

Computational Modeling of the AKT Pathway

Koh Yeow Nam, Geoffrey Teong Huey Fern David Hsu
Marie-Veronique Clement P.S. Thiagarajan

Abstract

Computational modeling and analysis of biological networks are a good way to gain understanding of cellular functions at the molecular level. Common techniques include differential equations to model molecular kinetics. More recently, models arising from engineering have been finding their way in biology, such as Finite State Machines, Boolean Networks, Petri Nets and their many variants. However, quantitative modeling faces some basic problems, one of them being lack of information concerning parameters. Here, we use a variant of Petri Net to model the AKT pathway and its interaction with the ERK cascade. We then look on issues relating to the parameter estimation problem and present our current methods to deal with them.

1 Introduction

Computational modeling has been gaining acceptance in the study and understanding of biological networks. The methods and tools of modeling vary but the process and problems faced by the researchers remain largely the same, such as the desired levels of abstraction, the underlying kinetic models and the lack of knowledge on the parameters. One of the more common methods is to represent the chemical reactions as a network and use ordinary differential equations to drive their dynamics. Several such pathways have been modeled and studied, each concentrating on a particular aspect of cellular activity, such as the canonical *Wnt* signaling pathway, the *mitogen-activated protein kinase (MAPK)* cascade, and *Caspase* action in apoptosis [14, 27, 46]. In the recent years, models of computation that were originally restricted to engineering domains are finding themselves being employed to model biological entities [12, 15, 30, 32]. One of the first such model is by McAdams and Shapiro, who represented regulatory networks as

electrical circuits [31]. Since then, other models such as Finite State Machines and Petri Nets have been used, all with the aim of finding out how a cell works from a systems point of view.

In this work, we employ a variant of the Petri Net methodology to model two pathways that are involved in programmed cell death, or apoptosis - the *AKT* and *ERK* pathways. We are interested not only in the pathways themselves, but also in understanding how they interact with each other via cross-interaction, as well as their influence on common downstream targets. We present a systematic way to translate information obtained from laboratory experiments and various literature sources into our model. In addition, we use an algorithm that is based on Evolutionary Strategies [3] to fill up the gaps in the model - namely parameter estimation. Thereafter, we look into some issues concerning the validation of our model.

2 Models of Computation

A model is a representation of a physical entity. In this case, it would be the molecular functions of a cell. It can range from being purely visual (symbols, diagrams etc.) to having formal semantics such that it is executable. The use of ordinary differential equations has long been the popular choice for modeling the molecular dynamics of the cell. In this work, we use a variant of the Petri Net methodology - the *Hybrid Functional Petri Net* [30], to model our pathway. The reason why we use the Hybrid Functional Petri Net is that it has a sound mathematical basis for all its components [37]. Also, there is already a simulator, the Cell Illustrator¹, that uses this methodology for modeling and simulation, allowing us to concentrate only on the pathways and not the execution semantics.

2.1 Hybrid Functional Petri Nets

The Hybrid Functional Petri Net (HFPN) is a modified version of the Petri Net model conceived by Carl Adam Petri in the 1960s, which has been used extensively to model concurrent processes. The original Petri Net consists of just two components - Places and Transitions. Places, usually denoted by a circular symbol, denotes passive entities such as buffers, or states of a system. Transitions, depicted by a rectangle, would then represent

¹Gene Networks Inc, <http://www.gene-networks.com>

active entities, such as reactions or operations. The places are then connected to the transitions (and vice versa) via directed arcs to form a network. The state of the system is then denoted by the number of tokens, or markings, within the places. Executing this involves discrete time steps, where in each step, tokens will be consumed or produced by the transitions that are connected to the places containing them. Note that arcs can only connect components of different types together (i.e. places to transitions and vice versa).

The HFPN adds more semantics and functionality to the Petri Net by allowing not only discrete components, but also continuous versions of them (hence explaining the term ‘Hybrid’). It also introduces two additional kinds of arcs - the Inhibitory arc and the Test arc. The graphical representation of the components are shown in Figure 1

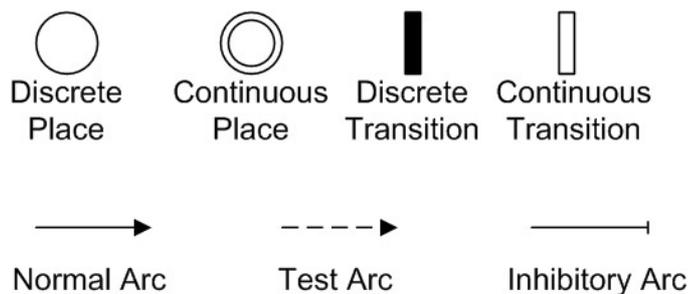


Figure 1: Components of the Hybrid Functional Petri Net.

As the name implies, discrete places can only hold non-negative integers (equivalent to the number of tokens in the Petri Net model) while continuous places can contain non-negative real numbers. A discrete transition can only activate, or fire, when its firing conditions (the number of tokens in its incoming places) are satisfied for a certain duration of time, specified by a *delay function*. A continuous transition, on the other hand, has a *firing function* which denotes the rate of consumption from its input places when its firing conditions are satisfied. Unlike the discrete transition which will only fire after a specified delay, the continuous transition fires instantaneously and continuously.

In addition to the normal arc, the HFPN introduces two additional arc types. The inhibitory arc performs the opposite role of a normal arc, preventing the connected transition from firing when the value of the input place satisfy a specified condition. In the biological context, this form of arc can be used to model molecular inhibition. Test

arc, on the other hand, behaves like a normal arc, except that no tokens or values are being consumed from the incoming place. Due to the additional components, there is a need for restrictions with regards to the possible ways places and transitions can be connected together. A normal arc can connect a place to a transition and vice versa (With the exception of connecting a discrete place to a continuous transition). Test and inhibitory arcs are restricted to only connect incoming places to transitions as they both involve satisfying a precondition.

By performing such connections, an entire network representing the reactions in a cell can be constructed.

2.2 Modeling Methodology

In this section, we look at the modeling of biological pathways using HFPN and present a systematic way to represent the various reactions.

One of the main assumptions for biochemical modeling is that the molecules are evenly distributed throughout the entire system, such that their concentration can be represented by a single variable. In the HFPN model, this would be represented by a continuous place whose value denotes the concentration of a particular protein type. The presence or the absence of certain conditions, such as serum, can be modeled using discrete places instead. Next we will consider the different types of reactions that can occur and the ways of representing them.

2.2.1 Association, Dissociation and Translocation

Association involves the binding of proteins to form complexes (dissociation being the opposite) while translocation involves the molecules moving from one region of the cell to another, such as from the cytoplasm to the nucleus. These three types of reactions do not chemically modify the proteins. Other than protein-mediated translocation, all of them follow the mass action law, which states that the rate of a reaction is dependent on the current concentration of its participating reactants. The equation denoting a binding reaction of two reactants and its HFPN equivalent is shown in Figure 2.

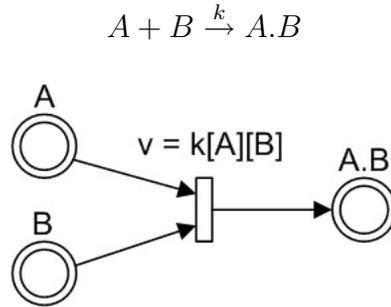


Figure 2: HFPN equivalent of a association reaction between two protein types A and B, and the equation v denoting the rate of reaction

2.2.2 Protein Modification

Protein modification often involves a catalyst, or enzyme, which essentially is not modified or consumed in the process. There are several considerations and variations for mathematically representing such reactions but for our model, we have decided to adopt the Michaelis-Menten model for enzyme kinetics, which is based on the quasi steady state approximation for chemical reactions. Under this scheme, most of the enzyme-catalyzed chemical reactions can be expressed as the form shown in Figure 3.

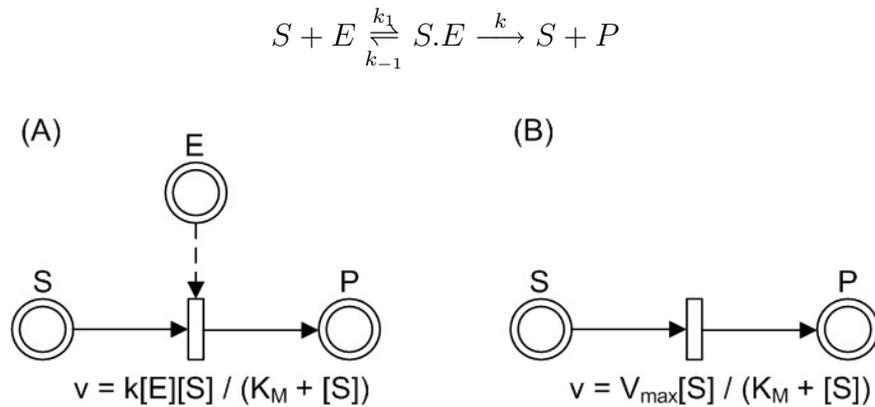


Figure 3: Panel A shows the HFPN representation of an enzyme-catalyzed chemical reaction. The arc from the place representing the enzyme E is a test arc as no enzyme molecules are consumed in a catalytic reaction. Panel B shows the similar reaction, but for the case where the identity of the enzyme is not known

In this scheme, S is the substrate, P is the product and E denotes the enzyme that catalyzes the reaction. The parameters k is the catalytic constant of the reaction and K_M is the Michaelis constant, which indicates the affinity of the substrate to the enzyme, with low values indicating that the ES complex is held together very tightly. In cases where we know that the reaction is an enzyme catalyzed reaction, but not the enzyme involved, the equation can be written in form depicted by Panel B of Figure 3. In this form, V_{max} is the maximal velocity of the reaction, and is an experimentally derived value.

There are other types of chemical reactions but for the pathway that follows, these types will be sufficient and we will see how to use them to systematically convert the pathway into its HFPN equivalent.

3 The Pathway

The *AKT* pathway has been studied extensively due to its role in apoptosis and cancer. In the prostate cancer cell line *LNCaP*, genetic defects cause such cells to be lacking in the active lipid phosphatase - *PTEN*, which is an inhibitor of the *AKT* pathway [17]. Hence this cell line has unusually high amounts of activated *AKT*, promoting cell survival. In experiments to knockdown *PDK1*, it was noticed that activity of other pathways are also affected, specifically the *MAPK* pathway, which also plays a part in cell growth and proliferation.

Hence, the model describing our pathway will consist of two main parts - the *AKT/PKB* pathway and the *MAPK* pathway. These two pathways are stimulated by external growth factors, which in this case will be modeled as a discrete element - serum. Several models representing these two pathways have already been created separately [16, 38, 36]. In our model, we are interested not only in the individual pathways but also in their cross interaction and effects on common downstream targets. The following sections will describe the main portions that make up the entire pathway and show how their HFPN model is derived.

3.1 AKT/PKB Pathway

The mechanism behind the activation of the *AKT* pathway is the membrane recruitment and activation of the signaling proteins. Binding of growth factors to the cell surface receptors will cause them (the receptors) to act as scaffolds for specific binding interactions with cytosolic proteins. This leads to the activation of *phosphoinositide 3-kinase* (*PI3K*), which will catalyze the phosphorylation of *phosphatidylinositol di-phosphate* (*PIP*₂) at the inositol ring, forming *PIP*₃. This reaction can be reversed by the lipid phosphatase *PTEN*. However, the cell line we are working with is the *LNCaP* cell line, which is lacking in *PTEN*. Hence during modeling, this concentration level will be kept to a minimal. Also, activation of *PI3K* can be eliminated by treating the cell with *LY294002*, a selective *PI3K* inhibitor [49].

Activated *PIP*₃ will then recruit proteins containing the *pleckstrin homology* (PH) domains, which includes the protein *AKT*. At the cell membrane, *AKT* will then be able to interact with *3-Phosphoinositide-dependent protein kinase 1* (*PDK1*). Through this interaction, *AKT* is activated by a sequential phosphorylation of *Thr*³⁰⁸ and *Ser*⁴⁷³ [13, 35], the former being catalyzed by *PDK1* and the latter by a yet-to-be-identified kinase, and is presently dubbed *PDK2* [21]. Activated *AKT* will then be released into the cytoplasm where it will further activate downstream targets such as the *Bcl-2 Antagonist of Cell Death* (*Bad*) and the *forkhead transcription factor* (*FKHR*). Activated *AKT* is regulated by *protein phosphatase 2A* (*PP2A*), which deactivates it by removing the phosphate groups. In our experiments involving the absence of serum, there is still slight *AKT* activity (Figure 13), implying that there might be basal activation of *PI3K*, which we will model as a separate reaction. The HFPN model of this scheme is shown in Figure 4.

3.2 MAPK Pathway

The *MAPK* pathway is highly conserved across several species. It involves several levels of kinases, each activating the subsequent level (hence the name cascade). Such an arrangement has an effect of ultra-sensitivity, where the cascade would behave like a discrete switch, turning from off to on over a very narrow range of stimuli, while responding less to other amounts of stimulation [19].

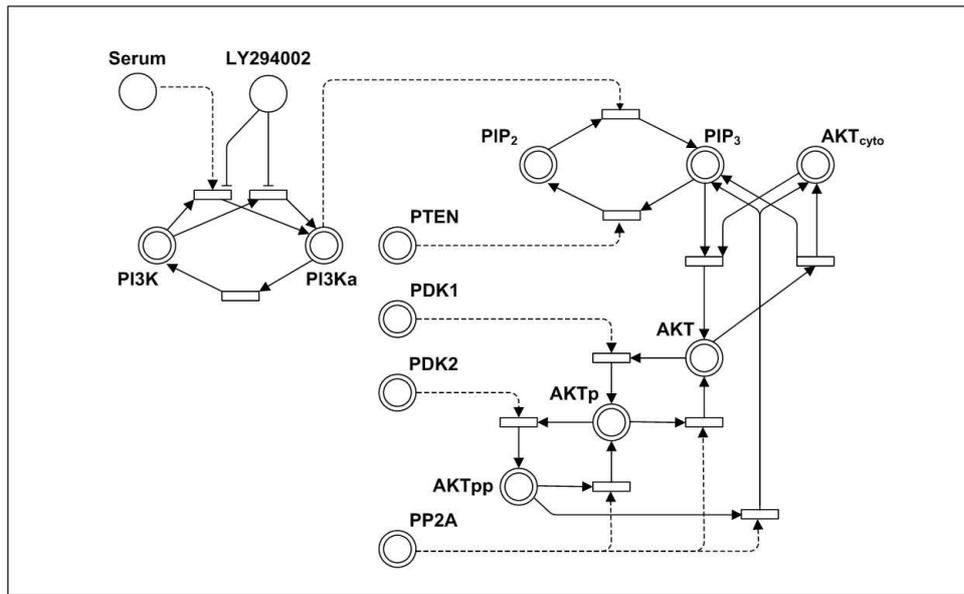


Figure 4: The HFPN model of the AKT pathway. Phosphorylated proteins are appended with ‘p’, with AKT having two of them to indicate double phosphorylation. Proteins that are activated but not via phosphorylation are appended with ‘a’. For PI3K, there are two activating transitions to denote serum and basal activation. For AKTpp, there are two ways to remove its phosphate group. It can be dephosphorylated either at the cell membrane, becoming singly phosphorylated, or after it has been released into the cytoplasm.

Similar to the *AKT* pathway, it is stimulated by the binding of external growth factors to the cell membrane. This causes the proto-oncogene *Ras* to be converted to its active conformation, having *GTP* bounded to it instead of *GDP*. In this conformation, its binding affinity to the *Raf* kinase is increased, causing *Raf* to be recruited to the membrane and become activated. This, in turn, causes it to activate the dual specific protein *MEK* by phosphorylating it at its serine/threonine residues. This is shown by the double phosphorylation of *MEK* in Figure 5. The activated *MEK* then carries on to phosphorylate *ERK* [26]. Finally, activated *ERK* will induce cell growth and proliferation by activating a variety of downstream transcription factors.

Several mathematical models describing this pathway have been created, such as the Kholodenko model and the Schoeberl model [1, 8, 16, 43, 50]. However it is not very clear which one is more accurate as they differ in the way certain reactions are being modeled as well as the proteins that are involved. For this portion of our pathway, it

follows largely the one presented in [16].

Apart from the regular components that make up the pathway, it has also been reported that phosphorylated *AKT* is a negative regulator of the *MAPK* pathway. It does this by phosphorylating *Raf* at *Ser*²⁵⁹, thereby inhibiting its activity along the *MAPK* pathway [40]. This shows that biological pathways do not always work in isolation, i.e. their dynamics are not shielded from one another. For our model, we represent it as a de-activation of *Raf*, following the model in [16]. In addition, the knockdown experiments that are being performed on *PDK1* also seemed to indicate that *PDK1* does play a positive role in the activation of *ERK* in a *MEK* dependent manner (Figure 11).

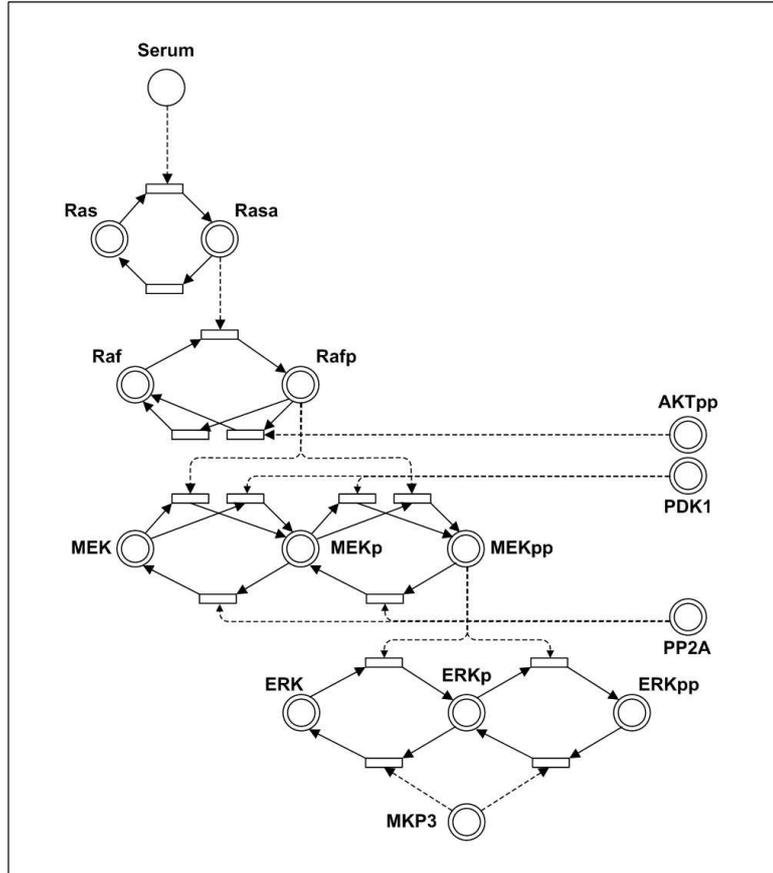


Figure 5: The HFPN model of the MAPK pathway. Presence of serum will lead to the activation of Ras, resulting in the phosphorylation of Raf and downstream targets MEK and ERK. Activated MEK is regulated by PP2A while ERK is de-activated by MKP3 [22, 24]

3.3 Downstream Targets - Bcl-2 Family

One of the ways *AKT* and *ERK* affect apoptosis is via their regulation of *Bad*. *Bad* is a protein that is involved in regulating apoptosis by the competitive heterodimerization of *Bcl-2* and *Bcl-XL* with itself and *Bax*. *Bcl-2* and *Bcl-XL* are anti-apoptotic proteins which inhibit the release of *cytochrome c* from the mitochondria [14]. By releasing mitochondrial *cytochrome c* into the cytosol, *procaspase 9*, a protein directly involved in cell death, will be activated, leading to apoptosis. This release of *cytochrome c* is induced by *Bax*. *Bcl-2* and *Bcl-XL* prevents this by binding to *Bax*, preventing its localization at the mitochondria. However, when *Bad* is sequestered to either *Bcl-2* or *Bcl-XL*, it displaces *Bax* from them, allowing *Bax* to promote cell death. Therefore, phosphorylating *Bad* will abrogate the apoptotic function of *Bax*, since phosphorylated *Bad* is unable to dimerize with either *Bcl-2* or *Bcl-XL*, allowing them to associate with *Bax* [39, 51].

AKT and *ERK* regulate the apoptotic function of the *Bcl-2* family by phosphorylating *Bad* at two sites - *Ser*¹³⁶ and *Ser*¹¹² respectively. *ERK* does not phosphorylate *Bad* directly. Instead, it does so via another kinase - the *P90RSK*, which is found to phosphorylate *Bad* at *Ser*¹¹² both *in vitro* and *in vivo* [47]. At the same time, activated *PI3K* can also regulate apoptosis by preventing *Bax* localization to the mitochondria, although the actual mechanism is not known. Hence the influence of *PI3K* on *Bax* is considered an indirect interaction [48] although we have modeled it as an enzyme that displaces *Bax* from the mitochondria.

3.4 Reactive Oxygen Species and PAK1

Several endogenous and exogenous activities in the cell as well as the action of certain protein scavengers influence and control the levels of reactive oxygen species (*ROS*) in the cell. Members of this species include the superoxide O_2^- and hydrogen peroxide H_2O_2 . Experiments have shown that prostate cancer cells produce substantial amounts of *ROS* and one of the sources could be *NOX5 NAD(P)H* oxidase [7]. This generation of *ROS* may be inhibited by the flavoprotein-dependent *NAD(P)H* oxidase inhibitor - *diphenylene iodonium* (*DPI*) [7]. *ROS* is needed in the activation of *P90RSK* to stimulate the expression of genes involved in cell growth and survival as well as activation of Na^+/H^+ Exchangers (*NHE*) [42]. *ROS* also mediates *MEK* and *ERK* activity, possible via activation of the *p21-activated kinase 1* (*PAK1*) [11]. Activated *PAK1* can then

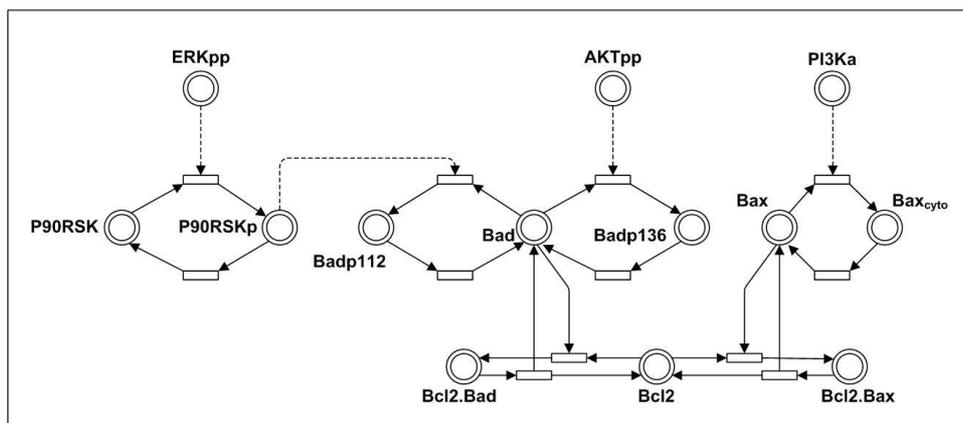


Figure 6: The interaction of the various proteins with the Bcl-2 family members - Bax, Bad and Bcl-2. There is competitive binding of Bad and Bax with Bcl-2. Phosphorylation of Bad at Ser^{112} or Ser^{136} causes more Bcl-2 to be able to bind to Bax, preventing it from causing apoptosis

regulate the *MAPK* pathway by regulating the phosphorylation of *Raf* [9]. However it is to be noted that the levels of *ROS* in the cell must be carefully regulated. Although known to promote cell survival, too high a concentration of *ROS* in the cell will also lead to cell death due to oxidative stress. Related to the *PAK1* mediated *ERK* activation is the role of *PI3K*, adding on another possible cross-interaction between the *AKT* and the *MAPK* pathways. Experiments have shown that *PI3K* can regulate the activity of *PAK1* by phosphorylation [9]. *PAK1* is also shown to be able to phosphorylate *Bad* at Ser^{112} and Ser^{136} both *in vitro* and *in vivo* [44].

The entire model, including the above-mentioned pathways and their cross-interactions is shown in Figure 8. However, even with the structure of the model being worked out, the model is still not complete as we do not know the rates of the various reactions that drive the model. This brings us to the next section to alleviate this issue - Parameter Estimation.

4 Parameter Estimation

Like any modeling endeavour, the greatest bottleneck for such quantitative modeling is the estimation of the parameters for the various rate reactions. Technical difficulties and huge resource requirements make the experimental determination of all the param-

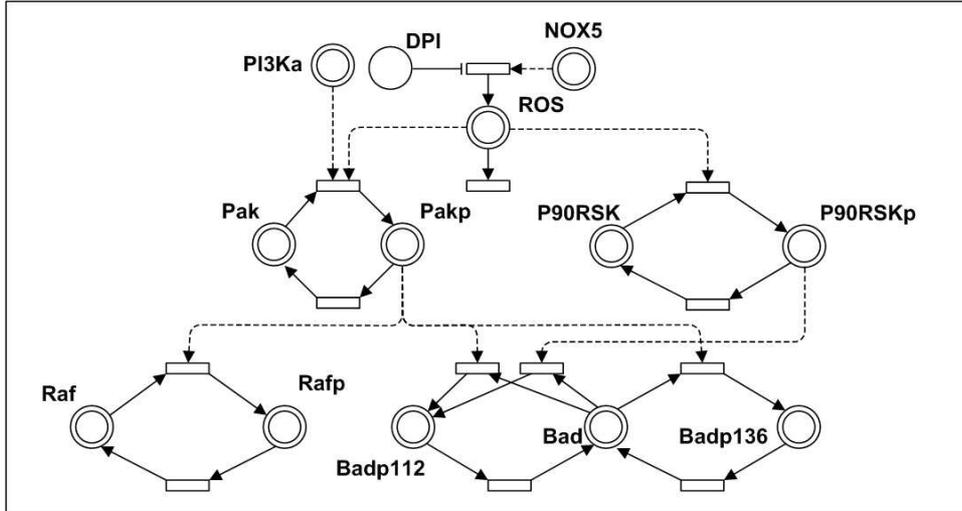


Figure 7: Role of ROS in the regulation of P90RSK and MAPK pathways. In our model, only ROS has reactions denoting the generation and degradation of ROS

eters impossible. Other than obtaining some of them through an extensive literature search, the rest would have been estimated. This amounts to a nonlinear global optimization problem and an in-depth review of the various approaches has been reported in [33]. Several parameter estimation algorithms have been considered and one of the best performing algorithm is a variant of the *Evolutionary Strategies* approach [41].

Every algorithm to estimate the parameters require searching through the solution space to find suitable parameter values. One of the main issues, especially when estimating the parameters for a huge network, is that the search space is too large. Our proposed method to alleviate this problem is to exploit the structure of the network, such that instead of searching the entire set of parameters, we concentrate only on a few parameters at any given time. This form of optimization is also known as *Alternating Optimization* [4].

4.1 Pathway Topology

Looking at the structure of the network in Figure 8, we can see that most of the reactions are catalytic reactions. Since we are adopting the Michaelis-Menten kinetics for enzymatic reactions, such reactions do not consume the enzyme. The concentration of the enzymes do not change with respect to the reactions that they catalyze. As such, the concentration flow of the molecules in Figure 8 can be viewed as being contained

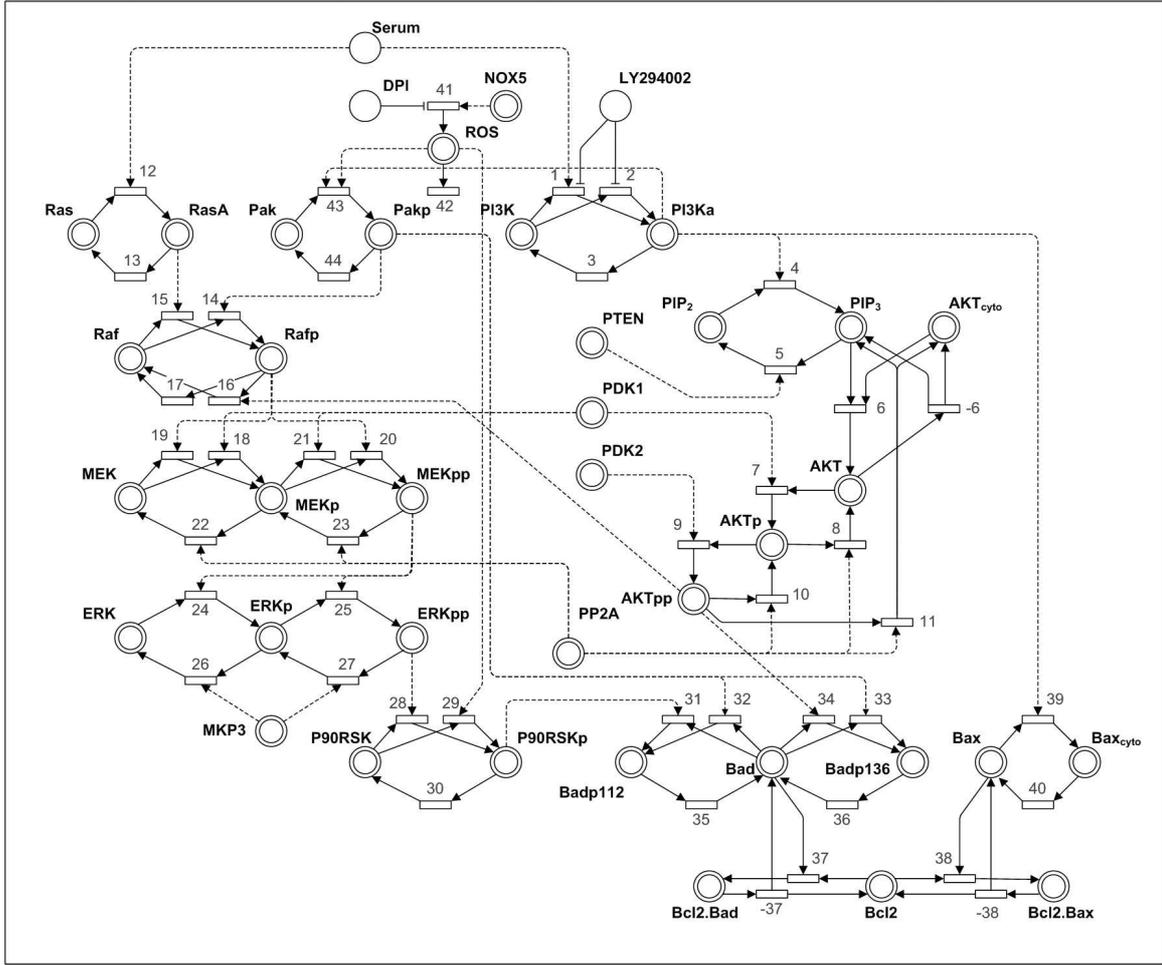


Figure 8: Scheme of the entire pathway, with the cross-interactions between the AKT and MAPK pathway and their influence on downstream targets

within separate modules, e.g. *PI3K* activity is contained within its own loop consisting of the places *PI3K* and *PI3Ka* and molecules do not flow into other places.

We let X be the set of variables (or places) representing protein concentration and V be the set of chemical reactions. Each variable is denoted by x_i and its rate of change is

$$\frac{dx_i}{dt} = \sum_{j \in J} c_{ij} v_j$$

where v_j is the reaction for which x_i is either a reactant or product of, and c_{ij} is its stoichiometric coefficient.

Definition 1 (Module) A module is a smallest subset $M' = (X', V')$ such that it satisfies the following properties. For all variables x_i where $x_i \in X'$ and $dx_i/dt = \sum_{j \in J} c_{ij} v_j$, it

must be the case that $v_j \in V'$. Also for any two modules $M' = (X', V')$, $M'' = (X'', V'')$, $X' \cap X'' = \emptyset$ and $V' \cap V'' = \emptyset$

Hence we can simplify the entire network into a Directed Acyclic Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where the nodes $\mathcal{V}_i \in \mathcal{V}$ are the individual modules and the edges $\mathcal{E}_i \in \mathcal{E}$ are the catalytic activation/inhibition between the modules. An example of how the scheme in Figure 8 can be simplified into a graph is shown in Figure 9. Note that not all pathways are amenable to such simplification. Here we assume that there are no major feedback loops in the model. Enzymes that do not change in concentration, such as *PDK1* are left out for brevity.

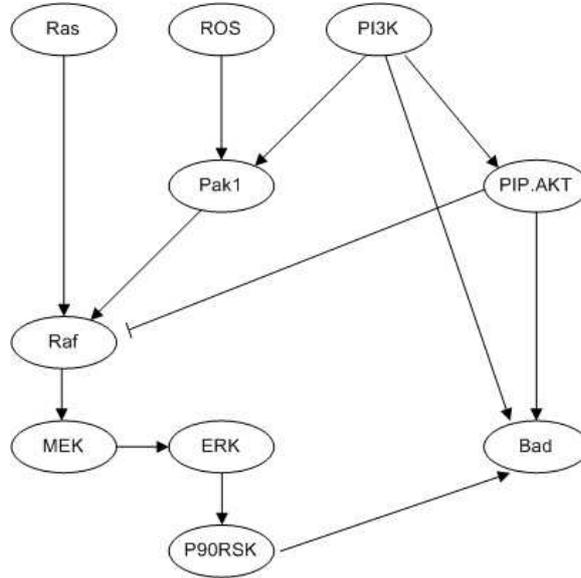


Figure 9: Directed acyclic graph of the individual modules that make up the pathway. For brevity, enzymes that do not change in concentration such as the PP2A and PDK1 are not shown in the diagram

4.2 Topological Ordering

Estimating the parameters of the pathway involves fitting the model with available experimental data, usually steady state protein concentration levels, or less frequently, time-series profiles. Most parameter estimation algorithms regard the system in its entirety. For reasons already mentioned, we shall focus only on optimization via Evolutionary Strategies. Such algorithms require several model evaluations to assess the fitness of the parameters or solutions.

In Evolutionary Strategies, current solutions are being ‘recombined’ and ‘mutated’ in every iteration. Hence for each iteration, or generation, the algorithm has to work with a solution space with exponential dimensionality. Whilst this is inevitable, we can reduce it by fitting the separate modules individually.

Taking the pathway in Figure 9 as an example, one can see that the choice of parameters in the Pak1 module does not affect the choice of parameters in the PI3K module. However, the same cannot be said for the reverse, where simulating the system with differing parameters in the PI3K module will lead to changes in enzyme profiles that will ultimately affect how the Pak1 module is to be fitted. Hence to work with the modules individually, the PI3K module must be fitted prior to Pak1, leading to the notion of precedence constraints. The nodes in the DAG can therefore be topologically ordered and their parameters will then be estimated in that order. There are already several efficient algorithms to perform topological sort on a DAG and hence it will not be further elaborated [45].

4.3 Rank

Despite being arranged in topological order, it is sometimes not possible to estimate the parameters due to unavailable data for fitting them with. As such, there is a need to group multiple modules together and assign them a rank to denote which modules are to be fitted together. Usually for fitting, there will not be one, but a series of experiments, each providing information on certain variables. For a variable x_i , we denote (if available) its profile for experiment j as $x_{ij}^{(e)}$.

Definition 2 (*Rank*) *The rank is a grouping of the modules in their linear extension such that for each group $\{M'\}$ of the same rank, there exists exactly one M'_k that has at least a variable x_i with an experimental profile $x_{ij}^{(e)}$.*

Additional requirements are that other than the last rank, the module with the experimental profile must be the last module within its group when arranged in topological order. When representing the network as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each vertex \mathcal{V}_k represents the module M_k , vertices with no experimental profiles take on the rank of the closest vertex \mathcal{V}_l where there is a path $\mathcal{V}_k \dots \mathcal{V}_l$. This will impose a slight restriction on the partial order of the modules that do not have experimental profiles to fit them to. The requirement for each rank to have only one module with a profile is to minimize the

number of modules in a rank, so that the solution space to search from is also kept to a minimal.

For the graph in Figure 9, we have the experimental profiles for the following protein types - *AKT*, *Raf*, *MEK*, *ERK*, *P90RSK* and *Bad*. Hence the modules and their grouping will be in this order - {PI3K, AKT}, {ROS, Pak1, Ras, Raf }, {MEK}, {ERK}, {P90RSK}, {Bad}, with the highest rank of 1 being the group {PI3K, AKT} and the lowest rank of 6 for the group {Bad}.

4.4 Evolutionary Strategy

The choice of estimation algorithm being used is independent on the ranking, i.e. one can choose to perform deterministic optimization algorithms such as the Levenberg - Marquardt method, or stochastic methods such as Genetic Algorithms, or even a mix of them for different ranks, depending on the characteristics of the modules within that rank. However, for reasons already mentioned, we have chosen Evolutionary Strategies for our work, specifically the $(\mu + \lambda) - ES$ algorithm [3]. The following steps show the basic implementation of the algorithm.

1. Generate μ parent vectors, each being a set of parameters $P_i = (p_{i1}, \dots, p_{in})$.
2. Create λ new offsprings, with each child being formed from the recombination of two randomly selected parent parameters.
3. Mutate the offsprings.
4. For each child parameter set, evaluate the model to assess its fitness score.
5. Select the μ most fit parameters from the $(\mu + \lambda)$ parameters to form the next generation.
6. Repeat steps 2 to 5 for a predetermined number of generations, or when no better parameters could be obtained.

This process is repeated in rank order. In addition, we need to maintain a set of fixed parameters obtained after estimating parameters of the previous ranks or those that have already been determined beforehand. We denote this set as P_f and for each generated parent parameters P_i in step (1), $P_f \subseteq P_i$. Recombination and mutation

are then performed on the parameters p_i where $p_i \notin P_f$. To assess the fitness of the parameters, the cost function must also take the current rank k into consideration.

$$J = \sum_{i,j,t} w_{ij} \sqrt{(x_{ij}(t) - x_{ij}^{(e)}(t))^2}$$

where $x_{ij}(t)$ is the predicted value of x_i in experiment j at time t . $x_{ij}^{(e)}(t)$ is its corresponding experimental value. w_{ij} is the weight used to normalize the contributions of each term to the cost function and it is usually taken to be $w_{ij} = 1/\max(x_{ij})$. Also, for all x_{ij} used in the cost function, $\text{Rank}(x_{ij}) \leq k$. After estimating the parameters for a particular rank k , we append the best parameters with rank k to the set of fixed parameters P_f .

Based on this algorithm, we estimate the parameters for the rate reactions that govern the network in Figure 8. The possible values for the parameters are bounded within physiological range, with association constants $k_i \in [10^{-6}, 10^{-2}] \text{ nM}^{-1}.\text{s}^{-1}$, dissociation constants $k_{-i} \in [10^{-6}, 10^{-1}] \text{ s}^{-1}$ and catalytic constants $k_d \in [10^{-6}, 1000] \text{ s}^{-1}$. For Michaelis Menton parameters, $K_M \in (0, 10^9] \text{ nM}$ and $V_{max} \in (0, 10^5] \text{ nM}.\text{s}^{-1}$. Production and degradation rate constants are kept between 0 and 1. The estimated parameters are shown in Table 1.

4.5 Experimental Data

It is never the case where we will have one set of complete data on all the molecule types obtained from a single experiment. Rather, it will more often come from different experiments with different set-up and the data profiles can either be steady state or time series. In addition to that, we can measure the profiles of only a few types of molecules, hence requiring grouping as described previously. For our pathway, we estimated the parameters based on three sets of experimental data, two of them being steady state values and the third, time-series data.

4.5.1 Experiment Set 1 - Knock-down Experiments

The first set of data is obtained from knockdown experiments. *LNCaP* cells are being transfected with small interfering RNA (siRNA) to knockdown the genes via RNA interference. In this set of experiments, cells are treated with 20 *nM*, 50 *nM* and 100 *nM* of siRNAs for the proteins *AKT* and *PDK1* to reduce their expression levels.

The siRNAs are transfected using the calcium phosphate method and the cells are then incubated for either 48 or 72 hours. The expression levels of selected proteins are then assayed and visualized via Western blotting.

Panels A, B and C of Figure 10 show the expression levels of *AKT* and its activation levels under various knock-down conditions. The three experiments are carried out independently.

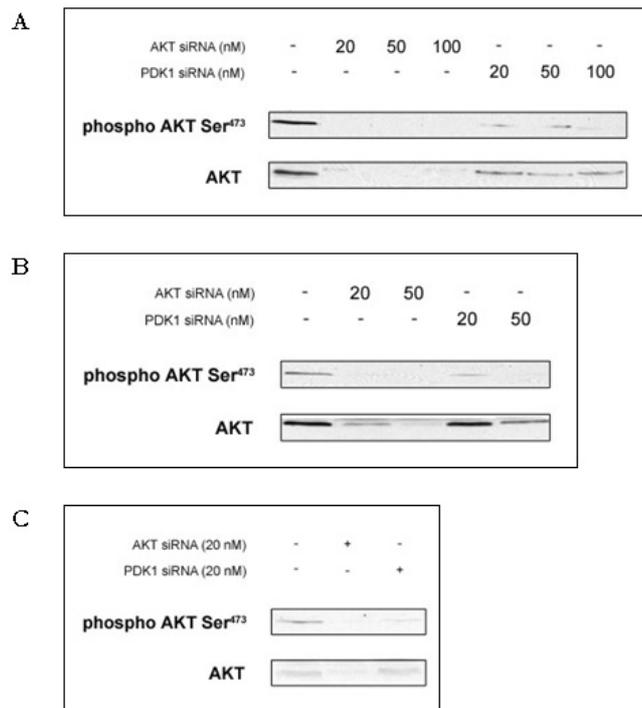


Figure 10: AKT and PDK1 knockdown on AKT expression and activation levels.

The effects of both treatments on the expression levels of *AKT* is already known, and such effects are seen from the experiments. Needless to say, *AKT* siRNA is suppose to reduce the total amount of *AKT* concentration, as can be seen in the figure. 20 nM of *AKT* siRNA is enough to reduce the total concentration of *AKT* to an almost negligible level. *PDK1* is an activator of *AKT*, phosphorylating it at its Thr³⁰⁸ residue. Reducing *PDK1* expression levels via siRNA will down-regulate *AKT* activation levels, as can be seen in Figure 10. However, it is noted that the total *AKT* concentration is also being reduced with *PDK1* siRNA treatment. This observation is consistent with Panels A and B of Figure 10. However, this interaction has not been observed anywhere in the model

and could be attributed to experimental inaccuracies.

The next set of data shows the expression levels of the various proteins involved in the *MAPK* pathway, namely *Raf*, *MEK* and *ERK*.

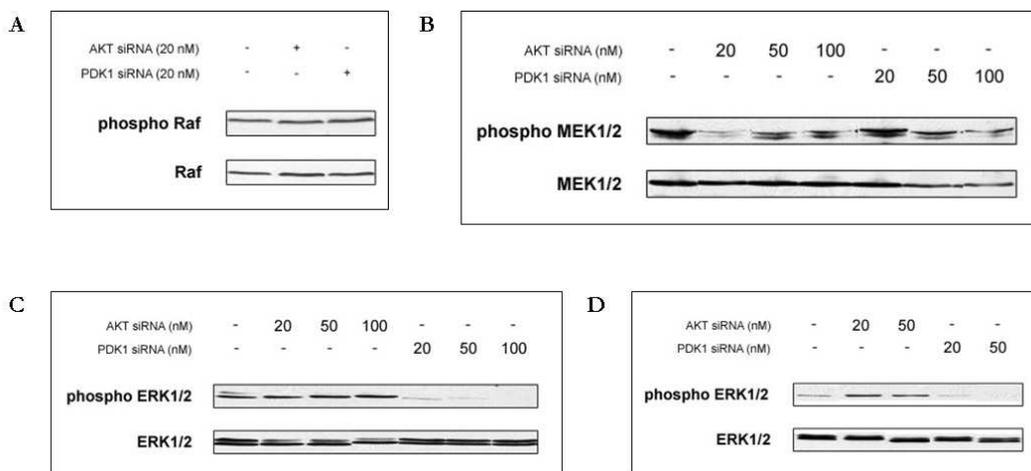


Figure 11: Effects of the knockdown experiments on the various parts of the *MAPK* pathway. Panel A shows *Raf* expression and activation levels upon knocking down *AKT* and *PDK1* while Panel B shows that of *MEK*. Panels C and D are two independent experiments for the activation levels of *ERK*.

From Figure 11, it is observed that variations of *PDK1* levels has not much effects on the activation of *Raf*. However, it has been reported in [52] that *AKT* can regulate the *MAPK* pathway by phosphorylating *Raf* at *Ser*²⁵⁹. Although this is not readily observed in Panel A of Figure 11, its downstream effects on *MEK* and *ERK* can be seen, with increasing levels of *MEK* and *ERK* activation when the expression levels of *AKT* decreases (Panels B to D). On the other hand, decreasing the levels of *PDK1* does have an effect of attenuating *MEK* and *ERK* activity. This leads to a somewhat interesting manner in the way various components of the *AKT* pathway play a part in regulating the *MAPK* pathway, possibly as a mechanism to restrict the influence of *PDK1* on the *MAPK* pathway where prolonged *PDK1* activation will lead to increased *AKT* activity, which in turns down-regulate the activation of *ERK* in the *MAPK* pathway.

P90RSK plays an essential role in cell growth by activating several transcription factors, including the *Na*⁺/*H*⁺ exchangers. This protein is also being implicated by the

activation of the *MAPK* pathway, and hence is one of the proteins that are being assayed in the experiments. One of the downstream targets of *P90RSK* is *Bad*. *P90RSK* blocks *Bad*-mediated cell death by phosphorylating *Bad* at *Ser*¹¹² [47]. At the same time, *Bad* is also one of the downstream targets of *AKT*, being phosphorylated at *Ser*¹³⁶ in a *PI3K*-dependent manner [47, 52].

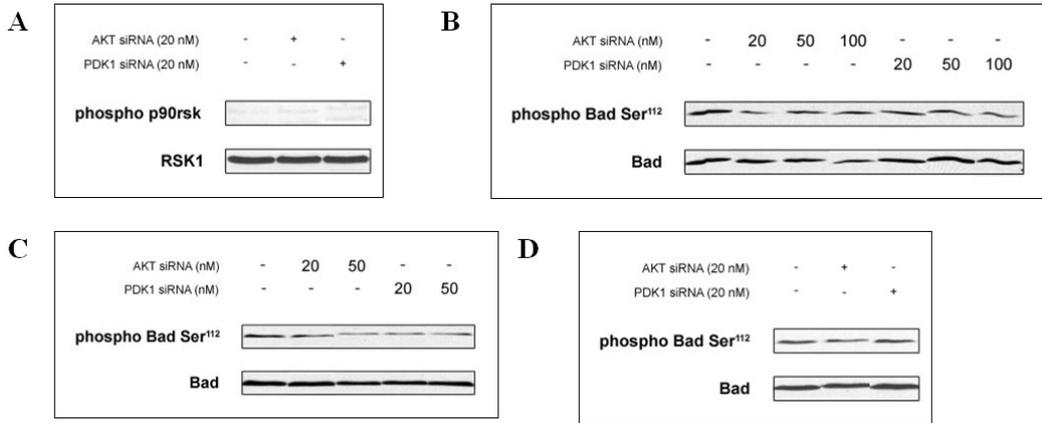


Figure 12: AKT and PDK1 knockdown effects on P90RSK(Panel A) and Bad (Panels B, C, D) activation

Results from Panel A in Figure 12 is rather inconclusive, with no apparent change in *P90RSK* activation under any conditions. *Bad* activity, however, is reduced with both knockdowns (Panel C). It is not clear how *AKT* siRNA transfection will lead to decreased *Bad* activation at *Ser*¹¹². Perhaps more experiments are needed to shed more light on this. On the other hand, *PDK1* siRNA would most probably have regulated *Bad* phosphorylation via the *MAPK* pathway.

4.5.2 Experiment Set 2 - PP2A Knockdown and Post Transfection Treatment

The second set of experiments involves not only RNA interference (for this set, *PP2A* siRNAs are being used instead), but also the treatment of cells with reagents *LY294002* and *DPI*. The cells are exposed to the various conditions for 1 hour before being assayed for protein expression.

PP2A is a family of phosphatase that reverses the activation of the various proteins by removing the phosphate groups. Some of these proteins include *AKT* and *MEK*. It

is also reported that *PP2A* deactivates *ERK*. However, this regulation is dominated by the highly selective activity of *MKP3* [22]. Nevertheless, *PP2A* is still crucial in the regulation of the *MAPK* pathway via *MEK* dephosphorylation.

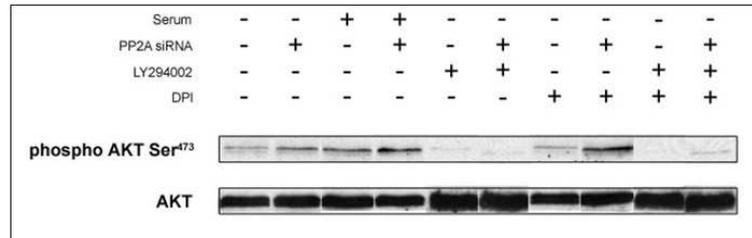


Figure 13: Effects of the various treatment on the activity of AKT

The *LNCaP* cell line lacks the active lipid phosphatase *PTEN*, a negative regulator of *PIP₃* and hence will display high levels of *AKT* activity [34]. With that, we can see that even in the absence of serum, there will be some activation of the *AKT* pathway, with its level probably being kept moderate by *PP2A*. The importance of *PI3K* in the *AKT* pathway can be seen in the treatment of the cells by *LY294002*, a selective *PI3K* inhibitor. In Figure 13, adding of *LY294002* appears to abrogate *AKT* phosphorylation, hence inhibiting this pathway. *DPI* on the other hand, does not affect *AKT* activity much, showing that *ROS* does not have any profound impact on the activity of the *AKT* pathway.

The role of *PP2A* in deactivating the *AKT* pathway is shown in Figure 13. For almost every experiment, its corresponding counterpart with *PP2A* siRNA appears to have a higher level of *AKT* phosphorylation, which would show that *PP2A* indeed plays a role in negating the effects of *AKT* activation.

Next are the effects of *PP2A* knockdown and the various treatments on the components of the *MAPK* pathway.

The effects of serum on *MEK* activation is very obvious, with its activity increasing as much as 3-fold in the presence of serum. The knocking down of *PP2A* also has the effect of increasing *MEK* phosphorylation. The addition of *DPI* will affect the activation for both *MEK* and *ERK*, inhibiting their activity. This re-affirms the role of *ROS* as second messengers in signaling pathways, although we only show them in the context of

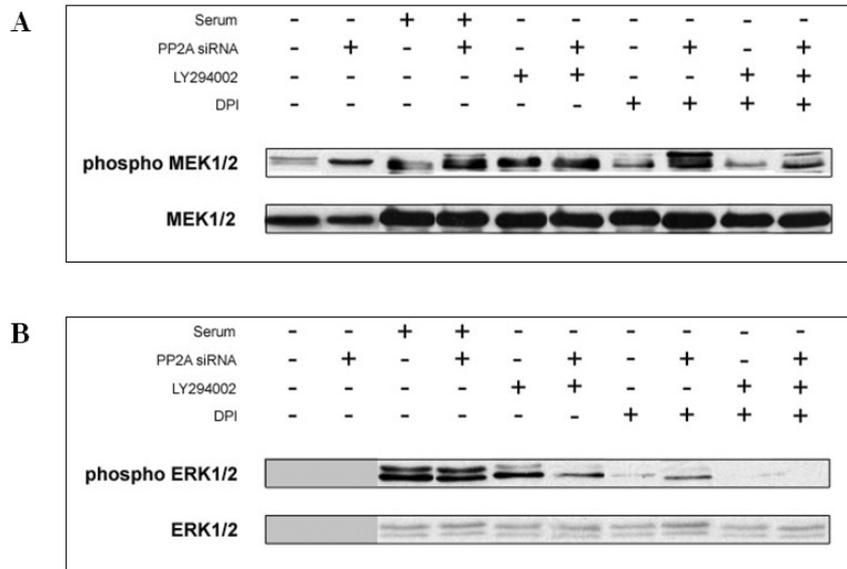


Figure 14: Effects of the various treatments on MEK and ERK activity. Greyed out regions means that the protein types were not assayed for that particular experiment

the *MAPK* pathway. One important thing to note about *ROS* is that its concentration and ratio must be kept within a certain level. Upsetting that proportion will lead to cell death either via oxidative stress, or by reductive stress. Further implications of the effects of *ROS* (as well as the ratio of its corresponding family members will be looked into in future studies).

4.5.3 Experiment Set 3 - Time Series Experiment

The third and final set of experiments is aimed primarily at finding out how protein activity changes with time under various conditions. (Different combinations of serum, *LY294002* and *DPI*). For the cells, they are initially kept in a serum starved state. They are then exposed to the different treatments and the readings of the various protein activation levels are taken at times 0 *min*, 10 *min*, 20 *min*, 30 *min*, 45 *min*, 60 *min* and 120 *min* after the start of the treatment. The proteins of interest in this set of experiments are *MEK*, *ERK*, *P90RSK* and *Bad* phosphorylation at *Ser*¹¹². The relative activation levels of the proteins are shown in Figure 16 as triangles and squares. The relative endpoints are already expected from the model, but having time profiles will enable us to have a tighter fit of the parameters.

Using all these three sets of experiments, we then estimate the parameters of the model such that it can reproduce most, if not all, of the observations made.

5 Simulation and Validation

The model is constructed using Cell Illustrator on a Pentium IV personal computer. To estimate the parameters, we broke the pathway down into modules separately and assigned them ranks according to the scheme presented in the previous sections. We then run the estimation algorithm, written in C++, on a PC cluster system. The algorithm was run a few times, and the most representative results (i.e. whose simulation profile matches the experimental ones the best) are used.

5.1 Effects of PDK1 siRNA on MEK and ERK Activation

The observation that prompted this investigation is the treatment of *LNCaP* cells with *PDK1* siRNA. Usually not thought to interfere directly with the *ERK* pathway, decreased concentration of *PDK1* shows a significant drop in both *MEK* and *ERK* activity, suggesting that *PDK1* might have been affecting the *ERK* pathway via an *MEK* dependent manner. This effect has been captured in our model. Figure 15 shows the western blot of *MEK* and *ERK* activation with various levels of *PDK1* siRNA treatment, as well as their corresponding simulation profiles. We simulate the effect of *PDK1* siRNA treatment by reducing its total concentration from 1000 *nM* (Control) down to 0 *nM* (100 *nM* siRNA treatment). In the control experiment, the level of *ERK* phosphorylation is moderate but once siRNA treatment is being administered, it shows an acute inhibition, dropping to nearly 0 *nM*, as can be seen in our simulations.

5.2 Effects of LY294002 and DPI

The experimental time series profiles are obtained mainly from cells that have been treated with the *PI3K* inhibitor *LY294002* and the *ROS* inhibitor *DPI*, both in the presence and absence of serum. Experiments show that both *DPI* and *LY294002* affects the *ERK* pathway in similar ways, suggesting that other than the *PDK1* and *AKT* interaction, there might still be some other forms of interaction. In our model, we realise this effect via an indirect action of activated *PI3K*. It has been reported in [9] that *PI3K* is one of the regulators of *Pak1* activation, which in turns regulates *Raf* activation. Hence,

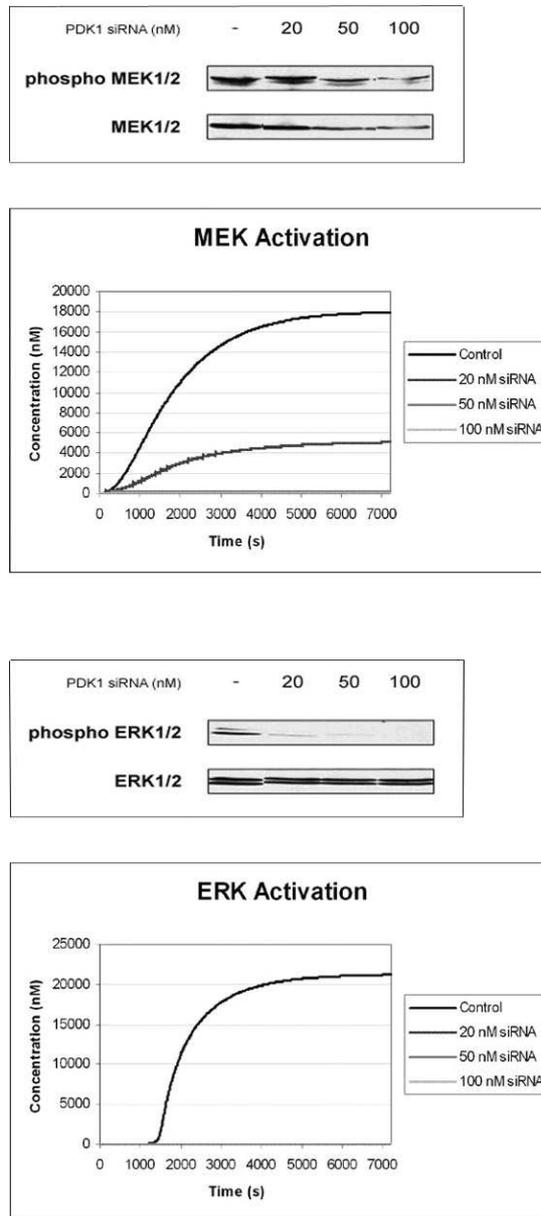


Figure 15: Profiles of the model simulation compared to the validation set

inhibition of *Pak1* activation could result in lower levels of *MEK* and *ERK* activity, and this can be achieved by adding either *LY294002* or *DPI*. The profiles of the simulation, compared to experimental data are shown in Figure 16.

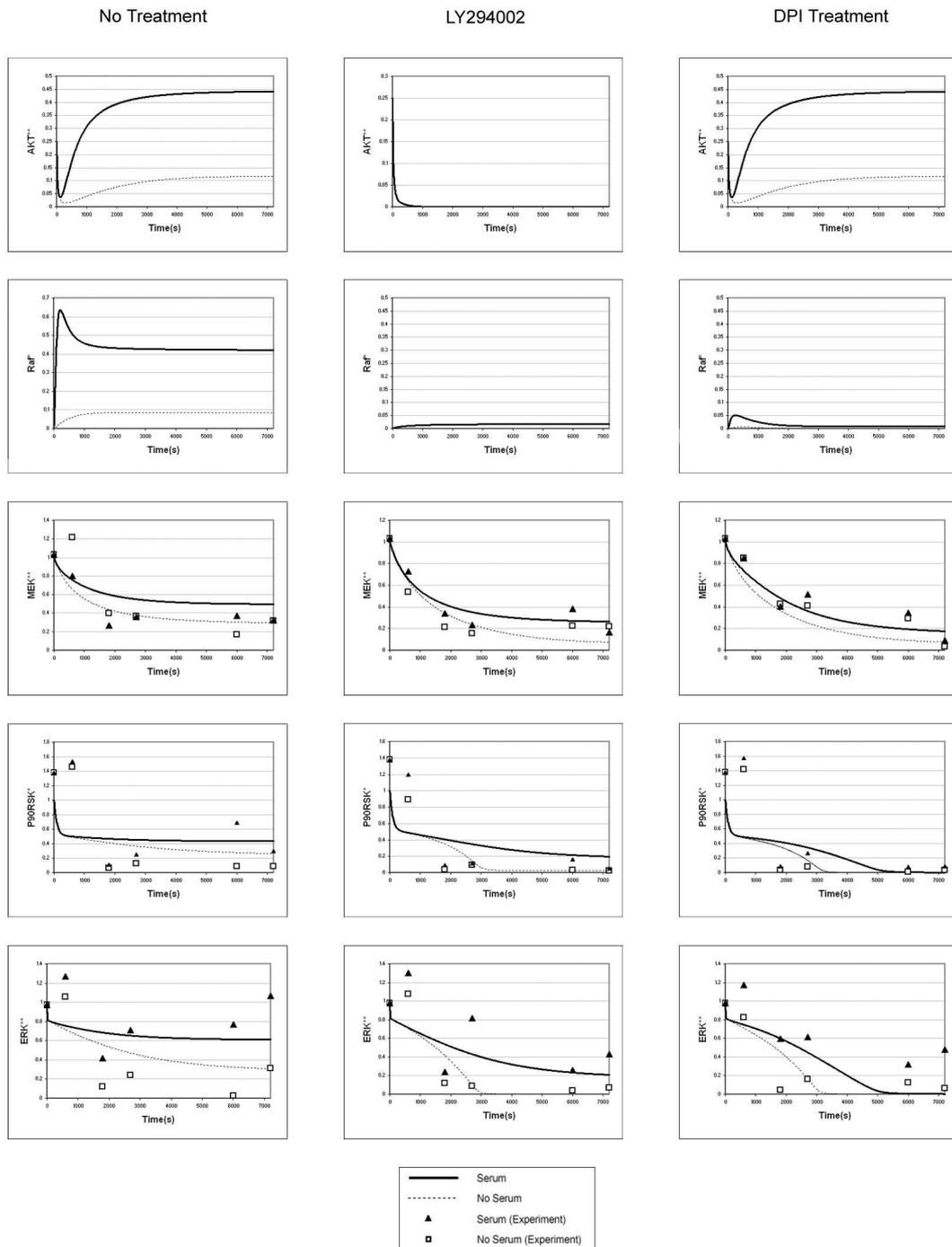


Figure 16: Profiles of the various activity levels under different treatment. Each graph shows the activity of the particular protein type in the presence (solid lines) or absence (dashed lines) of serum. In addition, some of the profiles include those obtained from experiments, with serum (triangles) and no serum (squares).

5.3 Validation

To validate the model and parameters, we simulate the system again, this time comparing it with another separate set of experimental data, or the validation data set. This experiment is a re-creation of the cell treatment with *LY294002*. However, due to experimental variations, the initial conditions of the various protein types are different and this is reflected in the simulation. This form of validation is a simplified form of model validation - the holdout technique [25]. Although it is not the best validation method, we use it due to the little and sparse nature of data sets that are obtained from laboratory experiments. The profiles of the following activated proteins *Raf*, *MEK*, *ERK*, *AKT* and *P90RSK* are shown in Figure 17.

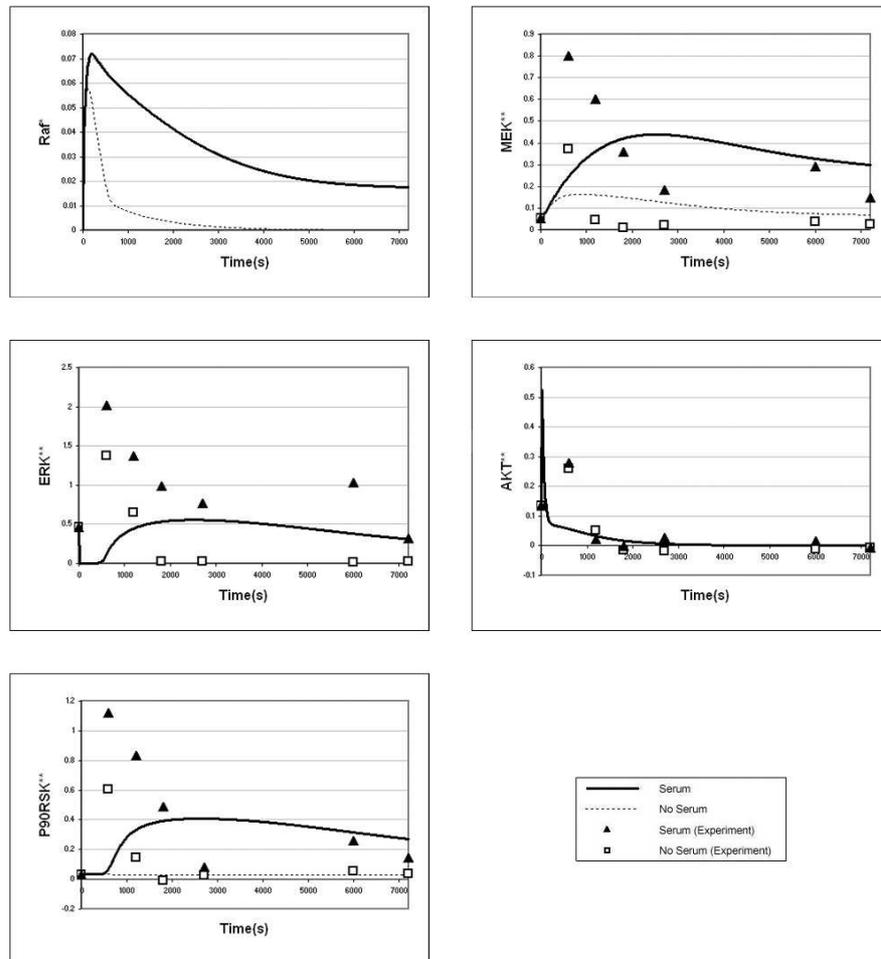


Figure 17: Profiles of the model simulation compared to the validation set

From the validation profiles, it is hard to conclude that the parameters are an exact fit. The *AKT* profile shows an almost exact fit while there are deviations in the activation levels of *MEK* and *ERK*. However, similarities in qualitative effects can still be seen, lending some level of confidence to the structure of the model. However, more work needs to be done to explain some of the dynamic profiles that can be seen from the experiments.

5.4 Parameters and Rate Reactions

The derived equations describing all the chemical reactions, and their estimated parameters, are shown in the Table 1. Cellular protein concentration were, whenever possible, taken from various literature sources. The remaining were then estimated over the range of $1 - 50000$ *nM*. However, note that since all the experiments were performed on the *LNCaP* cell line, the concentration of the protein *PTEN* is kept at a very low value of 0.1 *nM*.

6 Discussion

Computational modeling is indeed a useful tool to the biologists in understanding and re-creating network kinetics, as well as testing of hypotheses, as we have done so in this paper. Although being able to capture most of the effects, more tests are needed to be carried out to ascertain the validity of the hypotheses, especially the link between *PDK1* and *MEK*, as well as *PI3K* on *ERK* activation. The current ways in which we derive the model structure and its parameters are based on experimental data sets that we have and prior knowledge. Such data sets are sparse and diverse, differing not only in experimental conditions, but also in the cell type. For this model, we ensure consistency by only using data obtained from experiments that are being performed on *LNCaP* cells. In addition, when using the data for estimation, we ensure that the initial conditions mimic, as close as possible, the experimental set-up, i.e. the amount of already activated *AKT*, *MEK* and *ERK* at the start of the experiment. Even so, to reuse those data in validation will result in a form of a circular argument - i.e. using the data to estimate the parameters (or structure) and using their level of correlation as proof of correctness.

We try to alleviate this problem by using data sets from other experiments which are not used in estimating the parameters as the basis for comparison. This is also known as

| No | Rate Equation | Parameter | Ref |
|----|---|---|----------|
| 1 | $V_1[\text{PI3K}]/(K_1 + [\text{PI3K}])$ | $V_1 = 130, K_1 = 57800$ | Estimate |
| 2 | $k_2[\text{PI3K}]$ | $k_2 = 1.5 \times 10^{-4}$ | Estimate |
| 3 | $V_3[\text{PI3K}^*]/K_3 + [\text{PI3K}^*]$ | $V_3 = 7, K_3 = 8000$ | Estimate |
| 4 | $k_4[\text{PI3K}^*][\text{PIP}_2]/K_4 + [\text{PIP}_2]$ | $k_4 = 2.52, K_4 = 85200$ | Estimate |
| 5 | $k_5[\text{PTEN}][\text{PIP}_3]/K_5 + [\text{PIP}_3]$ | $k_5 = 15, K_5 = 306$ | Estimate |
| 6 | $k_6[\text{PIP}_3][\text{AKT}_{\text{cyto}}] - k_{-6}[\text{AKT}]$ | $k_6 = 1.25 \times 10^{-5}, k_{-6} = 4.22 \times 10^{-3}$ | Estimate |
| 7 | $k_7[\text{PDK1}][\text{AKT}]/(K_7 + [\text{AKT}])$ | $k_7 = 20, K_7 = 80000$ | [6] |
| 8 | $k_8[\text{PP2A}][\text{AKT}^*]/(K_8 + [\text{AKT}^*])$ | $k_8 = 35, K_8 = 34000$ | Estimate |
| 9 | $k_9[\text{PDK2}][\text{AKT}^*]/(K_9 + [\text{AKT}^*])$ | $k_9 = 20, K_9 = 80000$ | [6] |
| 10 | $k_{10}[\text{PP2A}][\text{AKT}^{**}]/(K_{10} + [\text{AKT}^{**}])$ | $k_{10} = 30, K_{10} = 64000$ | Estimate |
| 11 | $k_{11}[\text{PP2A}][\text{AKT}^{**}]/(K_{11} + [\text{AKT}^{**}])$ | $k_{11} = 8.5, K_{11} = 45100$ | Estimate |
| 12 | $V_{12}[\text{Ras}]/(K_{12} + [\text{Ras}])$ | $V_{12} = 0.825, K_{12} = 74500$ | Estimate |
| 13 | $V_{13}[\text{Ras}^*]/(K_{13} + [\text{Ras}^*])$ | $V_{13} = 0.198, K_{13} = 5.52$ | Estimate |
| 14 | $k_{14}[\text{Pak1}^*][\text{Raf}]/(K_{14} + [\text{Raf}])$ | $k_{14} = 0.18, K_{14} = 185$ | Estimate |
| 15 | $k_{15}[\text{Ras}^*][\text{Raf}]/(K_{15} + [\text{Raf}])$ | $k_{15} = 0.47, K_{15} = 150$ | Estimate |
| 16 | $k_{16}[\text{AKT}^{**}][\text{Raf}^*]/(K_{16} + [\text{Raf}^*])$ | $k_{16} = 3.3, K_{16} = 77.5$ | Estimate |
| 17 | $V_{17}[\text{Raf}^*]/(K_{17} + [\text{Raf}^*])$ | $V_{17} = 66, K_{17} = 16.7$ | [43] |
| 18 | $k_{18}[\text{PDK1}][\text{MEK}]/(K_{18} + [\text{MEK}])$ | $k_{18} = 5.53 \times 10^{-2}, K_{18} = 14700$ | Estimate |
| 19 | $k_{19}[\text{Raf}^*][\text{MEK}]/(K_{19} + [\text{MEK}])$ | $k_{19} = 3.5, K_{19} = 317$ | [43] |
| 20 | $k_{20}[\text{Raf}^*][\text{MEK}^*]/(K_{20} + [\text{MEK}^*])$ | $k_{20} = 2.9, K_{20} = 263$ | [43] |
| 21 | $k_{21}[\text{PDK1}][\text{MEK}^*]/(K_{21} + [\text{MEK}^*])$ | $k_{21} = 0.055, K_{21} = 14470$ | Estimate |
| 22 | $k_{22}[\text{PP2A}][\text{MEK}^*]/(K_{22} + [\text{MEK}^*])$ | $k_{22} = 0.058, K_{22} = 2232$ | [43] |
| 23 | $k_{23}[\text{PP2A}][\text{MEK}^{**}]/(K_{23} + [\text{MEK}^{**}])$ | $k_{23} = 0.058, K_{23} = 60$ | [43] |
| 24 | $k_{24}[\text{MEK}^{**}][\text{ERK}]/(K_{24} + [\text{ERK}])$ | $k_{24} = 16, K_{24} = 1.46 \times 10^5$ | [43] |
| 25 | $k_{25}[\text{MEK}^{**}][\text{ERK}^*]/(K_{25} + [\text{ERK}^*])$ | $k_{25} = 5.7, K_{25} = 5.21 \times 10^4$ | [43] |
| 26 | $k_{26}[\text{MKP3}][\text{ERK}^*]/(K_{26} + [\text{ERK}^*])$ | $k_{26} = 0.3, K_{26} = 160$ | [43] |
| 27 | $k_{27}[\text{MKP3}][\text{ERK}^{**}]/(K_{27} + [\text{ERK}^{**}])$ | $k_{27} = 0.27, K_{27} = 60$ | [43] |
| 28 | $k_{28}[\text{ERK}^{**}][\text{p90RSK}]/(K_{28} + [\text{p90RSK}])$ | $k_{28} = 2.5 \times 10^{-5}, K_{28} = 97.6$ | Estimate |
| 29 | $k_{29}[\text{ROS}][\text{p90RSK}]/(K_{29} + [\text{p90RSK}])$ | $k_{29} = 1.6 \times 10^{-5}, K_{29} = 81.4$ | Estimate |
| 30 | $V_{30}[\text{p90RSK}^*]/(K_{30} + [\text{p90RSK}^*])$ | $V_{30} = 0.058, K_{30} = 5.8$ | Estimate |
| 31 | $k_{31}[\text{p90RSK}^*][\text{Bad}]/(K_{31} + [\text{Bad}])$ | $k_{31} = 0.002, K_{31} = 346$ | Estimate |
| 32 | $k_{32}[\text{Pak1}^*][\text{Bad}]/(K_{32} + [\text{Bad}])$ | $k_{32} = 0.46, K_{32} = 710$ | Estimate |
| 33 | $k_{33}[\text{Pak1}^*][\text{Bad}]/(K_{33} + [\text{Bad}])$ | $k_{33} = 0.125, K_{33} = 1310$ | Estimate |
| 34 | $k_{34}[\text{AKT}^{**}][\text{Bad}]/(K_{34} + [\text{Bad}])$ | $k_{34} = 3.46, K_{34} = 307$ | Estimate |
| 35 | $V_{35}[\text{Bads}^{112}]/(K_{35} + [\text{Bads}^{112}])$ | $V_{35} = 5.8, K_{35} = 3450$ | Estimate |
| 36 | $V_{36}[\text{Bads}^{136}]/(K_{36} + [\text{Bads}^{136}])$ | $V_{36} = 180, K_{36} = 797$ | Estimate |
| 37 | $k_{37}[\text{Bad}][\text{Bcl} - 2] - k_{-37}[\text{Bcl} - 2.\text{Bad}]$ | $k_{37} = 4.78 \times 10^{-2}, k_{-37} = 1.35 \times 10^{-2}$ | Estimate |
| 38 | $k_{38}[\text{Bax}][\text{Bcl} - 2] - k_{-38}[\text{Bcl} - 2.\text{Bax}]$ | $k_{38} = 2 \times 10^{-3}, k_{-38} = 0.02$ | [18] |
| 39 | $k_{39}[\text{PI3K}^*][\text{Bax}]/(K_{39} + [\text{Bax}])$ | $k_{39} = 640, K_{39} = 38900$ | Estimate |
| 40 | $V_{40}[\text{Bax}_{\text{cyto}}]/(K_{40} + [\text{Bax}_{\text{cyto}}])$ | $V_{40} = 79800, K_{40} = 29200$ | Estimate |
| 41 | $k_{41}[\text{NOX5}]$ | $k_{41} = 1.26 \times 10^{-3}$ | Estimate |
| 42 | $k_{42}[\text{ROS}]$ | $k_{42} = 1.7 \times 10^{-3}$ | Estimate |
| 43 | $k_{43}[\text{ROS}][\text{PI3K}^*][\text{Pak1}]/(K_{43} + [\text{Pak1}])$ | $k_{43} = 0.33, K_{43} = 22600$ | Estimate |
| 44 | $V_{44}[\text{Pak1}^*]/(K_{44} + [\text{Pak1}^*])$ | $V_{44} = 270, K_{44} = 895$ | Estimate |

Table 1: Rate Reactions and the parameters being used. The first order and second order rate constants are given in s^{-1} and $nM^{-1}.s^{-1}$ respectively while Michaelis Constants V and K are given in $nM.s^{-1}$ and nM respectively.

| Reactant | Concentration (nM) | Ref |
|------------------|--------------------|----------|
| PI3K | 30 | [10] |
| PTEN | 0.1 | Estimate |
| PIP ₂ | 7000 | [10] |
| AKT | 200 | [2] |
| PDK1 | 1000 | Estimate |
| PDK2 | 1000 | Estimate |
| PP2A | 150 | [23] |
| NOX5 | 2000 | Estimate |
| PAK1 | 500 | [28] |
| ROS | 2000 | Estimate |
| Ras | 18900 | [43] |
| Raf | 66.4 | [43] |
| MEK | 36500 | [43] |
| ERK | 34900 | [43] |
| MKP3 | 44000 | Estimate |
| P90RSK | 5 | [5] |
| Bad | 100 | Estimate |
| Bax | 100 | [20] |
| Bcl2 | 100 | Estimate |

Table 2: Total concentration of the cellular components

the holdout technique [25] in model validation. However there still remains the problem of too few data sets to confidently assert the correctness of our model and its parameters. Tackling this problem of model validation with the least number of experimental data sets will be one of our directions in future research.

The role of the *Bcl-2* family needs to be further developed for the model to be more informative. In general there is a correlation between increase *AKT* and *ERK* activation and cell survival, but the actual mechanism that affects cell death is partly due to the interactions between members of the *Bcl-2* protein family, especially *Bad*, *Bax*, *Bcl-2* and *Bcl-XL*. These interactions depend not only on their physiological states (phosphorylated or unphosphorylated) but also in their physical locations in the cell. For example, signaling events that increase the likelihood of *Bax* being translocated from the cytoplasm to the mitochondria will have the higher chance of triggering apoptosis due to the release of *cytochrome c* from the mitochondria. In our model, there is still no clear indication of the rate of cell survival from the simulation results, other than the ratio of *Bcl-2* and *Bax* to serve as our guide. Experiments have shown that the *Bax* to *Bcl-2* ratio is a good indication of whether or not a cell will undergo apoptosis [29]. A *Bax/Bcl-2* ratio of more than 1 will imply that the cells are more susceptible to apoptosis while a ratio of less than 1 will imply that the cells are more resistant.

One of the main assumptions in the modeling process is that in an enzyme catalyzed chemical reaction, the concentration of the enzyme does not change with respect to that reaction. Although this does not necessarily happen in real life, it is only by applying such an assumption that we are able to use the Michaelis-Menten model to represent the chemical kinetics. As it turns out, such an assumption has a huge impact on the estimation of the missing parameters. For this model, much of the work is concentrated on parameter estimation. As this is a huge model to work with, to estimate the entire model in its entirety would be difficult, given the few data set to fit the model with. Hence, it would be desirable to break down the model into smaller modules. One of our guidelines in the determination of modules is that there should be no concentration flow in, or out of the modules. The enzymatic assumption stated above plays an important role in deciding where the boundaries of the modules are. Also, exploiting the fact that our model does not have any huge feedback loops, we are able to arrange the modules into a Directed Acyclic Graph, and estimate the parameters of the individual modules in topological order. By finding some form of dependency relationship, we can then say that the solution for a module depends only on the solution of its immediate parent modules.

Though not yet implemented in code, we do not have to execute the entire model in order to find out the fitness of the current parameters that are being estimated. Suppose the algorithm is currently handling rank 3 parameters, it only has to execute the modules that are ranked 3 and higher during model evaluation. Although not reported, test of this method on a smaller model - the *MAPK* model, results in consistently faster convergence of the solution, as compared to estimating the entire network. Although this has not been analyzed in depth, there have been instances of such algorithms out-performing the best optimization techniques for those instances [4].

We have seen how to use mathematical models to model a signaling network, and some of the issues underlying it, namely model derivation, parameter estimation and a bit on model validation. In future works, we plan to look at ways to model not just a network, but a class of networks that can sufficiently explain the results obtained from experiments. At minimum we should be able to assign confidence levels on the model based on the raw data provided.

References

- [1] I. Aksan and M. L. Kurnaz. A Computer-Based Model for the Regulation of Mitogen-Activated Protein Kinase (MAPK) Activation. *J. Recept. Signal Transduction*, 23(2-3):197–209, 2003.
- [2] S. F. Barnett, D. Defeo-Jones, S. Fu, P. J. Hancock, K. M. Haskell, R. E. Jones, J. A. Kahana, A. M. Kral, K. Leander, L. L. Lee, J. Malinowski, E. M. McAvoy, D. D. Nahas, R. G. Robinson, and H. E. Huber. Identification and Characterization of Pleckstrin Homology Domain Dependent and Isozyme Specific Akt Inhibitors. *Biochem J*, 385:399–408, 2005.
- [3] H.-G. Beyer and H.-P. Schewefel. Evolutionary strategies - A comprehensive introduction. In *Natural Computing*, pages 3–52. Springer-Verlag, 2002.
- [4] J. C. Bezdek and R. J. Hathaway. Some Notes on Alternating Optimization. In *Lecture Notes in Computer Science*, volume 2275, pages 289–300. Springer-Verlag, 2002.
- [5] R. R. Bhatt and J. E. F. Jr. Cloning and Characterization of Xenopus Rsk2, the Predominant p90 Rsk Isozyme in Oocytes and Eggs. *J Biol Chem*, 275(42):32983–32990, 2000.
- [6] R. M. Biondi, P. C. Cheung, A. Casamayor, M. Deak, R. A. Currie, and D. R. Alessi. Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA. *The EMBO Journal*, 19(5):979–988, 2000.
- [7] S. S. Brar, Z. Corbin, T. P. Kennedy, R. Hemendinger, L. Thornton, B. Bommarius, R. S. Arnold, A. R. Whorton, A. B. Sturrock, T. P. Huecksteadt, M. T. Quinn, K. Krenetsky, K. G. Ardie, J. D. Lambeth, and J. R. Hoidal. NOX5 NAD(P)H oxidase regulates growth and apoptosis in DU 145 prostate cancer cells. *Am J Physiol Cell Physiol*, 285(2):C353–C369, 2003.
- [8] F. A. Brightman and D. A. Fell. Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signaling in PC12 cells. *FEBS Letters*, 482(3):169–174, 2000.

- [9] A. Chaudhary, W. King, M. Mattaliano, J. Frost, B. Diaz, D. Morrison, M. Cobb, M. Marshall, and J. Brugge. Phosphatidylinositol 3-kinase regulates Raf1 through Pak phosphorylation of serine 338. *Current Biology*, 10(9):551–554, 2000.
- [10] D. Chodniewicz, A. M. Alteraifi, and D. V. Zhelev. Experimental Evidence for the Limiting Role of Enzymatic Reactions in Chemoattractant-induced Pseudopod Extension in Human Neutrophils. *J Biol Chem*, 279(23):24460–24466, 2004.
- [11] A. A. Deora, T. Win, B. Vanhaesebroeck, and H. M. Lander. A Redox-triggered Ras-Effector Interaction - Recruitment of Phosphatidylinositol 3-Kinase to Ras by Redox Stress. *J Biol Chem*, 273(45):29923–29928, 1998.
- [12] S. Eker, M. Knapp, K. Laderoute, P. Lincoln, , and C. Talcott. Pathway Logic: Executable Models of Biological Networks. In *4th International Workshop on Rewriting Logic and Its Applications (WRLA'2002)*, volume 71 of *Electronic Notes in Theoretical Computer Science*. Elsevier, 2002.
- [13] C. Farrar, C. R. Houser, and S. Clarke. Activation of the PI3K/Akt signal transduction pathway and increased levels of insulin receptor in protein repair deficient mice. *Aging Cell*, 4:1–12, Feb 2005.
- [14] M. Fussenegger, J. E. Bailey, and J. Varner. A mathematical model of caspase function in apoptosis. *Nature Biotechnology*, 18:768–774, July 2000.
- [15] H. Genrich, R. Küffner, and K. Voss. Executable Petri Net models for the analysis of metabolic pathways. *International Journal on Software Tools for Technology Transfer*, 3(4):394–404, 2001.
- [16] M. Hatakeyama, S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J.-H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama, and A. Konagaya. A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signaling. *Biochem J*, 373(2):451–463, July 2003.
- [17] J. Horoszewicz, S. Leong, T. Chu, Z. Wajsman, M. Friedman, L. Papsidero, U. Kim, L. Chai, S. Kakati, S. Arya, and A. Sandberg. The LNCaP cell line: a new model for studies on human prostatic carcinoma. *Prog Clin Biol Res*, 37:115–132, 1980.

- [18] F. Hua, M. G. Cornejo, M. H. Cardone, C. L. Stokes, and D. A. Lauffenburger. Effects of bcl-2 levels on fas signaling-induced caspase-3 activation: Molecular genetic tests of computational model predictions. *Journal of Immunology*, 175:985–995, 2005.
- [19] C. Y. F. Huang and J. E. F. Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci*, 93(19):10078–10083, Sept 1996.
- [20] J. M. Jürgensmeier, Z. Xie, Q. Deveraux, L. Ellerby, D. Bredesen, and J. C. Reed. Bax directly induces release of cytochrome c from isolated mitochondria. *Cell Biology*, 95(9):4997–5002, 1998.
- [21] Y. Kawakami, H. Nishimoto, J. Kitaura, M. Maede-Yamamoto, R. M. Kato, D. R. Littman, D. J. Rawlings, and T. Kawakami. Protein Kinase C β II Regulates Akt Phosphorylation on Ser-473 in an Cell Type and Stimulus specific Fashion. *J Biol Chem*, 279(46):47720–47725, 2004.
- [22] S. M. Keyse. Protein phosphatases and the regulation of mitogen-activated protein kinase signalling. *Current Opinion in Cell Biology*, 12(2):186–192, 2000.
- [23] S. Kikuchi, K. Fujimoto, N. Kitagawa, T. Fuchikawa, M. Abe, K. Oka, K. Takei, and M. Tomita. Kinetic simulation of signal transduction system in hippocampal longterm potentiation with dynamic modeling of protein phosphatase 2A. *Neural Networks*, 16(9):1389–1398, 2003.
- [24] Y. Kim, A. Rice, and J. Denu. Intramolecular dephosphorylation of ERK by MKP3. *Biochemistry*, 42(51):15197–15207, 2003.
- [25] R. Kohavi. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *Proc 14th Int Joint Conference on Artificial Intelligence*, pages 1137–1143, 1995.
- [26] W. Kolch. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem J*, 351:289–305, 2000.
- [27] E. Lee, A. Salic, R. Krüger, R. Heinrich, and M. W. Kirschner. The Roles of APC and Axin Derived from Experimental and Theoretical Analysis of the Wnt Pathway. *PLoS Biology*, 1(1):116–132, Oct 2003.

- [28] T.-H. Loo, Y.-W. Ng, L. Lim, and E. Manser. GIT1 Activates p21-Activated Kinase through a Mechanism Independent of p21 Binding. *Molecular and Cell Biology*, 24(9):3849–3859, May 2004.
- [29] T. Mackey, A. Borkowski, P. Amin, S. Jacobs, and N. Kyprianou. Bcl-2/Bax ratio as a predictive marker for therapeutic response to radiotherapy in patients with prostate cancer. *Urology*, 52(6):1085–1090, 1998.
- [30] H. Matsuno, A. Doi, M. Nagasaki, and S. Miyano. Hybrid Petri Net Representation of Gene Regulatory Network. *Pac Symp Biocomput*, pages 341–352, 2000.
- [31] H. H. McAdams and L. Shapiro. Circuit Simulation of Genetic Networks. *Science*, 269(5224):650–656, 1995.
- [32] B. Mishra and A. Policriti. Systems biology and automata. In *Proceedings of the 3rd Workshop on Computation of Biochemical Pathways and Genetic Networks*, Oct 2003.
- [33] C. G. Moles, P. Mendes, and J. R. Banga. Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods. *Genome Research*, 13(11):2467–2474, Nov 2003.
- [34] A. Nesterov, X. Lu, G. J. Miller, Y. Ivashchenko, and A. S. Kraft. Elevated AKT Activity Protects the Prostate Cancer Cell Line LNCaP from TRAIL-induced Apoptosis. *J Biol Chem*, 276(14):10767–10774, 2001.
- [35] K. M. Nicholson and N. G. Anderson. The protein kinase B/Akt signaling pathway in human malignancy. *Cellular Signaling*, 14(5):381–395, May 2000.
- [36] C. S. Park, I. C. Schneider, and J. M. Haugh. Kinetic Analysis of Platelet-derived Growth Factor Receptor/Phosphoinositide 3-Kinase/Akt Signaling in Fibroblasts. *J Biol Chem*, 278(39):37064–37073, 2003.
- [37] J. Pinney, D. Westhead, and G. McConkey. Petri Net representations in systems biology. *Biochemical Society Transactions*, 31(6):1513–1515, 2003.
- [38] Y. Pommier, O. Sordet, S. Antony, R. L. Hayward, and K. W. Kohn. Apoptosis defects and chemotherapy resistance: molecular interaction maps and networks. *Oncogene*, 23(16):2934–2949, 2004.

- [39] J. C. Reed. Double Identity for proteins of the Bcl-2 family. *Nature*, 387:773–776, 1997.
- [40] H. P. Reusch, S. Zimmermann, M. Schaefer, M. Paul, and K. Moelling. Regulation of Raf by Akt Controls Growth and Differentiation in Vascular Smooth Muscle Cells. *J Biol Chem*, 276(36):33630–33637, 2001.
- [41] T. Runarsson and X. Yao. Stochastic ranking for constrained evolutionary optimization. *IEEE Trans Evol Comput*, 4:284–294, 2000.
- [42] H. Sauer, B. Klimm, J. Hescheler, and M. Wartenberg. Activation of p90RSK and growth stimulation of multicellular tumor spheroids are dependent on reactive oxygen species generated after purinergic receptor stimulation by ATP. *FASEB J*, 15:2539–2541, 2001.
- [43] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, and G. Müller. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnology*, 20(4):370–375, 2002.
- [44] A. Schürmann and A.F. Mooney and L.C. Sanders and M.A. Sells and H.G. Wang and J.C. Reed and G.M. Bokoch. p21-Activated Kinase 1 Phosphorylates the Death Agonist Bad and Protects Cells from Apoptosis. *Mol Cell Biol*, 20(2):453–461, 2000.
- [45] S. V. Skiena. *The Algorithm Design Manual*. Springer, 1st edition, 1997.
- [46] J. W. Stucki and H.-U. Simon. Mathematical modeling of the regulation of caspase-3 activation and degradation. *Journal of Theoretical Biology*, 234:123–131, 2005.
- [47] Y. Tan, H. Ruan, M. R. Demeter, and M. J. Comb. p90RSK Blocks Bad-mediated Cell Death via a Protein Kinase C-dependent Pathway. *J Bio Chem*, 274(49):34859–34867, 1999.
- [48] F. Tsuruta, N. Masuyama, and Y. Gotoh. The Phosphatidylinositol 3-Kinase (PI3K)-Akt Pathway Suppresses Bax Translocation to Mitochondria. *J Biol Chem*, 277:14040–14047, 2002.
- [49] C. Vlahos, W. Matter, K. Hui, and R. Brown. A Specific Inhibitor of Phosphatidylinositol 3-kinase, 2-(4-morpholinyl)-8-phenyl-4H-1-benzopyran-4-one (LY294002). *J Bio Chem*, 269:5241–5248, 1994.

- [50] S. Yamada, T. Taketomi, and A. Yoshimura. Model analysis of difference between EGF pathway and FGF pathway. *Biochem Biophys Res Comm*, 314(4):1113–1120, Feb 2004.
- [51] E. Young, J. Zha, J. Jockel, L. H. Boise, C. B. Thompson, and S. J. Korsmeyer. Bad, a Heterodimeric Partner for Bcl-XL, and Bcl-2, Displaces Bax and Promotes Cell Death. *Cell*, 80(2), 1995.
- [52] S. Zimmermann and K. Moelling. Phosphorylation and Regulation of Raf by Akt (Protein Kinase B). *Science*, 286(5445):1741–1744, 1999.