A COMPUTATIONAL MODEL FOR THE METABOLISM OF *MYO*-INOSITOL HEXAKISPHOSPHATE

by

James Evan Smith

A thesis submitted in partial fulfillment of the requirements for the degree of

BACHELOR of SCIENCE in
BIOMEDICAL and HEALTH SCIENCES ENGINEERING

with HIGHEST HONORS

at

The UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

and NORTH CAROLINA STATE UNIVERSITY

2016

The UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

and NORTH CAROLINA STATE UNIVERSITY

We hereby approve the following thesis

A COMPUTATIONAL MODEL FOR THE METABOLISM
OF *MYO*-INOSITOL HEXAKISPHOSPHATE

by

James Evan Smith

ABSTRACT

This project formulates a computational model for the biochemical network interlacing inositol phosphates (InsPs) critical in the metabolism of *myo*-inositol hexakisphosphate ($InsP_6$). The role these InsPs are believed to play in developmental and homeostatic regulation underlines the need for novel analysis in model organisms like *Arabidopsis thaliana*. While some effects with apparent cause may be accessible by means of conventional analytics, a computational model for InsP synthesis could facilitate the resolution of far more complex queries and provide a comprehensive assessment of potential, even unintuitable, behavior. The primary goal of this project was to formulate an *in silico* model and to verify if, and to what extent, borrowed reaction kinetics for the InsP synthesis network recapitulate the behavior seen in traditional, *in vivo* studies. Preliminary results suggest that while our model's response to loss of function mutations emulates some aspects of the expected results, divergent phenomena indicate the need for deliberate modification. A subsidiary goal of this project was to demonstrate the utility of the model through its application in a variety of perturbation schemes. The model was able to, among other goal-oriented objectives, verify behavioral expectations in response to an indirect implementation of an energy-sensing kinase, SnRK1, and thereby corroborate contemporary understanding of certain protein-protein interactions without the need for expensive testing. Clearly, the flexibility and heuristic capacity of this model are of value and suggest that this project may very well represent the state of the art insofar as computational modeling of InsP signaling is concerned. In the future, iterative revision will only help to improve this model's efficacy and predictive power.

To my parents, who have always believed in me.

Thank you for your love and support.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Contributions made by members of the Gillaspy, Williams, Ducoste, and Perera labs are greatly appreciated.

This work was funded by the National Science Foundation (MCB #1052034).

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the view of the National Science Foundation.

LIST OF FIGURES

A COMPUTATIONAL MODEL FOR THE METABOLISM OF *MYO*-INOSITOL HEXAKISPHOSPHATE

Introduction

Myo-inositol phosphates (InsPs) are signaling molecules with diverse function that can be found in variety of organisms [1]. In animals, InsPs have been shown to demonstrate antineoplastic influence by means of mitigated proliferation and modulated cellular differentiation [2-3]. In yeast, the behavior of one particular InsP has been linked to mRNA export and chromatin remodeling [4]. In plants, the most prominent function of InsPs could be debated, but salient cases include roles in developmental transitions [5], homeostasis [6, 7], and secondary messaging—a process central to the continued transduction of extracellular stimuli [8]. These activities clearly illustrate the basic, if not translational, significance of InsPs in a broad contextual sense and underline the need for novel analysis in model systems like *Arabidopsis thaliana*.

In plants, the regulatory mechanisms supporting inositol signaling are of particular interest. Unfortunately, very little is known about the gene regulatory network (GRN) responsible for controlling the expression and activity of enzymes pertinent to InsP synthesis [1]. Network inference is convoluted, or otherwise complicated, by the stochastic quality of molecular interaction, the expense required to mitigate noise via the abstraction of high-throughput data, and the crosstalk that occurs between "levels" of interaction, where the lattermost element here describes edges that extend between nodes of the central dogma or detail catalysis. Phytogenetics, even in the kingdom's most studied organism, has not yet developed to the extent required to map interplay between relevant transcription factors and the sequences they affect. This knowledge would be of great utility in developing genetically modified organisms with characteristics optimized for a variety of applications. To achieve this end, however, would require that the scientific community engage in research that is traditionally hindered by costly experimentation—tests and assays that, without adequate support, can be largely immaterial. Consequently, most related work has instead focused, at least for the time being, on developing an understanding of the metabolic network InsPs comprise, which is a logical precursor to resolving "deep" connections [9].

Some current research questions concerning InsPs might appear fairly elementary to those unfamiliar with the subject. For instance, a legitimate question might ask, although there are existing theories, when and where in cellular dynamics is *myo*-inositol hexakisphosphate ($InsP_6$) synthesized? Other questions might merit more intricate consideration. For instance, how and why might $InsP_6$ be transported into vacuolar space (a form of storage), presuming its synthesis occurred in a cytosolic compartment? What effect could the energy status of a cell or a perturbed enzymatic balance have on said transport? Could a missing link in the metabolic network explain conceptual discrepancies in the response to loss of function mutations? What regulatory quotient might push the system toward a specific target state? These questions are just a few examples of the catechisms that this project hopes to influence.

While some effects with apparent cause may be accessible by means of conventional analytics, a computational model for InsP synthesis could facilitate the resolution of far more complex queries and provide a comprehensive assessment of potential, even unintuitable, behavior [10, 11]. The purpose of this research, then, is to employ a bottom-up approach to the resolution of hierarchical influence in InsP signaling by first synthesizing prevailing theoretical schemes and measurable parameters in an *in silico* environment before migrating into implicit dimensions or developing a sense of systemic behavior. In the simplest sense, this framework will predict metabolic dynamics as a function of enzyme concentrations. This operational link creates an opportunity for meta-heuristics—in this case, system-level insight that exceeds simply evaluating time-course kinetics. If this model demonstrates sufficient predictive power and has been agreeably verified, we may gain new intuition concerning the enzymatic balance needed to push the system's states toward some target and peer, as it were, into regulatory space. If in nothing else, this approach finds merit in that computational modeling tends to be relatively exhaustive and yet financially modest, so that any cost can be outweighed by substantial benefit.

There are a variety of mathematical paradigms that could be utilized in this situation to reconstruct subcellular interplay. The simplicity of flux balance analysis is attractive for systems with gaps in documented kinetics, but this structure lacks analytical flexibility in its constraints [12]. We turn, instead, to frameworks that boast resilience and familiarity in their use of ordinary differential equations, for which there are many well-documented approaches to systems analysis. Control theory and biochemical systems theory are robust schemes but rely on information that is either unknown or limiting in terms of the heuristic scope this project could have [13]. Mass action kinetics and Michaelian kinetics are differentiated by their popularity in biochemical modeling and the mechanistic significance of Michaelis constants *in vitro* [14].

Unfortunately, most kinetic models require a large number of empirically derived constants to accurately emulate *in vivo* behavior. Mass action models use "turnover" numbers or catalytic rate constants for each reaction to estimate dynamics through a linear sum. Michaelian models employ differential algebraic equations that rely on pairs of so-called "Michaelis constants." Together, these metrics describe substrate-enzyme affinity and reaction momentum. Notably, these mechanistic equations involve knowledge of the absolute concentration of a network's enzymes and metabolites. While the experiments used to quantify these characteristics can be complicated and expensive, in recent years the state of the art has witnessed a growing precedence for parameter estimation and data curation [15, 16]. With the advent of BRENDA and other online resources for information concerning biochemical reaction kinetics, computational models may instead rely on data abstracted from literature [17]. The primary goal of this project is to, first, formulate an *in silico* model of the biochemical network interlacing InsPs critical in the metabolism of $InsP_6$ and, then, to verify if, and to what extent, borrowed reaction kinetics recapitulate the behavior seen in traditional *in vivo* studies.

Background

Chemically, InsPs differ according to the number and position of phosphate groups bound to a *myo*-inositol ring, which is the most common stereoisomer of cyclohexanehexol. Together, these characteristics (i.e., number and position) form the basis for a chemical language [9]. Figure 1 illustrates that $InsP_6$ can be synthesized by phosphorylating the second position of $Ins(1,3,4,5,6)P_5$, a specific constitutional isomer of $InsP_5$. Note that the rate of each kinase reaction is dependent on the state of cellular energetics, since ATP is called upon to donate a phosphate group.

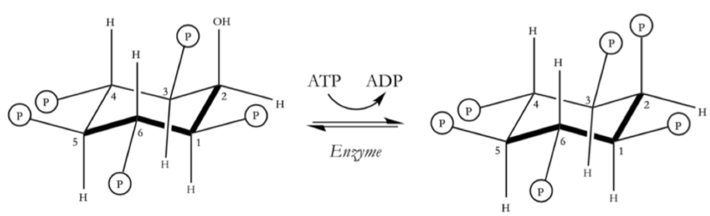

Figure 1

Phosphorylation of C2 on $Ins(1,3,4,5,6)P_5$

Figure 2 depicts the section of the InsP synthesis network that we have chosen to include in our pilot model. The selected window represents each path leading to the production and storage of $InsP_6$, beginning with either lower InsPs or inositol pyrophosphates (PPx-InsPs). To be clear, PPx-InsPs are similar to InsPs but instead carry two phosphate groups at one (or more) position(s). Our system includes six explicit reactions, three implicit "reactions" or channels for material flux, and one avenue for "transportation," or transmembrane migration. Despite an effort to account for every salient reaction between the metabolites that appear, there may be additional interactions that have not been described. For instance, while a single "reaction" has been employed to describe material influx (a source of mass) at the top of Figure 2, Ins(1,3,4,5,6)P_5 is actually synthesized via one of two evolutionarily discrete channels, termed the lipid-dependent and lipid-independent pathways [18]. The former sequence is shared with other organisms (e.g., animals and yeast) while the latter is unique to plants [1]. That the model can readily integrate "new" or missing reactions on the basis of biochemical theory is a testament to its flexibility. The effect of this capacity will be demonstrated in rectifying a discrepancy between theoretical expectations and experimental observations.

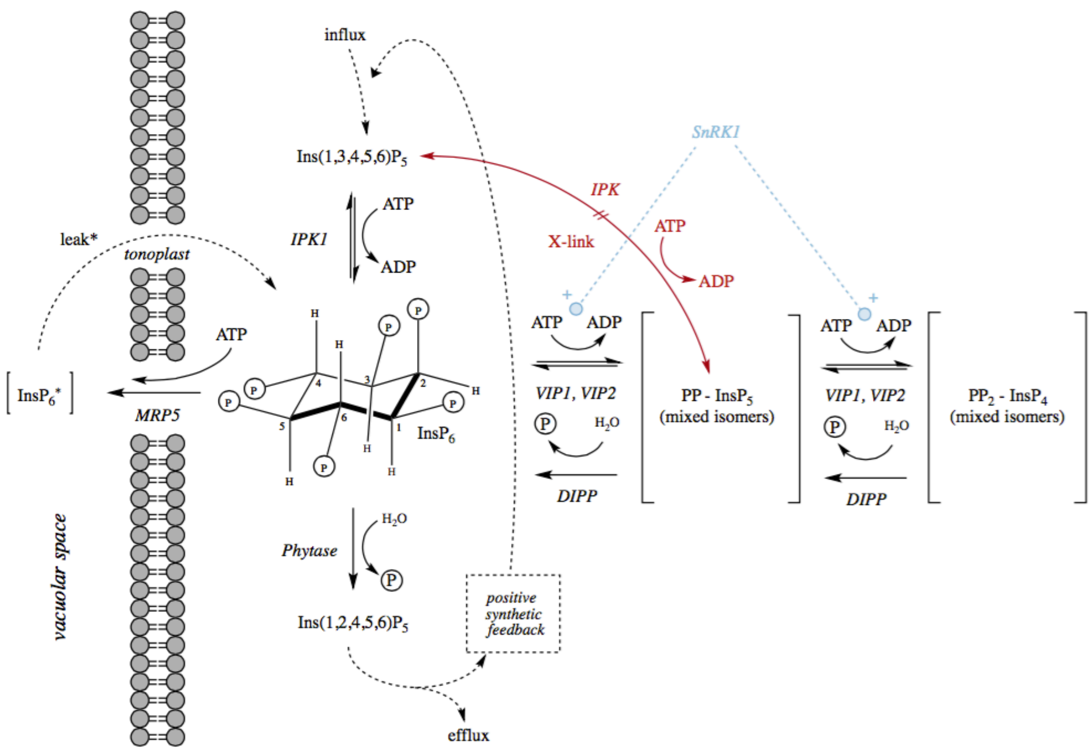

Figure 2

The InsP synthesis network

Once Ins(1,3,4,5,6)P_5 has been synthesized, irrespective of its origin, a specific kinase (IPK1) may readily phosphorylate its second carbon position, as depicted in Figure 1. In seeds and vegetative tissue alike, this reaction results in the accumulation of $InsP_6$ in subcellular space. At this juncture, mass will flow in one of five "directions," excluding the course just described.

That is, mass may:

Option 1. Shift back toward Ins(1,3,4,5,6)P_5 if the reaction quotient has outstripped the concomitant equilibrium constant,

Option 2. Migrate into vacuolar space (storage) via an ABC transporter, MRP5,
Note: this action does not alter chemical identity.

Option 3. Shift (a) to **or** (b) from the PPx-InsP branch via VIP1/2 **or** DIPP, respectively, or

Option 4. Shift "forward" into degradation, wherein Phytase is purportedly responsible for facilitating a reclamation process or "positive synthetic feedback."

Option 2 is of particular interest in biochemical research. As noted in the introduction, InsPs are biologically diverse in function but can be thought of, in the simplest sense, as being molecular storage units for phosphate groups. In the context of buffering the bioavailability of phosphate, the ability to bind any number of these key substituents is of great importance and is only augmented by creating sequestered pools of $InsP_6$ in vacuolar space. In theory, plants could retain this particular polyphosphate—$InsP_6$—in nutrient-rich intervals, only to mobilize these deposits as incident nutrients are drained. In this arrangement, transport to and from vacuolar space would be a function of and reflect not only development, but also environmental conditions. There is evidence for similar phenomena in *Arabidopsis thaliana*, the model organism we hope to emulate [19].

A member of the ATP-binding cassette (ABC) superfamily termed multidrug resistance protein 5 (MRP5) has been identified as the most likely macromolecule responsible for the transmembrane migration of $InsP_6$. This is significant because this particular ABC transporter can only export mass from the intracellular compartment [20]. After $InsP_6$ has migrated through some imagined tonoplast, it may be integrated into one or another "globoid," a spherical inclusion in protein bodies that serves to store phytate and other nutrients. While the exact mechanism through which $InsP_6$ is released may not be understood and has only been theorized, nutrient discharge has been documented throughout seedling development [19]. With this information in mind, it is possible to ask informed questions. Specifically, what effect could the energy status of a cell or a perturbed enzymatic balance have on this transport?

It is important to understand that options 1, 3b, and 4 entail degradation or the removal of one or another phosphate group while options 2 (no change) and 3a (an addition) do not. Likewise, it is critical that we distinguish each reaction on the basis of its complexity. In an attempt to reduce the number of kinetic parameters required for each reaction, we have chosen to exclude kinetic metrics for water and inorganic phosphate from all reactions, irrespective of volatility, but will include ATP and ADP as important substrates tied to cellular energetics. Consequently, the hydrolysis reactions catalyzed by either Phytase or DIPP are necessarily "uni-uni" reactions, meaning these are reactions with a single substrate and a lone product. Fortunately, these reactions are, at least when catalyzed by Phytase and DIPP, inherently irreversible. This scheme is somewhat complicated by the reversible behavior IPK1 and the VIP kinases demonstrate, in that we classify their reactions as "bi-bi," despite the potential involvement, again, of water and inorganic phosphate. Finally, we consider the transmembrane migration facilitated by MRP5 to be, at least conceptually, a "bi-uni" reaction [14]. This is because while the concentration of ATP is critical in determining the activity of the pump, its function is purportedly irreversible, making the concentration and Michaelian kinetics of ADP inconsequential.

In mammals and in yeast, option 3a is achieved through the action of at most two enzyme classes: the IP6 kinases (IP6Ks) and/or the VIP kinase sequence described above. The genes for the former class do not appear in plants, establishing AtVIP1 and AtVIP2 as those enzymes most pertinent to this and the neighboring PPx-InsP reaction: from $InsP_7$ to $InsP_8$ [21]. The specific constitutional isomers of the PPx-InsP branch have not yet been definitively resolved in *Arabidopsis*. A comparison with phylogenetically proximal organisms might suggest that the phosphorylation of $InsP_6$ and $InsP_7$ is most likely to occur at the 1st or 5th position of the *myo*-inositol ring [22]. At any rate, this project will instead delineate any diphosphoinositol pentakisphosphate as $InsP_7$ and any bis-diphosphoinositol tetrakisphosphate as $InsP_8$.

In general, option 4 presents difficulty in that while Phytase may theoretically reduce $InsP_6$ to a simple polyol via intermediates that share an identity with substrates driving influx—thereby creating some form of positive synthetic feedback—the specific identity of those intermediates is largely unknown. Inasmuch, the influx, synthetic feedback, and efflux of the selected window must be accounted for indirectly and represent the three "implicit" reactions mentioned earlier. A recent article in the New Phytologist by Panzeri, et al. has suggested that in mutants with a defective ABC transporter, this loop becomes "futile" and negative feedback inhibition represses IPK1 and a handful of enzymes upstream [23]. Given this project's limited resources and the modest number of reactions it considers, addressing this theory directly is likely beyond scope of our pilot model but could hopefully be investigated with future revisions. It is worth noting that MRP5 mutants are also known to present decreased and increased concentrations of $InsP_6$ and $InsP_7$, respectively [24]. The mechanism for this is not yet fully understood, but it is hoped that this model might provide insight concerning what mechanism may be needed to rectify this and similar, unexpected observations.

For instance, a correlation between the proliferation of PPx-InsPs, low energy status, and the necessity of VIP genes in responding to low energy conditions has led to the supposition that AtVIP enzymes may function as metabolic regulators. One possible mechanistic link between PPx-

InsPs and elevated enzymatic activity could be the sucrose non-fermenting related kinase 1 (SnRK1) [unpublished]. This protein is a low-energy sensing kinase that, beneath a certain energetic threshold, forms a protein complex with VIP. This interaction results in a key phosphorylation that may increase VIP's activity and is believed to escalate the concentration of "higher" and pyrophosphorylated InsPs. We hope to verify these behavioral expectations in response to an indirect implementation of SnRK1 and thereby further corroborate contemporary understanding of this protein-protein interaction without the need for expensive testing.

In reviewing the *in vivo* response to various double VIP mutations, we noticed that the presence of $InsP_7$ was more or less unaffected by what should amount to pinching the source of mass for the PPx-InsP branch. Naturally, this discrepancy engendered a series of questions that culminated with the following query: could a missing link in the metabolic network we have proposed explain conceptual discrepancies in the response to loss of function mutations? More specifically, and on the basis of expert knowledge concerning this system, could a reaction between $Ins(1,3,4,5,6)P_5$ and $InsP_7$ more or less bypass the activity of VIP in its absence? This hypothetical reaction will herein be referred to as the system's "X-link."

Significance

Recall that Figure 1 illustrates the phosphorylation of $Ins(1,3,4,5,6)P_5$ to generate *myo*-inositol hexakisphosphate, or phytic acid. Decomposition of any InsP, by contrast, releases inorganic phosphate, a macronutrient, through hydrolysis. Systematic degradation of $InsP_6$ is, however, particularly favorable as phytic acid is considered an "antinutrient" [25]. This property of $InsP_6$ is a physical consequence of its affinity for metal ions as a highly charged molecule. Phytic acid can chelate (or bind) essential minerals like iron, zinc, magnesium, or manganese. This activity could lead to intriguing inorganic biochemical queries if, as discussed in the background section, $InsP_6$ is sequestered as a function of the cellular environmental or developmental stage and carries with it the aforementioned minerals. At any rate, chelation leads to nutrient deficiency because $InsP_6$ is, by itself, indigestible by non-ruminant animals [26]. In the context of agricultural ecosystems and nutritional balance, targeted decomposition of $InsP_6$ could therefore lead to a reduction in dietary interference in that these minerals would be more readily absorbed. Likewise, a deliberate shift in state away from phytic acid in genetically modified plants may decrease the need for supplementary inorganic phosphate in animal feed because endogenous bioavailability would necessarily increase. Finally, these effects may, in turn, reduce nutrient runoff in nearby aquatic ecosystems, provided that the inorganic phosphate released by hydrolysis is absorbed. The cascading consequence of the lattermost effect is of particular significance, as phosphate runoff is understood to cause eutrophication that threatens the survival of key endangered species and, at length, Earth's delicate biosphere [21, 27]. Despite the variety of potential translational applications for InsPs in general, our focus on plants and *Arabidopsis* specifically has limited the immediate significance of this project to applications that are likely agronomical in nature. If nothing else, this model will likely help to address some of the long-standing issues in basic science that have for years impeded our understanding of inositol phosphate signaling in plants.

State of the Art

Generating a resilient *in silico* model of complex biochemical dynamics using only broad conceptual notions and external references certainly presents a challenge. To date, many biochemical models with mathematical structures similar to ours have utilized in-house experiments to quantify key kinetic or system parameters. This process can be complicated and expensive. While the accuracy and applicability of pertinent information is certainly in direct proportion to success, measuring any and all potentially relevant parameters can be problematic from a practical perspective. Indeed, measuring the value of each parameter only to construct a model that indicates which metrics *may* need to be revised on the basis of verification seems a bit paradoxical. Why not first estimate the required kinetics and cross check sensitivity analysis with uncertainty metrics to determine which kinetic parameters are of relatively greater value and therefore worth resolving more precisely? The latter approach, here, relies on whether or not it is possible to synthesize a more or less functional system using information that has been synthesized from various sources.

In recent years there has been a growing precedence for parameter estimation and data curation [15, 16]. With the advent of BRENDA and other online resources for information concerning biochemical reaction kinetics, computational models may instead rely on data abstracted from literature [17]. These online tools simplify the otherwise impossibly laborious task of sifting through hundreds and maybe even thousands of papers to find a handful of key kinetic parameters. As the results section illustrates, the flexibility and heuristic capacity of this model are of value and suggest that this project may very well represent the state of the art insofar as computational modeling of InsP signaling is concerned, certainly in that its parameters have almost entirely been sourced from external assessments. In fact, to the best of our knowledge, there are no other computation models for InsP signaling that employ a mathematical framework of similar quality, particularly not any that consider this relatively unexplored set of reactions or have a capacity for system level meta-heuristics.

Research Goals

In light of the state of the art, this project aims to forgo undertaking measurements for all potentially relevant parameters by first synthesizing prevailing theoretical schemes and kinetic data from online resources. If the resulting model is used to cross check sensitivity analysis with uncertainty metrics, we will be able to determine which kinetic parameters are of relatively greater value in goal-oriented tasks and thereby create a shortlist of parameters that are worth resolving more precisely. This information can be used to optimize iterative improvements to the model by focusing primarily on revising data points with the most influence.

To appreciate its impact, it is important to position this work in the context of being preliminary to a larger cohort of related research. As a pilot model, this project need only produce a basis for future iterations that, having been informed by the returns this model generates, aim to improve the system's validation, predictive power, and resilience. Since this venture is highly interdisciplinary in nature, it may take time for future iterations to engender new developments as the result of continued collaboration.

For reference, a reasonable functional metric in this context may be simply whether or not the system's response to perturbation matches the expected phenomena in a qualitative sense. For instance, an acceptable outcome might be that, for one or another loss of function mutation, the system's states shift not only in the expected direction, but with a magnitude more or less comparable to experimental data.

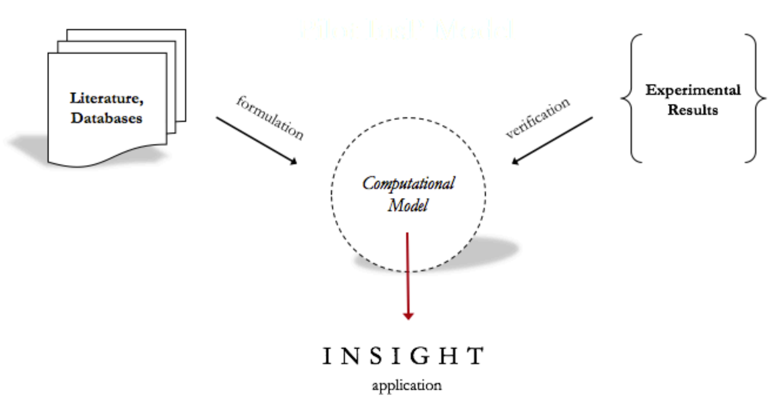

Figure 3

Diagram of interplay between key objectives

Figure 3 illustrates how the fundamental elements for this research project cooperate diagrammatically to achieve key objectives. By piecing together the appropriate substituent data and by corroborating early returns, we hope to assemble an *in silico* replica of the selected metabolic window capable of facilitating key applications with the hope that in doing so we may glean insight and new intuitions concerning system-level behavior in InsP signaling.

In summary, we hope to:

1. Formulate a computational model for InsP synthesis that will:
    a. Employ parametric data obtained from literature and various databases to determine if, and to what extent, borrowed reaction kinetics can recapitulate *in vivo* behavior
    b. Facilitate key system-level analytics such as objective-driven optimization and sensitivity analysis, among other more elementary tasks
2. Verify the pilot model by comparing the system's response to *in vivo* observations for knockout mutations in key enzymes
3. Demonstrate the utility of the model in various applications by:
    a. Predicting the regulatory balance concomitant to a low energy state
    b. Testing the response of the system to energetic modulation via SnRK1
    c. Investigating the model's behavior with a new reaction (the "X-link")

Methodology

There are a variety of mathematical paradigms that could be utilized in this situation to reconstruct subcellular interplay. Flux balance analysis is common [12]. Its structure only requires knowledge of stoichiometric coefficients for each reaction. This simplicity is attractive for networks with gaps in documented kinetics but sacrifices flexibility by assuming steady state conditions and optimality. In this framework, some objective function would describe a scenario wherein the relative amount of each metabolite is defined. The system would then return the (or, more problematically, a) flux distribution capable of maintaining the prescribed state. Subsequent variability analysis might elucidate interdependent flux domains from which the optimal solution can be obtained. Since analysis does not necessitate knowledge of kinetic parameters, the window of analysis can be extended greatly, so long as gaps in the network are avoided. Such gaps are, however, a shortcoming of research related to InsP signaling. Many reactions and their stoichiometric coefficients are only just now being theorized. If the window of considered reactions needs to be relatively limited to ensure accuracy, flux balance analysis would undoubtedly be outstripped by a more robust mathematical device.

Control theory, for example, may be used to explicate relationships beyond the scope of flux balance analysis by integrating information on feedback inhibition. More generally, biochemical systems theory (BST) constructs ordinary differential equations for metabolite concentrations $[M_i]$ as a function of system variables [13]:

$$\frac{d[M_i]}{dt} = \sum_j \mu_{ij} \cdot \lambda_j \prod_k [M]_k^{Q_{jk}}$$

—with stoichiometric coefficients μ, rate constants λ, and kinetic order Q. Much like flux balance analysis, however, BST relies on knowledge of stoichiometric coefficients and reaction order in a way that limits the window of considered reactions without providing a sufficient advantage over other methods that might warrant its use. One advantage of BST is the flexibility and familiarity of ordinary differential equations (ODEs). There are many well-documented approaches to systems analysis for ODEs. In contrast with Michaelian kinetics, BST does not utilize system parameters with measurable, mechanistic phenomenon *in vitro*. With this in mind, we turn to kinetic analysis.

In general, a kinetic system can be developed by assuming that the time derivative of any quantity X varies as the difference between the rate, v, of its production and depletion. That is:

$$\frac{dX}{dt} = \sum_{i} v_{production_X} - \sum_{i} v_{depletion_X}$$

Much in the same way, we have created a system of differential algebraic equations by considering this scheme for each metabolite within the InsP synthesis network—more specifically, the window delineated in Figure 2. For simple uni-uni, irreversible reactions like those catalyzed by Phytase and DIPP, we chose to employ the familiar Michaelis-Menten expression [14]. That is:

$$v = \frac{d[P]}{dt} = \frac{k_{cat_E}[E]_T[S]}{K_{M_{S,E}}+[S]}$$

Here, $K_{M_{S,\,E}}$ and k_{cat_E} represent the substrate-specific Michaelis constant and catalytic rate constant (or turn over number) for the general reaction S + E $\leftrightarrow$ ES $\rightarrow$ P + E, so that [S] and [P] represent the substrate and product concentrations, respectively, while E denotes the enzyme responsible for catalyzing the reaction. Since there may not be a one-to-one correspondence between active loci and protein bodies, the total concentration of catalytic sites is symbolized by $[E]_T$ to clarify this distinction. While rate equations are certainly simpler for uni-reactant systems, few reactions truly consist of a single substrate-product pair. For similar bi-substrate, <u>irreversible</u> reactions with the form A + B $\rightarrow$ P, we employ a related expression found in literature [14]:

$$v = \frac{k_{cat_E}[E]_T[A][B]}{[A][B]+K_{M_{A,E}}[B]+K_{M_{B,E}}[A]+K_{M_{A,E}}K_{M_{B,E}}}$$

With some inspection, it should be clear that this equation is simply an extension of the familiar Michaelis-Menten expression we employed for simple uni-uni reactions. Differences between the two forms have been highlighted in red to accentuate a sensitive trade off between the accuracy of mechanistic equations and their simplicity. To consider an additional substrate for an irreversible reaction, the expression must integrate knowledge of not only a new Michaelis constant, but also the absolute value for boundaries that describe the range of state for the concentration of the subsidiary substrate. The latter information is important for sampling initial conditions and gauging the viability of the model's response. The transmembrane migration of $InsP_6$ is the only network link that fits this description in the reactions this project considers. This is because while the concentration of ATP is critical in determining the activity of the ABC pump, its function is purportedly irreversible, making the concentration and Michaelian kinetics of ADP inconsequential. While this form does not, in the strictest sense, describe the mechanism of this transport, the selected expression should reasonably approximate the protein-substrate interactions required to ship phytic acid across some vacuolar tonoplast.

Given the complexity of explicit mechanistic expressions for fully reversible bi-bi reactions of the form A + B ↔ P + Q, we have opted to use a "simplified, generic" equation [28] developed by Rohwer, et al. Again, key differences between this and the prior form, insofar as parametric complexity is concerned, have been highlighted in red:

$$v_{net} = k_{cat_E}[E]_T \left[\frac{\frac{[A][B]}{K_{M_{A,E}} K_{M_{B,E}}} \left(1 - \left[\frac{[P][Q]}{K_{eq_n}[A][B]} \right] \right)}{\left(1 + \frac{[A]}{K_{M_{A,E}}} + \frac{[Q]}{K_{M_{Q,E}}} \right) \left(1 + \frac{[B]}{K_{M_{B,E}}} + \frac{[P]}{K_{M_{P,E}}} \right)} \right]$$

—where K_{eq_n} denotes the equilibrium constant for the reaction. In total, this expression requires knowledge of only seven additional values, two for each metabolite that together comprise the expected bounds for the substrate concentration in wild-type (WT) conditions. This is actually a favorable reduction as it eliminates the need for as many as twelve additional values. To be clear, this form was originally derived by Rohwer for its "relative simplicity in comparison to … mechanistic models." Rohwer believed that the modeling "process would … be aided if a generic equation was available that contained fewer parameters and yet described the kinetic behavior of [an] enzyme adequately." The value of his contribution, however, is not limited to the reduction of parametric data. Mechanistic equations are further complicated by differences in form that account for the order of substrate binding [14]. The primary mechanistic classifications include random, ordered-sequential, and ping-pong schemes that describe operations that fit nominal expectation. Since the mechanism for these reactions is unknown, that Rohwer's equation is not only "simple," but also "generic" is of great importance since it "fits data from both the ordered and … ping-pong models well." The exact mechanism is "irrelevant" to "the behavior of the pathway as a whole." In theory, the rate given by this expression matches the mechanistic rate within, at most, 3% of the "true" value for domains ranging well over 2 orders of magnitude [28].

The use of Michaelian kinetics is only further supported by its capacity to integrate explicit measures of inhibition and activation. While specific details concerning feedback are yet to be uncovered, these features can be integrated into the model by scaling the applicable Michaelis constant. That is, by scaling either a K_M value or $k_{cat}\,[E]_T = V_{max}$. For reference, modifications are feasible for competitive, non-competitive, and uncompetitive inhibition, respectively, where the latter-most modification alters both kinetic constants as follows [14]. Here, K_{M_i} has scaled its uninhibited counterpart K_{M_o} by some term with inhibitor I and its Michaelis constant K_i. Similar adjustments are possible for enzyme activation.

$$K_{M_i} = K_{M_o}\left(\frac{[I]}{K_i} + 1\right)$$

$$V_{max_i} = \frac{V_{max_o}}{\left(\frac{[I]}{K_i}+1\right)}$$

In theory, *Arabidopsis* should retain $InsP_6$ in nutrient-rich intervals only to mobilize these deposits as incident nutrients are drained. In this scenario, transport to and from vacuolar space is a function of development. For the sake of simplicity, this project will assume that phosphate accumulates in early development (that is, in seeds, if an organ-level distinction must be made) before being called upon throughout maturity (or in leaves). This is consistent with documentation of nutrient discharge throughout seedling development [19]. With this proposition, we may reasonably assign a leak from storage in leaves while constraining vacuolar capacity in seeds to the bounds of experimental observation.

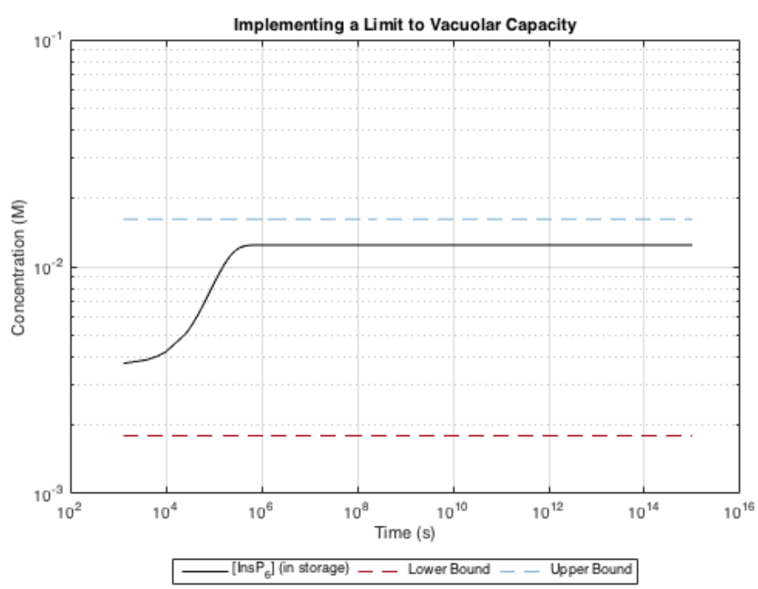

Figure 4

Limit to vacuolar capacity achieved by waning the activity of MRP5

This task is accomplished using a novel approach developed for this project. We begin by sampling a threshold for capacity between the nominal concentration of phytic acid in vacuolar space and half the distance to the upper bound of expectation. This sampling is repeated with each set of initial conditions to maintain some semblance of freedom in the terminal cap. From a practical perspective, once this threshold has been exceeded, the activity of MRP5 is diminished exponentially as a function of the distance between the rising concentration and said threshold. This modulation is made dynamically within the system of equations. Allegorically, this amounts to pinching some channel connecting intra- and extracellular space. Figure 4 illustrates the result.

While the mechanism that releases $InsP_6$ from storage may not be understood, a study by Bohn et al. has quantified the Michaelis constants for the degradation of $InsP_6$ found in isolated phytate globoids, the protein structures within which phytic acid is believed to nest while in storage. Additionally, Bohn identifies *myo*-inositol's third carbon position as being the target for catabolic activity in this context [29]. This suggests that, if this information is used to theorize some form of emission, the "product" of our leak should be $Ins(1,2,4,5,6)P_5$. Two caveats to this arrangement are that it necessarily assumes that the phytic acid in storage is found only in globoid deposits and that these complexes cannot be otherwise disintegrated.

Flux at the periphery of the selected window is accounted for using Hill functions with an adjustable parameter vector $\bar{\phi}$ and constants n [14]:

$$\frac{d[X_i]}{dt} = \frac{\phi_i^n}{\phi_i^n + [X_i]^n} \text{ OR } \frac{[Y_i]^n}{\phi_i^n + [Y_i]^n}$$

—where ϕ represents the "substrate" specific concentration concomitant to a reaction rate at half its maximum value. Here, X and Y are used to differentiate whether or not a rate is self-dependent. A limit of the first expression reveals that if the concentration of X is large, its rate of change approaches zero. The converse is also true. These characteristics are useful for modeling generic influx. Likewise, a limit of the second expression shows that if the concentration of Y is small, the rate of change for X approaches zero. Inasmuch, the latter expression is used to emulate not only a generic efflux, but also the positive synthetic feedback described in the background section on pages 6 and 8. This scheme was favored in deliberations for its flexibility, as the appropriate definitions for material flux can vary not only between seeds and other vegetative tissue, but also across development. In general, Hill functions are appropriate because the resulting rate of mass transfer may increase or decrease dynamically depending on what other strains the system is experiencing.

For these functions to stabilize the system in way that best emulates *in vivo* behavior, values for each constant ϕ must be resolved alongside some Hill constant n. Given that the purpose of these "reactions" is to prevent systemic mass from either proliferating or becoming depleted, a value of 3 was proposed for all Hill constants n on the basis that this balances the reactivity of these flux channels and is therefore best suited for this context. This was confirmed empirically. Meanwhile, the vector $\overline{\phi}$ was set using Monte Carlo runs that aimed to minimize the Euclidean distance between the present and wild-type steady state for select metabolites while maintaining a 9:1 ratio between synthetic feedback and material efflux. This proportion was favored in discussions relating this framework to biochemical experts. The algorithm responsible for pushing the system's states to a favorable location (i.e., 'phiset.m') would not be possible without empirical knowledge of the effects each flux channel has on the system. These characteristics were discovered by inspection. Figures 5 and 6 demonstrate the algorithm's ability to minimize this distance and maintain a 9:1 ratio between synthetic feedback and material efflux, respectively.

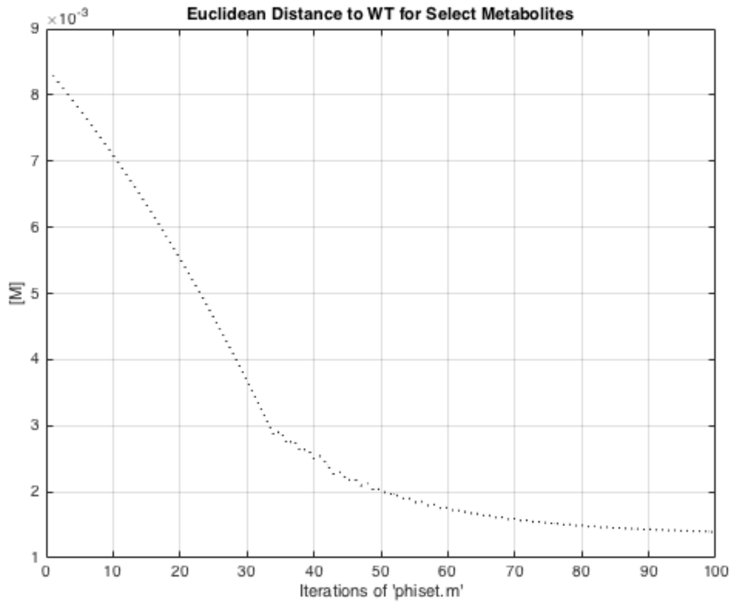

Figure 5

Resolving the vector $\overline{\phi}$

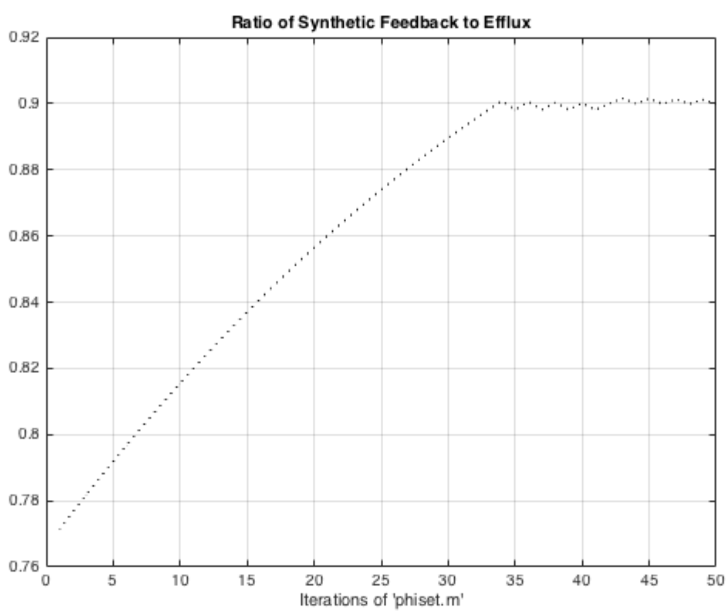

Figure 6

Maintaining a 9:1 ratio between synthetic feedback and material efflux

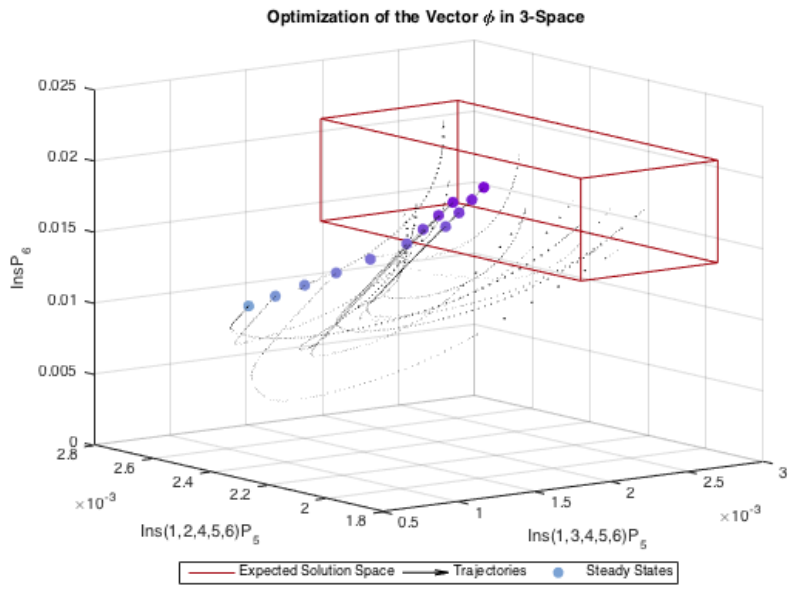

Figure 7

Resolving the vector $\overline{\phi}$

Figure 7 illustrates the effect of this approach in 3-space, where each circle denotes a steady state and each curve the trajectory from the system's initial condition. The red box encloses the expected space for solutions on the basis of probable boundaries for each concentration. Bounds were derived either from literature or our own measurements [30-38, unpublished].

Kinetic data for each of the six enzymes was obtained in a literature search expedited by online databases like BRENDA [20, 30-38]. Note that of the six, only two enzymes are currently being modeled using data from *Arabidopsis*, while the rest use parameters obtained using enzymes isolated from either other plants (1), yeast (1), or humans (2). Likewise, wild-type enzyme and metabolite levels were derived either from related literature or from our own measurements.

MATLAB was selected as the coding environment for this work due to its matrix-oriented manipulation of data. System parameters for seeds and leaves are indexed separately so that the user can quickly select the organ she would like to investigate. The steady-state solution space is determined with Monte Carlo runs that sample initial conditions within the bounds of expectation for each reactant and solve the system of equations over some interval. It is wise to include a safety factor here to avoid false returns in transience. The argument for this action must include a vector containing the concentration of each enzyme. In most cases, the WT profile should be provided. Redundant solutions are excluded on the basis of a tunable threshold for uniqueness. This approach to steady state determination may be supplemented with the Levenberg-Marqaurdt algorithm. Time-course data is then tabulated and plotted for some number of initial conditions. Solutions can be viewed with or without a variety of key markers. The stability of a solution is assessed by evaluating the sign of eigenvalues associated with the system's Jacobian matrix at that location. Loss of function mutations are modeled by executing the preceding tasks in quick succession, passing an increasingly modulated enzymatic profile with each iteration. An algorithm has been developed to automate this process and track the return.

Algorithms for System Analysis

Interest in predicting the regulatory balance concomitant to a particular steady state has led to the development of an optimization function. The resulting framework uses a "genetic algorithm" (GA) to assess the fitness of each solution and adjusts its operation accordingly [39]. Conceptually, a GA appeals to preeminent ideas in the theory of evolution to inform its mechanism for optimization. Much in the same way that key characteristics advantage certain individuals competing for, say, food and shelter, our GA assigns an elevated metric of "fitness" to enzyme profiles that correspond to states proximal to the target "location" in hyperspace. By using Euclidean distance as a metric for strength and by sifting through a batch of entries, the function can rank a set, or "generation," of profiles. The next generation is created by retaining elite individuals and by introducing variation through the stochastic combination of mutation and recombination. Mutation is accomplished by resampling random indices within a profile while recombination applies a complementary pair of binary masks to facilitate a random exchange between profiles in a process analogous to chromosomal crossover. If this technique is repeated for a sufficient number of generations, the function will likely reach an optimum. This process can be improved by adjusting the mutation rate and by increasing the number of individuals within the population.

To determine which kinetic parameters are of greatest value in goal-oriented tasks, the model needs to implement some form of sensitivity analysis. The Morris one-at-a-time (MOAT) method was selected for its flexible scope and efficiency [40]. Conceptually, the MOAT method calculates statistics for the so-called "elementary effects" of our parameter matrix. These elements are nothing more than a measure of the effect a small change in one input can have on the system's output. In the simplest sense, this function generates a rank-ordered list that indicates which parameters have the most influence in a particular context. Practically, this feat is accomplished by first reducing the range of k parameters to dimensionless intervals between zero and one. The resulting unit hypercube is then discretized along each dimension using a constant and predetermined interval. A trajectory is formed within this space by selecting a random initial condition and by changing one coordinate at a time using a premeditated step Δ_i k+1 times. Trajectories differ, then, according to both the initial condition and the order of modulation. For each step Δ_i , an elementary effect EE_i is calculated:

$$EE_i = \frac{[y(x+e_i\Delta_i)-y(x)]}{\Delta_i}$$

… with parameter vector x, output y, and unit vector e_i . The mean and standard deviation of elementary effects are then calculated for each parameter. The former is a straightforward measure of relative importance while the latter metric is said to reveal the prevalence of parametric interaction and divergence from linearity. This information can be used to optimize iterative improvements to the model by focusing primarily on revising data points with the most influence.

Results

A comparison of the model's response to key loss of function mutations with *in vivo* observations yielded promising markers for verification. In some cases, the model performed extraordinarily well by matching not only the direction of influence, but also its magnitude with accuracy well within an estimated range of uncertainty. Figure 8 depicts one such case for an IPK1 knockdown in seeds. Experimental data reports an 83% reduction of $InsP_6$ [20]. As is the case with most observations concerning loss of function mutations, this value is reported without an explicit metric for translational interference. If, however, the knockdown is presumed to affect about 90% of translation with, say, a conservative 10% estimate for uncertainty, then the model's prediction either intersects the biological measurement or nearly does so. In other cases, the model demonstrated a more qualitative capacity by tracking the direction of a response for one or another metabolite without strictly matching its magnitude.

Of course, this trend for verification was not without exception. Figure 9 details an interesting case in which the model predicts a steady decline in the concentration of both inositol pyrophosphates in response to a VIP1/2 double knockdown. This behavior is not in line with experimental observations that indicate that, in fact, little to no change occurs for these metabolites. Intriguingly, however, there is no intuitable reason to conclude that the PPx-InsP branch would be insensitive to a deliberate reduction of its only source. Could a missing link in the metabolic network we have proposed explain this conceptual discrepancy? More specifically, and on the basis of expert knowledge concerning this system, could a reaction between $Ins(1,3,4,5,6)P_5$ and $InsP_7$ more or less bypass the activity of VIP in its absence and stabilize the PPx-InsP branch? Figure 10 illustrates the effect of integrating this reaction (or, rather, set of reactions) using kinetic parameters for IPK, a likely candidate for this interaction should it exist in actuality. While the specifics for this X-link are, of course, almost entirely unknown and can only be theorized, the heuristic capacity of the model supports our suspicion by sustaining the stability of diphosphoinositol pentakisphosphate. That the model still predicts a steady decline in concentration for $InsP_8$ may be a shortcoming of the model to adequately capture material transfer without detailed kinetics or perhaps the consequence of yet another uncharted connection.

Much in the same way, we hope to investigate a correlation between the proliferation of PPx-InsPs and low energy status. One possible mechanistic link between the PPx-InsP branch and energetics could be SnRK1, a low-energy sensing kinase that phosphorylates VIP and augment's its activity [unpublished]. This cascade is believed to escalate the concentration of pyrophosphorylated InsPs. Indeed, Figure 11 depicts the expected phenomenon in response to an indirect implementation of SnRK1.

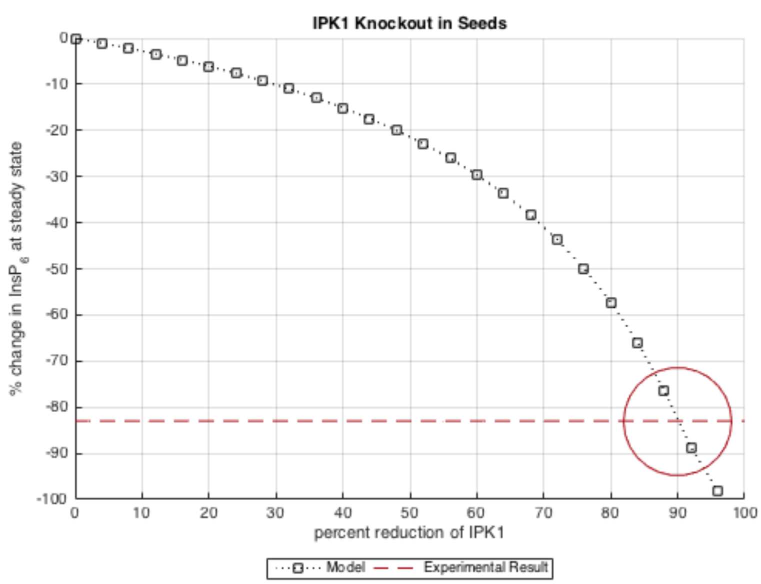

Figure 8

Intersection of in vivo and in silico behavior

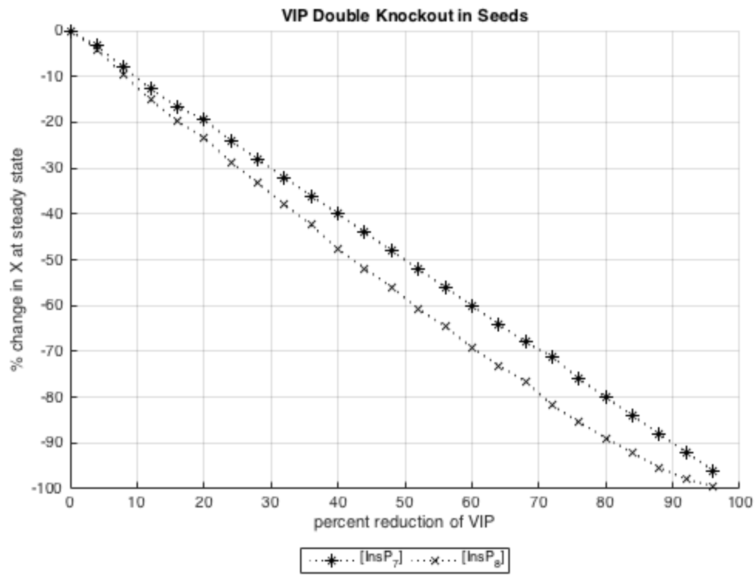

Figure 9

Divergent phenomena

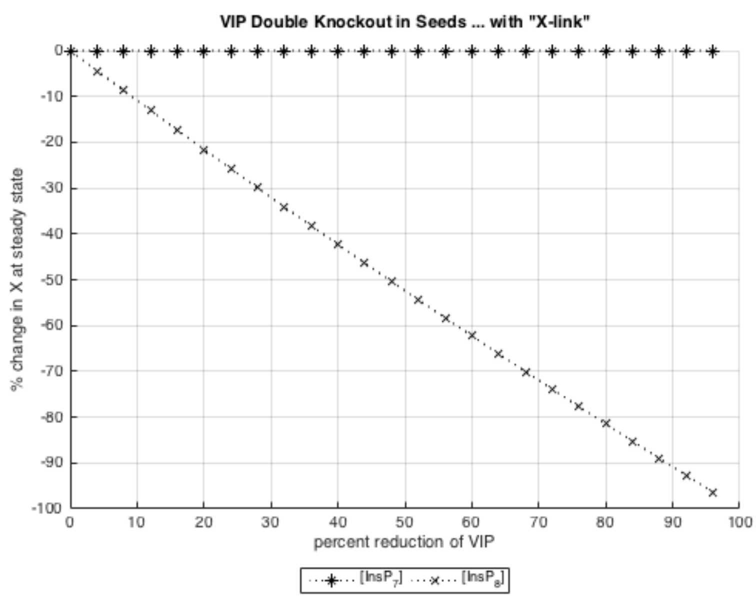

Figure 10

Impact of new reaction

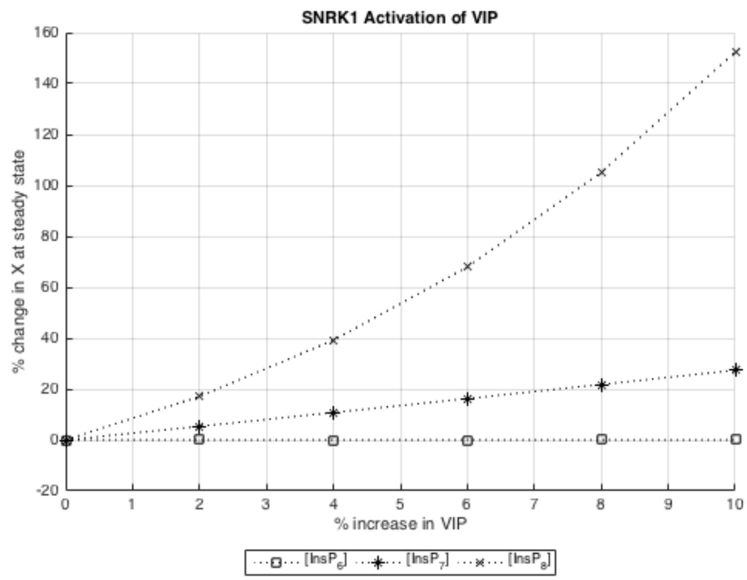

Figure 11

Indirect energy modulation

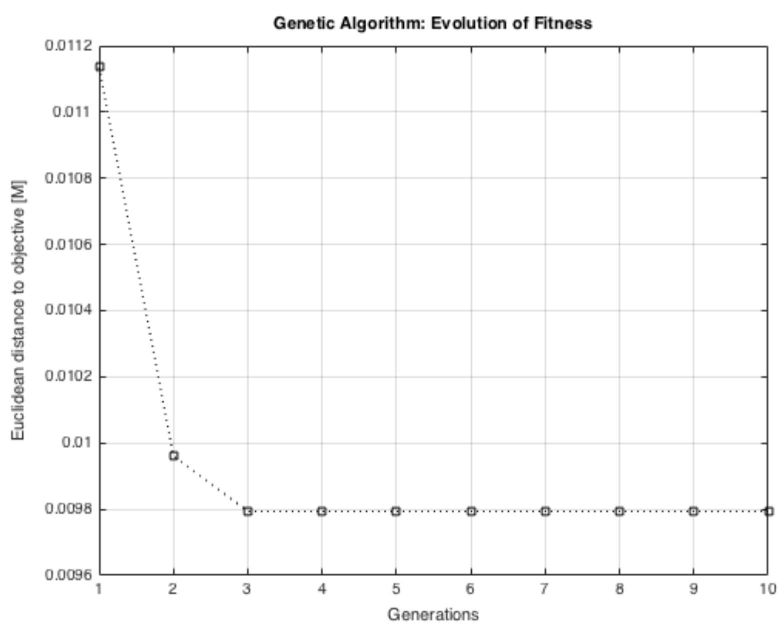

Figure 12

Tracking fitness of 20 individuals over 10 generations

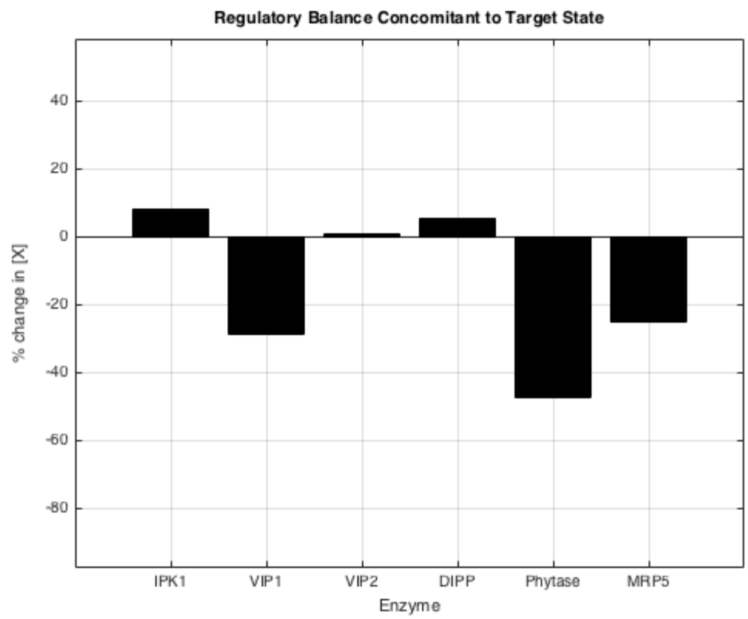

Figure 13

An optimum

While the activity of SnRK1 could be modeled and evaluated by mapping its mechanism to an increase in VIP's activity, it is often desirable to reverse this process and begin by specifying a target state. Since the low energy phenotype is of interest, we can use data on the relative abundance of InPs in plants 24 hours without sucrose to estimate an appropriate objective. Since this batch of data [unpublished] indicates relative abundance as a function of radioactive labeling, the time-course difference between the WT and low energy state must be used to extrapolate absolute values via percent difference. We ran optimization with the resulting target state using the genetic algorithm described above. Figure 12 illustrates how the "fitness" of 20 individuals improved over 10 generations. Figure 13 displays the corresponding optimum.

Discussion

Preliminary results suggest that while our model's response to loss of function mutations agreeably emulates some aspects of the expected results (Figure 8), divergent phenomena (Figure 9) indicate a need for deliberate revision. It is thought that the observed discrepancies result from:

1. Considering a limited or incomplete window of reactions,
2. Disparity in reaction kinetics for different organisms, and / or
3. Assuming a constant state for cellular energetics

Fortunately, each of these issues can be readily addressed using the framework this project has created. A key constraint to development has always been that the model be open to expansion. As the X-link application demonstrated, the flexibility to add new reactions to the existing model may very well be critical to reconciling differences between *in silico* and *in vivo* behavior and can even reveal novel connectivity, so long as new channels are added on the basis of robust biochemical theory and are implemented with caution (Figure 10).

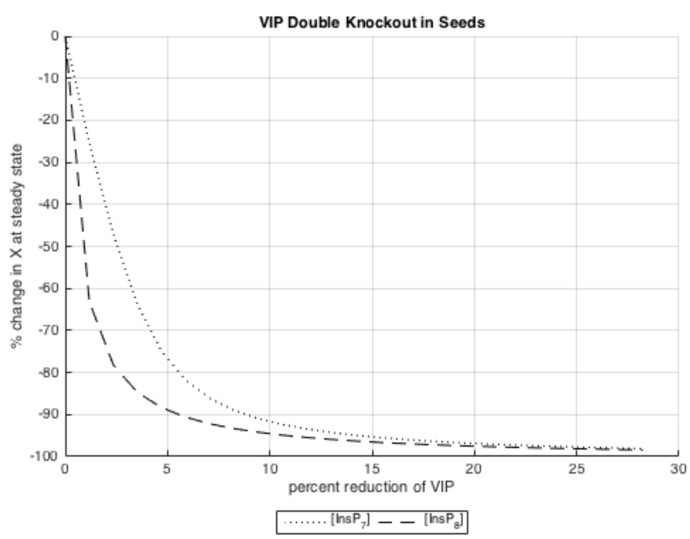

Figure 14

Parameter uncertainty as a key limitation

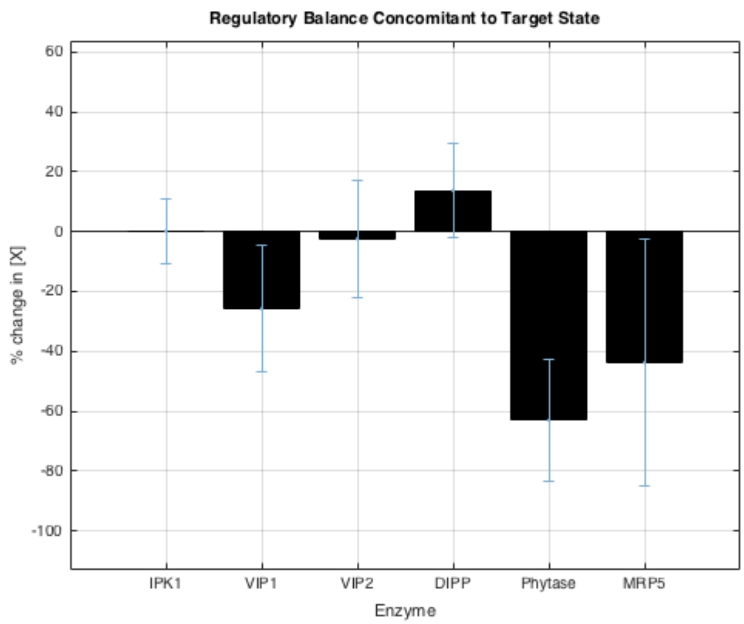

Figure 15

Organizing a multiplicity of local optimums

Since a computational approach to biological modeling is inherently indirect, there will inevitably be uncertainty propagation from input parameters to the results a system generates. This is particularly true for this model given that parameters were obtained from experiments that involve various, and often phylogenetically distant, organisms (e.g., yeast and humans). The influence of this uncertainty is best illustrated by considering the model's response to a VIP1/2 double knockdown before and after having revised a set of parameters critical to activity in the PPx-InsP branch (Figures 9 and 14). Clearly, this revision has a significant impact on the behavior we observe and most likely stems from variance furnished by a disparity in reaction kinetics for different organisms. Ideally, each parameter would be measured in the same environment using similar diagnostics by the same team within a fixed interval of time. Practically, this process would be expensive, time consuming, and prohibitive. Indeed, it is the daunting nature of this task itself that has led the state of the art to shift toward first estimating the required kinetics and cross checking sensitivity analysis with uncertainty metrics to determine which kinetic parameters are of value and therefore worth resolving more precisely. To this end, this project has synthesized prevailing theoretical schemes and measurable parameters into an *in silico* environment and customized a generic protocol for parameter screening so that this goal may be accomplished.

Of the various forms of utility this model provides, its optimization framework is likely the most attractive for its capacity to predict system-level organization and the regulatory balance concomitant to specific states. With that said, however, it important to understand that there may be a multiplicity of "optimal" profiles for a single target state, depending on how the algorithm is used and on the location of the objective. In the former case, if only a few "individuals" are allowed to evolve in the genetic algorithm over a handful of generations, the result will most likely be a local optimum. This limitation is analogous to searching for a function's critical points within a limited interval. As Figure 12 illustrates, a local optimum can be found after only a few generations. In the latter case, the model's inherent dimensional constraints may prohibit the system from reaching a global optimum, generating one or more possible locations that minimize the Euclidean distance between the system's state and its target without ever allowing the two to intersect. Singular value decomposition was used to confirm that although this system's state varies as a function of both its enzyme profile and the flux vector $\overline{\phi}$, this model is nevertheless constrained to a complex subspace.

The issue of multiplicity in this context can be approached in one of two ways. The first is to sacrifice computational cost for a deep search. Finding a global optimum may simply require the evolution of a large population over many generations. This process can be improved by adjusting the mutation rate. The second is to organize an array of optimums, found with the same argument, using basic statistics. Of course, the variance each enzyme has would be of particular interest. Figure 15 illustrates this technique for the same low-energy objective described for Figure 13 and organizes 30 optimums for 5 individuals over 10 generations.

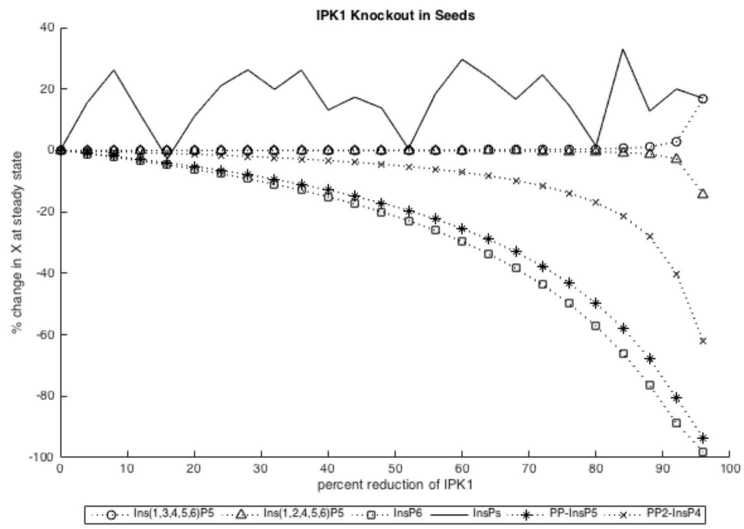

Figure 16

Vacuolar storage as a key limitation

Conclusion

As with any research project, this model has inspired a new set of questions for every answer it helps facilitate. One key question concerns the model's approach to emulating vacuolar storage of $InsP_6$. Since most translational applications and many basic research questions hinge on understanding the mechanism for and influence of having stored large quantities of phytate, we believe this feature should be emphasized as a focal point in future revision. Figure 16 illustrates deficiency in contrast with the success this model has had in emulating *in vivo* behavior outside of vacuolar space. The "noise" featured in this figure is the consequence of how the vesicle's cap is allowed to fluctuate with each initial condition. This measure ensures that the resulting steady states are realistic but renders this space more or less non-responsive to modulation of the enzymatic profile.

At any rate, iterative revision will help to improve the model's efficacy and predictive power. In the future, this model will likely help to address some of the long-standing issues that have for years impeded our understanding of inositol phosphate signaling. Ultimately, this construct may even serve as a tool in resolving certain features of the gene regulatory network that governs InsP and PPx-InsP activity. As the results section illustrates, the flexibility and heuristic capacity of this model are of value and suggest that this emulator may very well represent the state of the art insofar as computational modeling of InsP signaling is concerned, certainly in that its parameters have almost entirely been sourced from external assessments. In fact, to the best of our knowledge, there are no other computation models for InsP signaling that employ a mathematical framework of similar quality, particularly not any that consider this relatively unexplored set of reactions or have a capacity for system level meta-heuristics.

REFERENCES

[1] Gillaspy GE. The role of phosphoinositides and inositol phosphates in plant cell signaling. Adv Exp Med Biol. 2013; 991:141-57.

[2] Shamsuddin AM, Vucenik I, Cole KE. IP6: a novel anti-cancer agent. Life Sci. 1997; 61(4): 343-354.

[3] Shamsuddin AM. Inositol phosphates have novel anticancer function. J Nutr. 1995; 125(3): 725-732.

[4] Monserrate JP, York JD. Inositol phosphate synthesis and the nuclear processes they affect. Curr Opin Cell Biol. 2010; 22(3): 365–373.

[5] Gillaspy GE. Signaling and the polyphosphoinositide phosphatase. In: Munnik T, editor. Lipid signaling in plants. Berlin (Germany): Springer-Verlag.

[6] Boss WF, et al. Basal signaling regulates plant growth and development. Plant Physiol. 2010; 154(2): 439–443.

[7] Kuo HF, et al. Arabidopsis inositol pentakisphosphate 2-kinase, AtIPK1, is required for growth and modulates phosphate homeostasis at the transcriptional level. Plant J. 2014; 80(3): 503–15.

[8] Munnik T, Vermeer JE. Osmotic stress-induced phosphoinositide and inositol phosphate signaling in plants. Plant Cell Environ. 2010; 33(4): 655–669.

[9] Gillaspy GE. The cellular language of myo-inositol signaling. New Phytologist. 2011; 192(4): 823–839.

[10] Baghalian K, Hajirezaei MR, Schreiber F. Plant Metabolic Modeling: Achieving New Insight into Metabolism and Metabolic Engineering. Plant Cell. 2014; 26(10): 3847-3866.

[11] Wang JP, et al. Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in Populus trichocarpa. Plant Cell. 2014; 26(3): 894-914.

[12] Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Computational Biology. 2010; 28(3): 245-8.

[13] Savagea MA, Voit EO, and Irvine DH. Biochemical Systems Theory and Metabolic Control Theory: Fundamental Similarities and Differences. Mathematical Biosciences. 1987; 87: 127-145.

[14] Bisswanger H. Enzyme Kinetics: Principles and Methods. Weinheim: Wiley-VCH; 2002.

[15] Wittig U, et al. SABIO-RK: Integration and curation of reaction kinetics data. In: Leser U, Naumann F, Eckman B, editors. DILS 2006, LNBI 4075, 94-103.

[16] Beard DA. Simulation of Cellular Biochemical System Kinetics. Wiley Interdiscip Rev Syst Biol Med. 2011; 3(2): 136–146.

[17] Brenda: The Comprehensive Enzyme Information System [Internet]. BMFB and MWK; c2016. Available from http://brenda-enzymes.info

[18] Valluru R. Myo-inositol and beyond—emerging networks under stress. Plant Science. 2011; 181(4): 387-400.

[19] Beecroft, P, Lott, JNA. Changes in the Element Composition of Globoids From Cucurbita maxima and Cucurbita andreana Cotyledons During Early Seedling Growth. Canadian Journal of Botany. (1996); 74 (6): 838–847.

[20] Nagy R, et al. The *Arabidopsis* ATP-binding cassette protein AtMRP5/AtABCC5 is a high affinity inositol hexakisphosphate transporter involved in guard cell signaling and phosphate storage. The Journal of Biological Chemistry. 2009; 284(48): 33614-33622.

[21] Abelson PH. A potential phosphate crisis. Science. 1999; 283(5410): 2015.

[22] Thomas MP, Potter BVL. The enzymes of human diphosphoinositol polyphosphate metabolism. The FEBS Journal. 2014; 281: 14-33.

[23] Panzeri, et al. A defective ABC transporter of the MRP family, responsible for the bean lpa1 mutation, affects the regulation of the phytic acid pathway, reduces seed myo-inositol and alters ABA sensitivity. 2011; 191(1): 70-83.

[24] Desai M, et al. Two inositol hexakisphosphate kinases drive inositol pyrophosphate synthesis in plants. Plant J. 2014; 80(4): 642–653.

[25] Stevenson-Paulik J, et al. Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. Proc Natl Acad Sci. 2005; 102(35): 12612–12617.

[26] Cowieson AJ, Acamovic T, Bedford MR. Phytic acid and phytase: implications for protein utilization by poultry. Poult Sci. 2006; 85(5): 878–885.

[27] Brinch-Pedersen H. Engineering crop plants: getting a handle on phosphate. Trends in Plant Science. 2002; 7(3): 118–125.

[28] Rohwer JM, et al. Evaluation of a simplified generic bi-substrate rate equation for computation systems biology. IEE Proc Syst Biol. 2006; 153(5): 338–341.

[29] Bohn L, et al. Quantitative Analysis of Phytate Globoids Isolated from Wheat Bran and Characterization of Their Sequential Dephosphorylation by Wheat Phytase. Journal of Agricultural and Food Chemistry. 2007; 55: 7547-7552.

[30] Phillippy BQ, Ullah AHJ, Ehrlich KC. Purification and some properties of inositol 1,3,4,5,6-pentakisphosphate 2-kinase from immature soybean seeds. The Journal of Biological Chemistry. 1994; 296(45): 28393-28399.

[31] Sweetman D. Characterization of an Arabidopsis inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). Biochem J. 2006; 394: 95-103.

[32] Sun Y. Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize: molecular and biochemical characterization. Plant Physiology. 2007; 144: 1278-1291.

[33] Fridy PC. Cloning and characterization of two human VIP1-like inositol hexakisphosphate and diphosphoinositol pentakisphosphate kinases. The Journal of Biological Chemistry. 2007; 282(42): 30754-30762.

[34] Weaver JD, Wang H, and Shears SB. The kinetic properties of a human PPIP5K reveal that its kinase activities are protected against the consequences of a deteriorating cellular bioenergetics environment. Biosci Rep. 2013; 33(2): 229-241.

[35] Choi JH. Purification, sequencing, and molecular identification of a mammalian PPIP5 kinase that is activated when cells are exposed to hyperosmotic stress. The Journal of Biological Chemistry. 2007; 282(42): 30763-30775.

[36] Huang CF. Identification and purification of diphosphoinositol pentakisphosphate kinase, which synthesizes the inositol pyrophosphate bis(diphospho)inositol tetrakisphosphate. Biochem J. 1998; 37: 14998-15004.

[37] Phillippy BQ. Purification and catalytic properties of a phytase from scallion leaves. J Agric Food Chem. 1998; 46: 3491-3496.

[38] Wundenberg T, Mayr GW. Synthesis and biological actions of diphosphoinositol phosphates (inositol pyrophosphates), regulators of cell homeostasis. Bio Chem. 2012; 393: 979-998.

[39] Dorronsoro B, Alba E. Cellular Genetic Algorithms. New York: Springer; 2008.

[40] Morris MD. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics. 1991; 33(2): 161-174.