{serine\_C}*\text{THF\_M})/(\text{SHMT\_K1}*\text{SHMT\_K2}*((\text{serine\_C}/\text{SHMT\_K1})+(\text{THF\_M}/\text{SHMT\_K2})+((\text{serine\_C}*\text{THF\_M})/(\text{SHMT\_K1}*\text{SHMT\_K2}))+1))$$

$$\text{ReactionFlux4} = (\text{folM\_Vmax}*\text{DHF\_M}*\text{NADPH\_C})/(\text{folM\_K1}*\text{folM\_K2}*((\text{DHF\_M}/\text{folM\_K1})+(\text{NADPH\_C}/\text{folM\_K2})+((\text{DHF\_M}*\text{NADPH\_C})/(\text{folM\_K1}*\text{folM\_K2}))+1))$$

$$\text{ReactionFlux5} = (21*\text{HCY\_vvms\_M}*\text{[5CH3THF\_M]}*\text{vvms\_k12})/(20*\text{vvms\_k13}*\text{vvms\_k14}*((\text{HCY\_vvms\_M}/\text{vvms\_k14})+1)*((\text{[5CH3THF\_M]}/\text{vvms\_k13})+1))$$

$$\text{ReactionFlux6} = (\text{THF\_M}*\text{Formate\_C}*\text{vch2\_k1})-(\text{[510mTHF\_M]}*\text{vch2\_k2})$$

$$\text{ReactionFlux7} = \text{k5.K5}*\text{DHF\_M}$$

$$\text{ReactionFlux8} = \text{F1}$$

$$\text{ReactionFlux9} = \text{MTHFR\_2\_R.Kcat}*\text{[5CH3THF\_M]}/(\text{MTHFR\_2\_R.[5CH3THF\_M\_Km]}*(1+\text{[5CH3THF\_M]}/\text{MTHFR\_2\_R.[5CH3THF\_M\_Ki]})+\text{[5CH3THF\_M]})$$

$$\text{ReactionFlux10} = \text{MTHFR\_1\_R.Kcat}*\text{[510mTHF\_M]}/(\text{[510mTHF\_M\_Km]}*(1+\text{[5CH3THF\_M]}/\text{MTHFR\_1\_R.[5CH3THF\_M\_Ki]})+\text{[510mTHF\_M]})$$

## B.1.2 Enzyme Kinetic Equations

The fluxes from above written out here in a much more readable form.

$$TS: \frac{TS\_V \max \times dUMP \times 510mTHF}{TS\_K1 \times TS\_K2 \times \left( \frac{dUMP}{TS\_K1} + \frac{510mTHF}{TS\_K2} + \frac{dUMP \times 510mTHF}{TS\_K1 \times TS\_K2} + 1 \right)}$$

$$f\text{olM}: \frac{f\text{olM}\_V \max \times DHF \times NADPH}{f\text{olM}\_K1 \times f\text{olM}\_K2 \times \left( \frac{DHF}{f\text{olM}\_K1} + \frac{NADPH}{f\text{olM}\_K2} + \frac{DHF \times NADPH}{f\text{olM}\_K1 \times f\text{olM}\_K2} + 1 \right)}$$

$$SHMT: \frac{SHMT\_V \max \times serine \times THF}{SHMT\_K1 \times SHMT\_K2 \times \left( \frac{serine}{SHMT\_K1} + \frac{THF}{SHMT\_K2} + \frac{serine \times THF}{SHMT\_K1 \times SHMT\_K2} + 1 \right)}$$

$$vch2: (THF \times Formate \times vchs\_k1) - (510mTHF \times vch2\_k2)$$

$$GCS: \frac{GCS\_V \max \times glycine \times THF}{GCS\_K1 \times GCS\_K2 \times \left( \frac{glycine}{GCS\_K1} + \frac{THF}{GCS\_K2} + \frac{glycine \times THF}{GCS\_K1 \times GCS\_K2} + 1 \right)}$$

$$MTHFR\_1: \frac{MTHFR\_Kcat \times 510mTHF}{510mTHF\_Km \times \left( 1 + \frac{5CH_3THF}{5CH_3THF\_Ki} \right) + 510mTHF}$$

$$MTHFR\_2: \frac{MTHFR\_Kcat \times 5CH_3THF}{5CH_3THF\_Km \times \left( 1 + \frac{5CH_3THF}{5CH_3THF\_Ki} \right) + 5CH_3THF}$$

$$v\text{vms}: \frac{21 \times v\text{vms}\_V \max \times HCY\_v\text{vms} \times 5CH_3THF}{20 \times v\text{vms}\_K13 \times v\text{vms}\_K14 \times \left( \frac{HCY\_v\text{vms}}{v\text{vms}\_K14} + 1 \right) \times \left( \frac{5CH_3THF}{v\text{vms}\_K13} + 1 \right)}$$

$$F1: F1$$

$$k5: K5 \times DHF$$

## B.1.3 Kinetic Constants

Table 3: Kinetic Constants in the Revised Model

Parameter	Parameter Value
TS_K1[11]	6.3
TS_K2[12]	14

GCS_K1 <sup>*</sup> [4]	3400
GCS_K2 <sup>*</sup> [4]	50
SHMT_K1[13]	600
SHMT_K2[13, 14]	50
folM_K1[15]	0.5
folM_K2[15]	4
vvms_k13 <sup>o</sup> [4]	25
vvms_k14 <sup>o</sup> [4]	0.1
MTHFR1_Km[16]	85
MTHFR1_Ki[17]	61
MTHFR2_Km[16]	0.5
MTHFR2_Ki[17]	61
vch2_K1 <sup>*</sup> [4]	0.01
vch2_K2 <sup>*</sup> [4]	0

---

<sup>\*</sup> The data was fitted for GCS to experimental data because no information was available in BRENDA.

<sup>o</sup> vvms is a combination of two separate reactions involving both 2.1.1.13 and 2.1.1.14. Therefore, there will be no exact match to literature values

<sup>\*</sup> vch2 is a non-enzymatic reaction. Therefore, there is no literature information because it is a fitted parameter.

## References

1. Bailey, S.W. and J.E. Ayling, *The extremely slow and variable activity of dihydrofolate reductase in human liver and its implications for high folic acid intake*. Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15424-9.
2. Weinstein, S.J., et al., *Null association between prostate cancer and serum folate, vitamin B(6), vitamin B(12), and homocysteine*. Cancer Epidemiol Biomarkers Prev, 2003. **12**(11 Pt 1): p. 1271-2.
3. Kwon, Y.K., et al., *A domino effect in antifolate drug action in Escherichia coli*. Nat Chem Biol, 2008. **4**(10): p. 602-608.
4. Leduc, D., et al., *Flavin-dependent thymidylate synthase ThyX activity: implications for the folate cycle in bacteria*. J Bacteriol, 2007. **189**(23): p. 8537-45.
5. Bekaert, S., et al., *Folate biofortification in food plants*. Trends in plant science, 2008. **13**(1): p. 28-35.
6. Segel, I.H., *Biochemical Calculations 2nd Edition*. 1976, New York: John Wiley & Sons.
7. *{the Concentration of "Folic Acid"}*. {Nutrition Reviews}, 2009. **{46}**: p. {324--325}.
8. *{Methotrexate: Studies on cellular metabolism. III.Effect on the transplasma-membrane redox activity and on ferricyanide-induced proton extrusion by HeLa cells}*. {Cell Biochemistry and Function}, 1989. **{7}**: p. {135--137}.
9. *{Mechanism of the antimicrobial drug trimethoprim revisited}*. {Faseb Journal}, 2000. **{14}**: p. {2519--2524}.
10. Schomburg, I., et al., *BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA*. Nucleic Acids Research, 2013. **41**(D1): p. D764-D772.
11. Chiericatti, G. and D.V. Santi, *Aspartate 221 of thymidylate synthase is involved in folate cofactor binding and in catalysis*. Biochemistry, 1998. **37**(25): p. 9038-42.
12. Haertle, T., F. Wohlrab, and W. Guschlbauer, *Thymidylate synthetase from Escherichia coli K12. Purification, and dependence of kinetic properties on sugar conformation and size of the 2' substituent*. Eur J Biochem, 1979. **102**(1): p. 223-30.
13. Fu, T.F., et al., *Role of proline residues in the folding of serine hydroxymethyltransferase*. J Biol Chem, 2003. **278**(33): p. 31088-94.
14. Schirch, V., et al., *Serine hydroxymethyltransferase from Escherichia coli: purification and properties*. J Bacteriol, 1985. **163**(1): p. 1-7.
15. Wilquet, V., et al., *Dihydropteridine reductase as an alternative to dihydrofolate reductase for synthesis of tetrahydrofolate in Thermus thermophilus*. J Bacteriol, 2004. **186**(2): p. 351-5.
16. Trimmer, E.E., et al., *Folate activation and catalysis in methylenetetrahydrofolate reductase from Escherichia coli: roles for aspartate 120 and glutamate 28*. Biochemistry, 2001. **40**(21): p. 6216-26.
17. Trimmer, E.E., et al., *Aspartate 120 of Escherichia coli methylenetetrahydrofolate reductase: evidence for major roles in folate binding and catalysis and a minor role in flavin reactivity*. Biochemistry, 2005. **44**(18): p. 6809-22.