

A PARAMETER ESTIMATION APPROACH TO FORM A PRELIMINARY MODEL DESCRIBING
THE TRANSCRIPTOMIC AND PROTEOMIC DYNAMICS IN RAT LIVER FROM A DOSE OF
CORTICOSTEROID

By

JONATHAN PAI

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Chemical and Biochemical Engineering

Written under the direction of

Ioannis P. Androulakis

And approved by

New Brunswick, New Jersey

January, 2016

ABSTRACT OF THESIS

A parameter estimation approach to form a preliminary model describing the transcriptomic and proteomic dynamics in rat liver from a dose of corticosteroid

by JONATHAN PAI

Thesis Director:

Ioannis Androulakis

Synthetic glucocorticoids are used widely in medicine for its effective anti-inflammatory and immunosuppressive effect. However, they have fairly negative adverse effects especially when used in the long term. Better understanding of the pharmacokinetics and pharmacodynamics (PK/PD) will provide better insights to their mechanism of action helping us to optimize its use as a clinical therapy. The onset of gene microarray presents quantitatively large data sets that can be very cumbersome to analyze. A parameter estimation approach using optimization methods can be used in order to help shape a PK/PD model for describing the transcriptomic and proteomic changes from CS. In this study, parameter estimation was used to form a preliminary model of the transcriptomic and proteomic gene interaction in rat liver from a dose of methylprednisolone (MPL), a commonly used synthetic glucocorticoid.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance of several individuals who contributed and helped in the preparation and completion of this study.

I would like extend my gratitude to my advisor Dr. Ioannis Androulakis for his patience and guidance in completing this study.

Kubra Kamisoglu for her help in handling the bioinformatics and more biology related questions I had about the project.

Pantelis Mavroudis for his help outlining the parameter estimation approach and MATLAB assistance.

Dr. Roth for making sure my academic requirements were on track.

Lynn DeCaprio and Debora Moon for handling the administrative issues in the Chemical and Biochemical engineering department to ensure the smooth operation of the school.

All of my classmates in the Chemical and Biochemical engineering departments who have supported and assisted me in the course of my study.

Finally, my family who have supported me and encouraged me to work hard during my time in the program.

Table of Contents

Abstract of Thesis.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables	v
List of Figures	vi
I. Introduction	1
II. Methods.....	3
Pharmacokinetic and pharmacodynamic model of methylprednisolone	3
Selecting a relatively simple subnetwork of genes for modeling	6
Modeling of the Genes	7
Parameter estimation model	9
Experimental Data	12
System Specs.....	12
III. Results.....	13
IV. Discussion	16
V. Conclusion.....	18
VI. Bibliography	19

List of Tables

Table 1. Pharmacokinetic and receptor dynamic parameters (Jin, Almon, Dubois, & Jusko, 2003).	4
Table 2. Estimated Parameter Values from model. Case A is evenly weighted ($w_i=1$) while Case B has a weight $w_{GPX1}=10$ for GPX1.....	14

List of Figures

Figure 1. Fifth-generation model of CS pharmacokinetics and receptor dynamics from Jin et al. Where R_m is $mRNAR$, $kdgrR$ is $kdRM$, $ksynR$ is ksR , $kdgrR$ is kdR (Jin, Almon, Dubois, & Jusko, 2003).	5
Figure 2. Plasma Concentration of MPL after 50 ml/mg dose solved using Jin et al.'s model.	5
Figure 3. Profiles of the GR mRNA (R_m), cytosolic GR density (R), cytosolic DR (DR), DR in nucleus ($DR(N)$) solved using Jin et al.'s model.	6
Figure 4. Graphviz representation of the interaction links between the six genes in the simplified network. Arrows represent activation and bars represent inhibition.	7
Figure 5. Schematic of the parameter estimation program in MATLAB. *.m files were files written by me and <code>fmincon()</code> and <code>ode45()</code> were built-in MATLAB functions.	10
Figure 6 Even weights for each gene and upper and lower bounds of 5 and -5 respectively for the parameter values. Solid black line is the solution of the system of ODEs and the blue circles represent each time point of the experimental data.	13
Figure 7. Weight of $w=10$ for GPX1 Gene and Protein only and $w=1$ for everything else. Upper and lower bounds of 5 and -5 respectively for the parameter values. Solid black line is the solution of the system of ODEs and the blue circles represent each data point.	14

I. Introduction

Synthetic glucocorticoids are part of a group of drugs known as corticosteroids (CS) and have an extremely potent anti-inflammatory/immunosuppressive effect (Jin, Almon, Dubois, & Jusko, 2003). Their ability to suppress inflammation and immune response make it a very effective therapy and commonly used to treat conditions such as organ transplantation, rheumatoid arthritis, inflammatory bowel disease, Crohn's disease, lupus, multiple sclerosis, dermatomyositis, arthritis, leukemia, non-Hodgkin's lymphoma, asthma, and glomerulonephritis (Almon, Dubois, Pearson, Stephan, & Jusko, 2003). However, long term use can bring about adverse effects such as electrolyte disturbances, cardiovascular effects, diabetes mellitus and loss of bone density and osteoporosis with concomitant vertebral fracture (Frauman, 1996).

Both the desirable and adverse effects of CS are caused by the binding of CS to the glucocorticoid receptor (GR). Virtually all tissue contain genes that are subject to regulation by activated GR resulting in enhanced or repressed expression of mRNA (Almon, Dubois, Pearson, Stephan, & Jusko, 2003). The liver in particular is one of the primary targets of CS action and plays a central role in maintaining systemic energy balance (Kamisoglu, et al., 2015). Alteration of a critical transcription factor can affect expression of many genes which mean a better understanding of CS pharmacogenomics and their mechanism of action will allow us to improve its use as a clinical therapy (Jin, Almon, Dubois, & Jusko, 2003). This stresses the importance of developing a mechanistic pharmacokinetics/pharmacodynamics (PK/PD) for quantitative understand of molecular and cellular mechanisms (Jin, Almon, Dubois, & Jusko, 2003).

Traditionally message quantification methods such as Northern blot and reverse transcriptase-polymerase chain reaction only measured single genes which limited the number of genomic changes that could be observed preventing efficacy and toxicity to be adequately

studied (Jin, Almon, Dubois, & Jusko, 2003). The onset of high throughput techniques such as gene microarrays enabled the collection of very large sets of quantitative data on biological substances in living cells. Cellular dynamic systems usually contain too many parameters and complex dynamics so a systematic way of inferring biological regulatory networks is required (Han, Yoon, & Cho, 2007). An optimization framework is able to be used for the design and analysis of regulatory networks by combining gene expression data and prior biological knowledge of regulatory interactions between genes and corresponding transcription factors (Foteinou, Yang, Saharidis, Ierapetritou, & Androulakis, 2009). In this study, a parameter estimation model is developed using MATLAB in order to provide an adaptable and robust method of constructing a PK/PD model of the transcriptomics and proteomics in rat liver after a dose of methylprednisolone (MPL) widely used as a corticosteroid.

II. Methods

Pharmacokinetic and pharmacodynamic model of methylprednisolone

The pharmacokinetics of methylprednisolone (MPL) can be described by the biexponential equation by Jin et al. (Jin, Almon, Dubois, & Jusko, 2003).

$$D = C_1 \cdot e_1^{-\lambda_1 t} + C_2 \cdot e^{-\lambda_2 t}$$

Where D is the plasma concentration of MPL in nanograms per milliliter and C_i and λ_i are the coefficients for the y-intercepts and slopes respectively of the semi-logarithmic plot shown in Figure 2.

Jin *et al.* also developed a pharmacodynamic model of receptor dynamics in rat liver after MPL treatment (Jin, Almon, Dubois, & Jusko, 2003). This model is depicted in Figure 1 and is described by the following differential equations.

$$\frac{dR_m}{dt} = k_{syn_{RM}} \cdot \left(1 - \frac{DR(N)}{IC_{50_{RM}} + DR(N)} \right) - k_{dgr_{RM}} \cdot R_m$$

$$\frac{dR}{dt} = k_{syn_R} \cdot R_m + R_f \cdot k_{re} DR(N) - k_{on} D \cdot R - k_{dgr_R} \cdot R$$

$$\frac{dDR}{dt} = k_{on} D \cdot R - k_T \cdot DR$$

$$\frac{dDR(N)}{dt} = k_T \cdot DR - k_{re} \cdot DR(N)$$

Where the terms are the plasma molar concentration of MPL (D), the receptor mRNA (R_m), the free cytosolic GR density (R), cytosolic drug-receptor complex (DR), and drug-receptor complex in nucleus [$DR(N)$]. Zero-order rate of GR mRNA synthesis ($k_{syn_{RM}}$). First-order rates of GR mRNA degradation ($k_{dgr_{RM}}$), receptor synthesis (k_{syn_R}) and degradation (k_{dgr_R}), translocation of the drug-receptor complex into the nucleus (k_T), and overall turnover of $DR(N)$ to cytosol (k_{re}). Second-order rate constant of drug-receptor association (k_{on}). $IC_{50_{RM}}$ is the

concentration of DR(N) at which the synthesis rate of receptor mRNA drops to 50% baseline and R_f is the fraction of free receptor being recycled.

Baselines defined using

$$k_{dgr_{RM}} = \frac{k_{syn_{RM}}}{R_{m_0}}$$

$$k_{syn_R} = \left(\frac{R^0}{R_{m_0}} \right) \cdot k_{dgr_R}$$

Where R_{m_0} and R^0 are baseline values of receptor mRNA and free cytosolic GR density.

Table 1. Pharmacokinetic and receptor dynamic parameters (Jin, Almon, Dubois, & Jusko, 2003).

Parameter	Value
Pharmacokinetics (fixed)	
C_1 (ng/ml)	39130
C_2 (ng/ml)	12670
λ_1 (h ⁻¹)	7.54
λ_2 (h ⁻¹)	1.20
Receptor dynamics (fixed)	
$k_{s_{RM}}$ (fmol/g liver/h)	2.90
$IC_{50_{RM}}$ (fmol/mg of protein)	26.2
k_{on} (l/nmol/h)	0.00329
k_T (h ⁻¹)	0.63
k_{re} (h ⁻¹)	0.57
R_f	0.49
k_{d_R} (h ⁻¹)	0.0572
R^0 (fmol/mg protein)	540.7

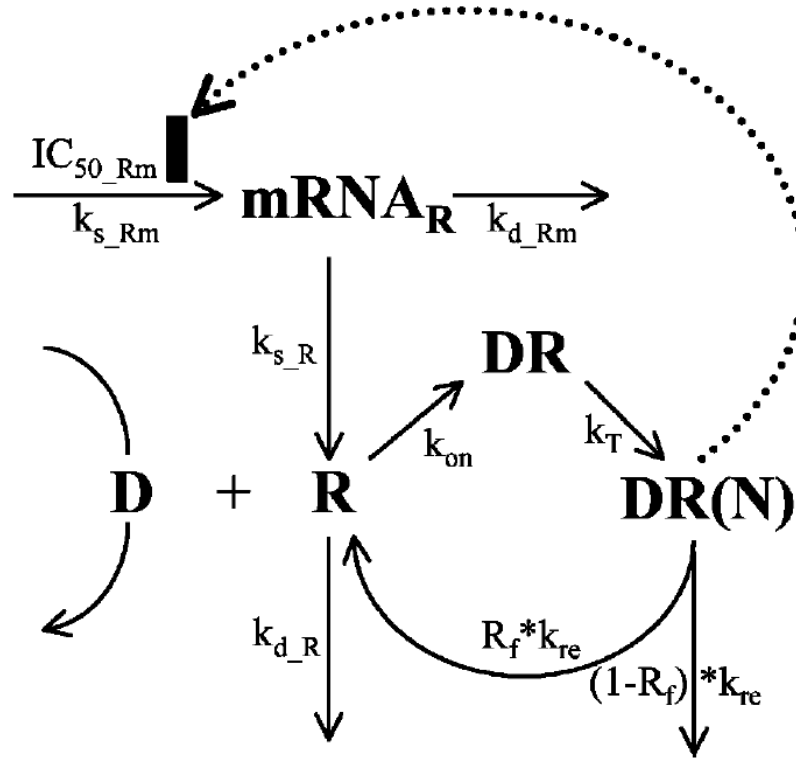


Figure 1. Fifth-generation model of CS pharmacokinetics and receptor dynamics from Jin et al. Where R_m is $mRNA_R$, k_{dgr_R} is $k_{d_{RM}}$, k_{syn_R} is k_{s_R} , k_{dgr_R} is k_{d_R} (Jin, Almon, Dubois, & Jusko, 2003).

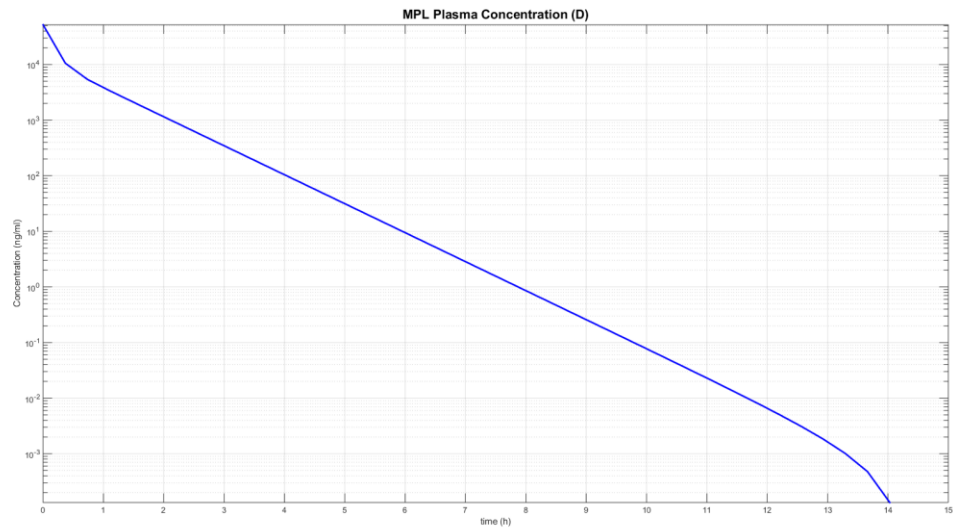


Figure 2. Plasma Concentration of MPL after 50 ml/mg dose solved using Jin et al.'s model.

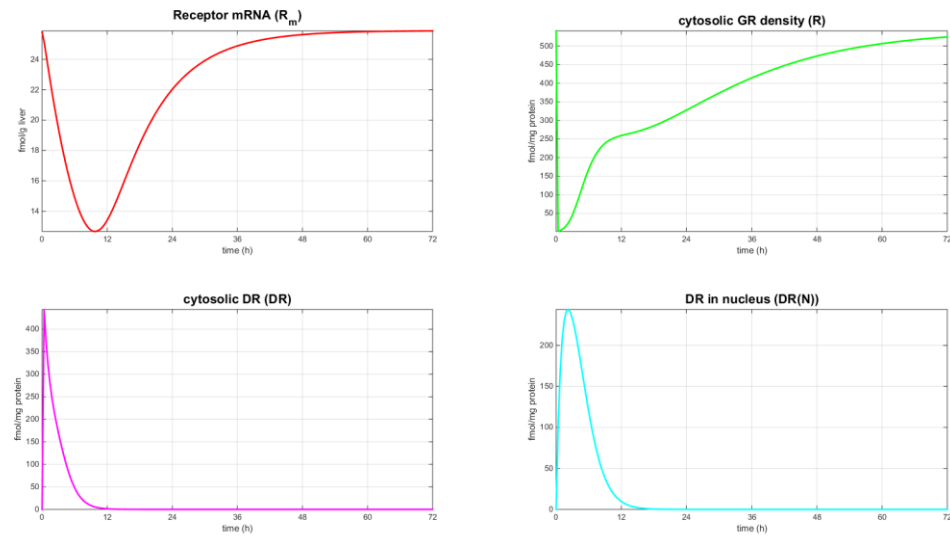


Figure 3. Profiles of the GR mRNA (R_m), cytosolic GR density (R), cytosolic DR (DR), DR in nucleus ($DR(N)$) solved using Jin et al.'s model.

Selecting a relatively simple subnetwork of genes for modeling

An interaction network of 163 genes was formed in a study by Tung (Nguyen, T, unpublished work) which drew from four popular pathway databases (KEGG, NCI, Biocarta, Reactome) as well as text mining from Pubmed abstracts for whether gene A activates or inhibits gene B. However, the network involved 1956 interaction links, which was too large for us to reasonably work with for this study. Kubra Kamisoglu simplified the network, using a functional approach, to a system of six genes and only 16 interaction links. This was done by determining the most important functions of MPL administration by functional enrichment analysis, focusing on the most important common pathways, and eliminating genes where data was missing for mRNA or protein because we needed to be able to compare both. The six genes chosen for this study were GSTK1, GSTM1, GPX1, PRDX6, ACACA, and MYL9 and the interaction matrix is shown in Figure 4.

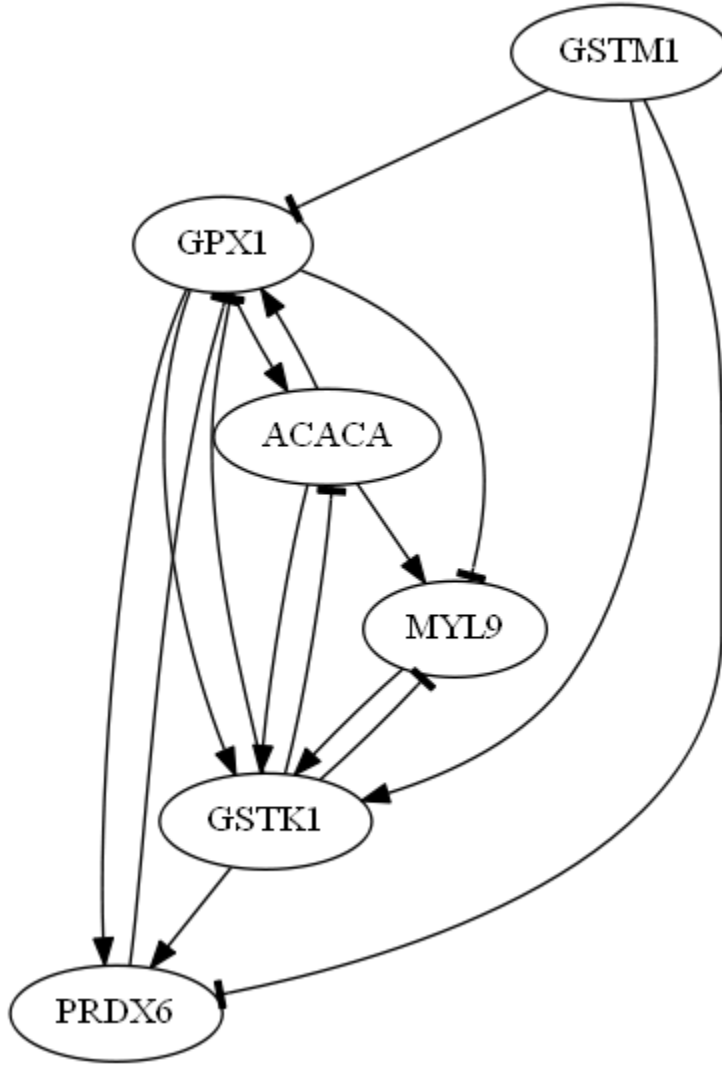


Figure 4. Graphviz representation of the interaction links between the six genes in the simplified network. Arrows represent activation and bars represent inhibition.

Modeling of the Genes

The rate of production of mRNA can be modeled using simple synthesis and degradation terms (Foteinou et al. 2007). The rate of production of protein was modeled in a similar way. In order to integrate the transcriptomic and proteomic models, the generation terms of each depend on the other.

$$\frac{dG_i}{dt} = f_i - k_{dgr_m} \cdot G_i$$

$$\frac{dP_i}{dt} = k_{syn_p} \cdot G_i - k_{dgr_p} \cdot P_i$$

Where G_i is the concentration of transcript and P_i is the concentration of protein. The f_i term was created as follows and is based on a proposed network interaction between the six genes shown in Figure 4 of the form

$$f_i = k_{syn_m} \sum_j^n a_{ij} \cdot P_j$$

Where a_{ij} is the interaction matrix of protein j regulating gene i with a value of $a = 1$ for each instance of activation, $a = -1$ for each instance of inhibition, and $a = 0$ for no regulation.

Interaction matrix a_{ij} is visually shown in Figure 4.

$$f_{GSTK1} = k_{GSTK1}DR(N) + [k_{1,1}P_{GSTM1} + k_{1,2}P_{GPX1} + k_{1,3}P_{ACACA} + k_{1,4}P_{MYL9}]$$

$$f_{GSTM1} = k_{GSTM1}[1 + DR(N)]$$

$$f_{GPX1} = k_{GPX1}DR(N) + [k_{3,1}P_{GSTK1} - k_{3,2}P_{GSTM1} - k_{3,3}P_{PRDX6} + k_{3,4}P_{ACACA}]$$

$$f_{PRDX6} = k_{PRDX6}DR(N) + [k_{4,1}P_{GSTK1} - k_{4,2}P_{GSTM1} + k_{4,3}P_{GPX1}]$$

$$f_{ACACA} = k_{ACACA}DR(N) + [-k_{5,1}P_{GSTK1} + k_{5,2}P_{GPX1}]$$

$$f_{MYL9} = k_{MYL9}DR(N) + [-k_{6,1}P_{GSTK1} - k_{6,2}P_{GPX1} + k_{6,3}P_{ACACA}]$$

Where k_{syn_m} is the rate of generation of transcript, k_{dgr_m} is the rate of degradation of transcript, k_{syn_p} is the rate of generation of protein, and k_{dgr_p} is the rate of degradation of protein. In the expanded f equations, the k_{syn_m} parameter is propagated to each term of the interaction matrix. This allows different genes to unequally affect the activation and inhibition of mRNA generation. These parameters affecting the generation and degradation of transcript and proteins are the unknown parameters we try to solve for in this problem. By being able to fit our model to the data, the parameters can tell us how well our model can fit the data and numerically what affects the generation and degradation more heavily such as which genes and

proteins may be more involved in activating or inhibiting production and can give us better insights to the mechanism of action.

This model was created under the assumptions

- 1) The gene-protein network interaction matrix a_{ij} is known.
- 2) Protein concentration regulates transcription only.
- 3) Protein act independently of each other.

Parameter estimation model

A parameter estimation model was developed in MATLAB in order to provide an easily adaptable and robust method of forming PK/PD models of the transcriptomics and proteomics in rat liver. The core of the program uses an optimization solver to minimize an objective function comparing the solution for a system of ordinary differential equations (ODEs) and the experimental data. This method involves an iterative approach of proposing a system of ODEs and solving it in the parameter estimation model to determine the fit and revising the model in order to try to make it fit better either quantitatively through objective function value or qualitatively by visually assessing the graph of the solution.

The methods used in this study were all of the simplest case in order to allow room for expansion in future studies because of the modular nature of the model. As such, the program was split into multiple parts as shown in Figure 5.

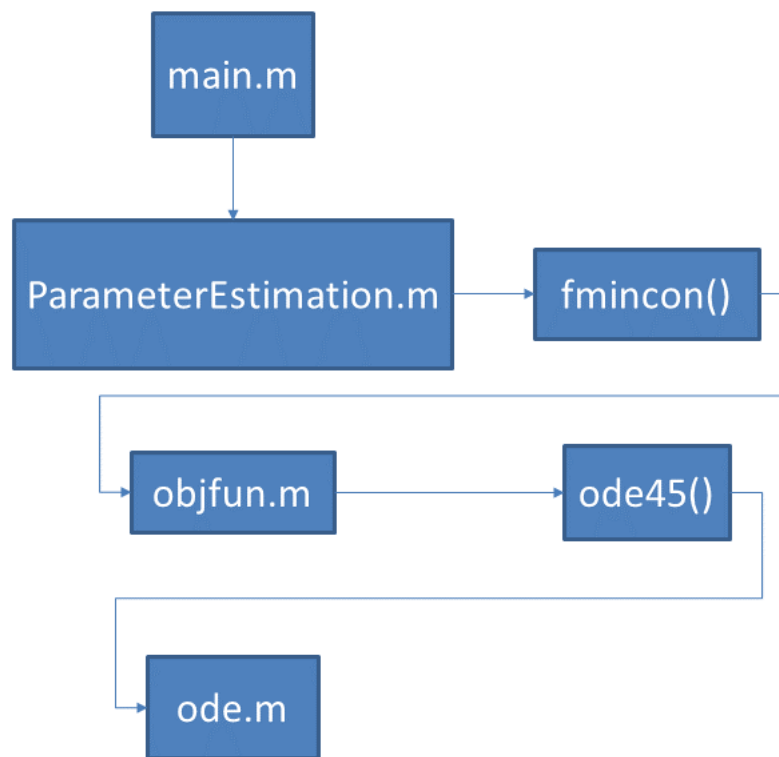


Figure 5. Schematic of the parameter estimation program in MATLAB. *.m files were files written by me and `fmincon()` and `ode45()` were built-in MATLAB functions.

The main overall package used to run everything was `main.m`. A helper function not shown in the schematic was used to read in the experimental data from an excel spreadsheet and store it as a cell array in order to address the issue of the transcriptomic having more time points than the proteomic data. `ParameterEstimation.m` was the function for parameter estimation calculation and outputs minimized parameter function values. This function took not only initial parameter values as inputs and their upper and lower bounds but also experimental data. The reason for using experimental data as one of the inputs was to later allow bootstrapping calculation to be done by varying the experimental data using a randomly determined value for each time point. However, due to time constraints bootstrapping was not done in this study.

The ParameterEstimation.m function called MATLAB's built-in fmincon() function for constrained nonlinear multivariable optimization and used the default interior-point algorithm capable of handling large, sparse, and small dense problems. Because our system of ODEs has many unknown parameters, many different initial values for the optimization need to be used. For this purpose, fmincon() was run using MATLAB's MultiStart. MultiStart generates a set of multiple starting points for fmincon(), runs them individually, and outputs the final minimum objective function value and solution. Parallelization was used in order to accelerate the calculation.

The objective function used for fmincon() was placed in objfn.m in order to allow easy modification. A simple difference of least squares was used as the objective function. Of the form

$$err = \sum_{i=1}^n (y_i - yexp_i)^2 \cdot w_i$$

Where *err* is the error between the model solution and the experimental data, *i* is the index of the gene, *n* is the total number of genes compared, *y* is the solution to the system of ODEs for each gene and *yexp* is the experimental data of the transcriptomic and proteomic data. A weight of *w_i* was included as a way of assigning weight values to certain genes to make them more important in the objective function calculation. A weight value of *w* = 1 can be normally set for all genes in order to treat them equally but if the fit is desired for a specific gene then a higher value can be input for that specific index. In this study a weight value of *w_{GPX1}* = 10 was set for GPX1 transcript and protein in order to compare how well the model could fit a gene and see how the others fit under those constraints.

The objective function called MATLAB's built-in ordinary differential equation solver, ode45() function, which Mathworks recommends is the first algorithm to try because it works

most of the time. In order to call ode45() the system of ODEs were written in ode.m in a way that makes editing and adjusting the ODEs in a simple manner. This is to facilitate the tweaking of the PK/PD model and running the parameter estimation calculation.

In order to visualize the results, a helper graphing method was written in order to view the solutions to the system of ODEs with the experimental data for each gene side by side.

Experimental Data

The parameter estimation was done using transcriptomic and proteomic data from two previous studies using liver samples excised from rats. The transcriptomic data was collected from 44 male adrenalectomized (ADX) Wistar rats that were given an intravenous bolus dose of 50 mg/kg MPL and four untreated rats serving as controls. Control group rats were sacrificed at a time point of 0 h and the dosed rats sacrificed at 0.25, 0.5, 0.75, 1, 2, 4, 5, 5.5, 6, 7, 8, 12, 18, 30, 48, 72 h. The liver samples were then analyzed using Affymetrix GeneChips Rat Genome U34A (Affymetrix, Inc.) (Jin, Almon, Dubois, & Jusko, 2003). Proteomic data was collected from 60 ADX Wistar rats given an intramuscular 50 mg/kg MPL dose and sacrificed at time points 0.5, 1, 2, 4, 5.5, 8, 12, 18, 30, 48, 66 h and analyzed using nano-LC/LTQ/Orbitrap instrument (Nouri-Nigjeh, et al., 2014).

System Specs

All computations done using MATLAB version 2014a on an Intel® Core™2 Quad CPU Q6600 @ 2.40GHz 4.00 GB 32-bit Operating System running Windows Vista™ Home Premium SP2.

III. Results

Running the model through the parameter estimation program produced positive results. The final model as shown in the previous section was able to fit the experimental data fairly well.

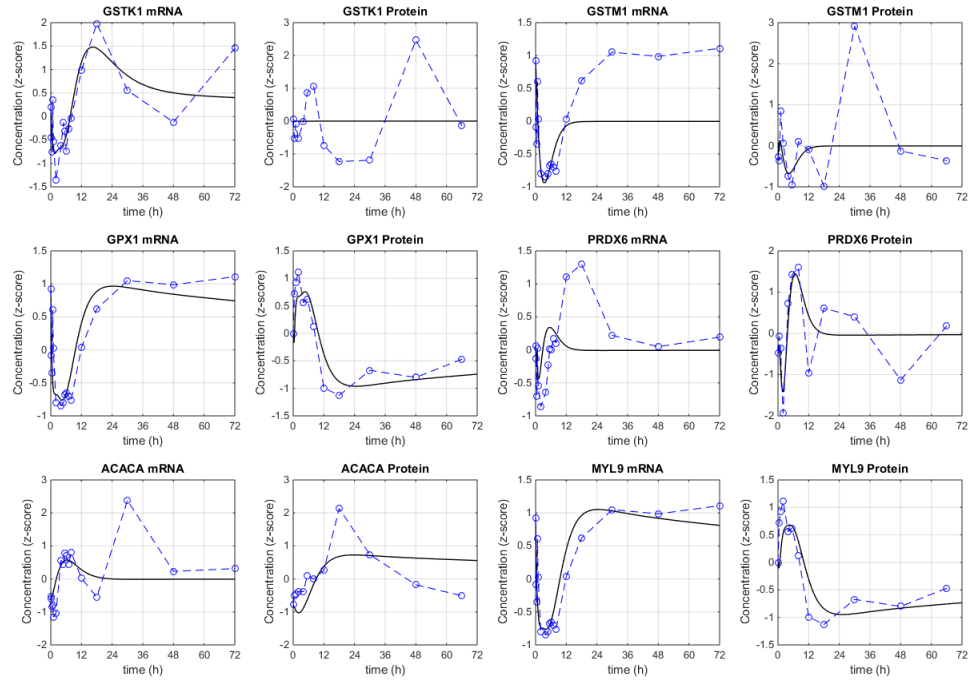


Figure 6 Even weights for each gene and upper and lower bounds of 5 and -5 respectively for the parameter values.

Solid black line is the solution of the system of ODEs and the blue circles represent each time point of the experimental data.

The MultiStart of the parameter estimation calculation was done using 400 randomly generated starting points of parameter values ranging from -5 to 5. As can be seen in Figure 6 the model solution fits the experimental data fairly well in almost every case.

A weight of $w_{GPX1} = 10$ was applied for the GPX1 gene and protein term in the objective function in order to observe the fit. This calculation was able to fit the data in an almost identical way to the previously shown calculation using all even weights but was able to

find it using only 200 starting points for MultiStart and half the computational time.

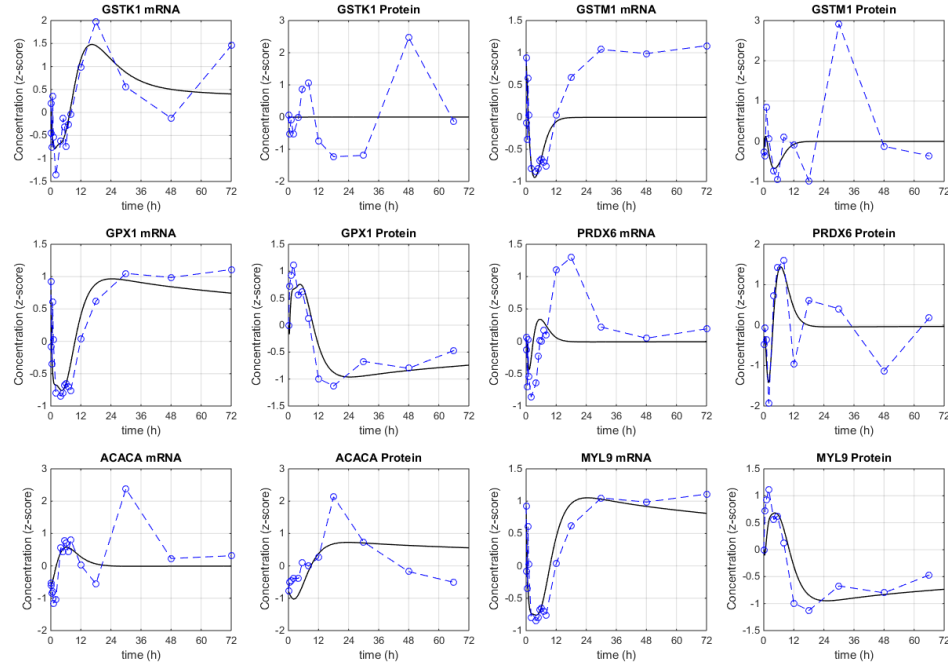


Figure 7. Weight of $w=10$ for GPX1 Gene and Protein only and $w=1$ for everything else. Upper and lower bounds of 5 and -5 respectively for the parameter values. Solid black line is the solution of the system of ODEs and the blue circles represent each data point.

Table 2. Estimated Parameter Values from model. Case A is evenly weighted ($w_i=1$) while Case B has a weight $w_{GPX1}=10$ for GPX1.

Case		A	B
Objective fn value		60.03	98.8
time (s)		208518	106278.9
GSTK1	k_{GSTK1} k1	0.00207	0.005433
	$k_{1,1}$ k25	0.58645	0.595412
	$k_{1,2}$ k26	0.01671	0.054527
	$k_{1,3}$ k27	1.55903	1.812849
	$k_{1,4}$ k28	0.67382	1.267736
	k_{dgr_m} k2	0.10154	0.101169
	k_{syn_p} k3	0.00098	0.004384
GSTM1	k_{dgr_p} k4	1.26385	4.656541
	k5	-0.00498	-0.00497
	k_{dgr_m} k6	1.17497	1.168711
	k_{syn_p} k7	1.17979	1.103182

GPX1	k _{dgr_p}	k8	1.57105	1.453446
	k _{GPX1}	k9	-0.0013	0.004443
	k _{3,1}	k29	4.72665	4.227765
	k _{3,2}	k30	0.41152	0.011318
	k _{3,3}	k31	0.32333	0.230179
	k _{3,4}	k32	1.61751	2.735257
PRDX6	k _{dgr_m}	k10	0.80213	2.061918
	k _{syn_p}	k11	-4.58607	-4.92862
	k _{dgr_p}	k12	4.86813	4.948451
	k _{PRDX6}	k13	-0.00736	-0.00874
	k _{4,1}	k33	4.08456	2.367895
	k _{4,2}	k34	3.42968	4.360083
ACACA	k _{4,3}	k35	0.01649	0.060699
	k _{dgr_m}	k14	2.76113	4.410379
	k _{syn_p}	k15	4.9834	4.877272
	k _{dgr_p}	k16	1.27971	1.052419
	k _{ACACA}	k17	0.0011	0.001227
	k _{5,1}	k36	4.87983	3.774845
MYL9	k _{5,2}	k37	0.00048	0.00023
	k _{dgr_m}	k18	0.00139	0.231157
	k _{syn_p}	k19	0.21351	0.334162
	k _{dgr_p}	k20	0.30481	0.000648
	k _{MYL9}	k21	0.0012	0.005478
	k _{6,1}	k38	4.47587	1.910702
	k _{6,2}	k39	3.68544	3.087072
	k _{6,3}	k40	3.12127	2.642957
	k _{dgr_m}	k22	4.94284	4.644076
	k _{syn_p}	k23	-4.72504	-1.4601
	k _{dgr_p}	k24	4.73273	1.619557

IV. Discussion

Looking at Figure 6 shows us that the model was able to capture the dynamics of the transcriptomic and proteomic changes fairly well. The biggest areas of discrepancy occur for GSTM1 and PRDX6 transcript data. In the case of GSTM1, while the experimental data begins near one, decreases, and then increases back to 1, our model ends on 0 and seems incapable of reaching 1. This shows a main error with our model and may suggest additional terms added to the equation describing the rate of production of mRNA. Physically, the experimental data makes sense because after the bolus dose of drug, the body will purge the drug and attempt to recover its initial state. This needs to be reflected in our model by possibly adding additional terms to allow as time approaches infinity, the model should approach the initial conditions.

Another point of concern is that in order to better fit that data, the bounds of the parameter values were expanded to allow negative values. While negative parameter values don't have physical meaning, the parameters that were calculated to have negative values show signs of where our model needs to be improved. Most noticeably is the parameter for protein synthesis (k_{syn_p}) for GPX1 and MYL9 as can be seen in Table 2. This may suggest that instead of mRNA having a positive effect on the protein level it could have a negative effect. The $k_{syn_p} mRNA$ term in the protein equation may need to be expanded into a term similar to the transcript equation involving some function of the level of mRNA, $f'(mRNA)$, that may be also affected by a similar interaction matrix, a'_{ij} .

Other reasons for our model to not fit the experimental data well could be the way we simplified the problem. The model we used was linear, which was a reasonable starting assumption to build our model and allow easier solving, however the underlying dynamics could be highly nonlinear. Also, the interaction matrix used was an extremely small subset in order to simplify our model. More work should be done expanding the subnetwork chosen to include a

few more genes or choosing a larger subnetwork. Eventually the goal would be to add all 163 genes in the original data to our model.

Comparing our evenly weighted objective function with the objective function weighting GPX1 ten times higher than the other genes as in Figure 7 show very similar results. While in this study weighting was mainly used to help the optimization solver to converge on a solution faster by focusing on a specific gene, weighting should be more correctly used by basing it off of the standard deviation of the experimental data as a confidence factor. The more confident experimental data points should have a larger impact on the model and allow the less confident data to have more room for error.

V. Conclusion

The results show how parameter estimation can be used to determine the fit our PK/PD model has to the experimental data. The objective function value can numerically evaluate how close our model is to the data and the parameters can tell which parts of the model more heavily affect the rate of production of mRNA or protein. By observing the fit visually and quantitatively the model can be tweaked in an iterative fashion in order to improve the model. From the results looking at the transcriptomic and proteomic data from rat liver after a dose of MPL our final model for the transcriptomic and proteomic dynamics fit the experimental data fairly well. Certain parts of the model are fairly lacking in describing the dynamics more specifically, as large times, our model approaches zero instead of the initial value. More terms need to be added in order to more accurately represent the mechanism.

VI. Bibliography

- Almon, R. R., Dubois, D. C., Pearson, K. E., Stephan, D. A., & Jusko, W. J. (2003). Gene arrays and temporal patterns of drug response: corticosteroid effects on rat liver. *Functional & integrative genomics*, 171-9.
- Foteinou, P. T., Yang, E., Saharidis, G. K., Ierapetritou, M. G., & Androulakis, I. P. (2009). A mixed-integer optimization framework for the synthesis and analysis of regulatory networks. *Journal of Global Optimization*, 43(2), 263-276.
- Frauman, A. (1996). An overview of the adverse reactions to adrenal corticosteroids. *Adverse Drug Reactions and Toxicological Reviews*, 15(4), 203-206.
- Han, S., Yoon, Y., & Cho, K.-H. (2007). Inferring biomolecular interaction networks based on convex optimization. *Computational Biology and Chemistry*, 31, 347-354.
- Jin, J. Y., Almon, R. R., Dubois, D. C., & Jusko, W. J. (2003). Modeling of Corticosteroid Pharmacogenomics in Rat Liver Using Gene Microarrays. *The Journal of Pharmacology and Experimental Therapeutics*, 307, 93-109.
- Kamisoglu, K., Sukumaran, S., Nouri-Nigjeh, E., Tu, C., Li, J., Shen, X., . . . Androulakis, I. P. (2015). Tandem Analysis of Transcriptome and Proteome Changes after a Single Dose of Corticosteroid: A Systems Approach to Liver Function in Pharmacogenomics. *A Journal of Integrative Biology*, 19(2), 80-91.
- Nouri-Nigjeh, E., Sukumaran, S., Tu, C., Li, J., Shen, X., Duan, X., . . . Qu, J. (2014). Highly multiplexed and reproducible ion-current-based strategy for large-scale quantitative proteomics and the application to protein expression dynamics induced by methylprednisolone in 60 rats. *Analytical Chemistry*, 86, 8149-8157.