# MODELING AND ANALYSIS OF THE DIFFERENTIATION PROGRAM OF A MODEL ADULT HUMAN STEM CELL

R. Tasseff[1], S. Nayak[1], D. Kudela[2], J. K. Siddiqui[1] J. Wang[3], M. Shen[3], Andrew Yen[3], Jeffrey D. Varner[1]
[1]School of Chemical and Biomolecular Engineering, Cornell University, [2]School of Chemistry, Cornell University, [3]Department of Biomedical Sciences, Cornell University

Corresponding author: J. Varner, School of Chemical and Biomolecular Engineering, Cornell University 120 Olin Hall, Ithaca, NY 14853 USA jdv27@cornell.edu

**Abstract.** Understanding the molecular basis of stem cell differentiation programs is one of the grand unmet challenges facing modern cell-biology. In this study, we integrated computational and experimental network analysis tools to unravel the response of HL-60 human myeloblastic leukemia cells to Retinoic Acid (RA). HL-60 is an uncommitted precursor cell-line that is responsive to RA signals. Thus, HL-60 is an archetype model to study the molecular architecture of human differentiation programs. Our initial studies have focused on the role of the BLR1 receptor in the transduction of RA signals. BLR1, a G-protein coupled receptor expressed following RA exposure, is required for RA-induced cell-cycle arrest and differentiation and leads to atypical persistent MAPK signaling. A dynamic mathematical model of the molecular programs governing RA induced cell-cycle arrest and differentiation was formulated and tested against BLR1 wild-type (wt), knock-out and knock-in HL-60 cell-lines with and without RA. The current HL-60 model architecture described the dynamics of 729 proteins and protein complexes interconnected by 1356 interactions. An ensemble strategy was used to compensate for uncertain model parameters. Consistent with previous experimental studies, the initial HL-60 model showed up-regulation of BLR1 expression following RA exposure along with sustained MAPK activation. The initial simulation studies led to several testable structural linkages between BLR1 expression and MAPK activation. When taken together, our modeling efforts have established a prototype organization the differentiation program in HL-60. BLR1 acts as part of a feed-forward control element which drives its own expression and that of other components required for differentiation. More broadly, we have demonstrated that modeling of molecular programs can help prioritize experimental directions and expand current biological knowledge despite model uncertainty.

## 1 Introduction

Understanding the molecular basis of stem cell proliferation and differentiation programs is one of the grand unmet challenges facing molecular cell-biology. If these programs could be manipulated at a granular level, advanced stem cell therapies could be developed for the treatment of a spectrum of human cancers, spinal cord injuries and neurodegenerative disorders. However, the fine-tuned control of proliferation and differentiation programs in complex embryonic and adult stem cells is currently not possible.

In this study, we constructed and analyzed the response of HL-60 human myeloblastic leukemia cells to environmental stimuli (Figure 1). HL-60 is an archetype *in-vitro* model studied since the 1970's [7, 5, 23]. HL-60 remains a durable experimental model because it can undergo myeloid or monocytic differentiation and G0 arrest, making it one of the few immature precursor cell lines that is uncommitted with respect to differentiation lineage. Retinoic Acid (RA) or DMSO cause G0 arrest and myeloid differentiation in HL-60, whereas 1,25-dihydroxy vitamin D3 (D3) or sodium butyrate causes arrest with monocytic differentiation. The onset of G0 arrest and terminal differentiation is slow requiring $\sim$ 48 hr of treatment during which HL-60 cells undergo $\sim$ 2 division cycles [30, 26, 24, 27]. The 48 hr treatment period segregates into discernible halves. First, treatment with either RA or D3 causes a precommitment state where cells are primed to differentiate without lineage specificity. Growing cells in bromodeoxyuridine for one cell cycle also results in this precommitment priming [23]. Cells remain primed even as they continue to proliferate with RA, D3 or BUdR removed. Late in the priming period, RA (or D3) drives MEK-dependent activation of the ERK2/MAPK pathway [32, 33, 13, 22]. Elevated MAPK signaling persists until cells G0 arrest and terminally differentiate. If MEK is blocked, ERK2 fails to activate and HL-60 does not arrest or differentiate. Interestingly, expression of a highly mutated polyoma middle T antigen, $\Delta$205-214, also causes ERK2 activation and priming. After HL-60 has been primed, a second treatment with RA or D3 causes arrest and myeloid or monocytic differentiation, respectively. The kinetics of arrest and differentiation for a RA/D3 sequence is the same as continuous exposure to D3, and likewise G0 arrest and differentiation for cells treated with a D3/RA sequence is the same as cells continuously exposed to RA. Thus, priming by RA is lineage non-specific and equivalent to priming by D3. Activation of both RAR and RXR is necessary for RA induced G0 arrest, MAPK signal activation [25, 2, 29] and myeloid differentiation [33, 13, 22, 3].

RA-induced MAPK signaling, which is slow and persistent unlike the prototypical growth factor-induced mitogenic signal, is not well understood. During priming RA causes the transcriptional up-regulation of the BLR1 (Burkitt's Lymphoma Receptor-1) receptor. BLR1, also known as CXCR5, is a putative serpentine heterotrimeric

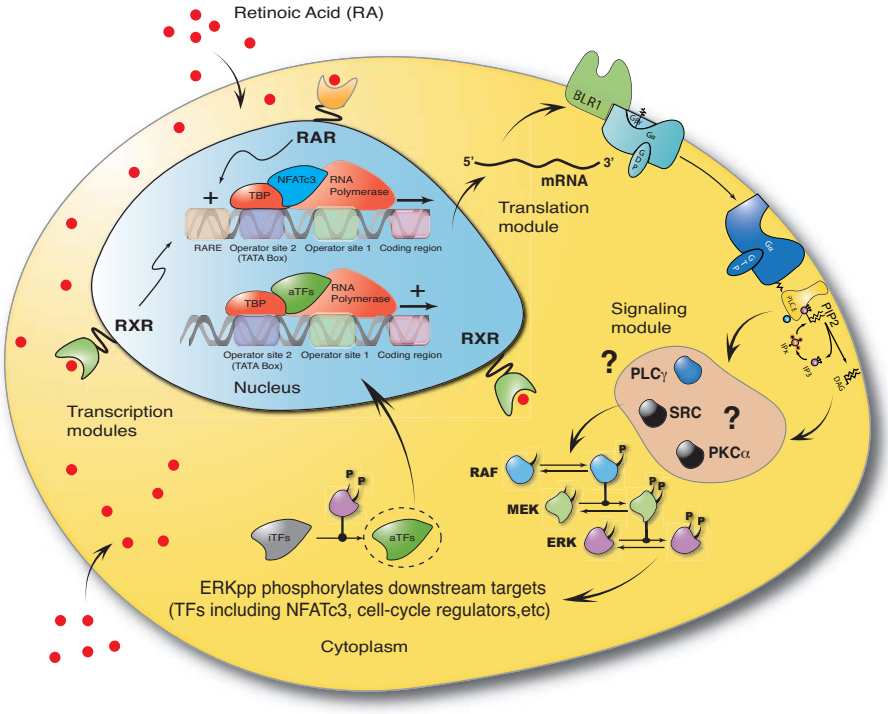


**Figure 1:** Schematic of the response of HL-60 to Retinoic Acid (RA). RA signals are intercepted by a family of RAR/RXR nuclear receptors which in turn drive the expression of genes with RARE promoter elements. One key RA-regulated protein is BLR1. BLR1 is a transmembrane surface receptor that is thought to be a G-protein coupled receptor. BLR1 expression, upregulated by RA, drives an atypical sustained MAPK signal which in turn activates the expression of genes required for the execution of the cell-cycle arrest and differentiation programs.

G-protein-coupled receptor, with a sequence similar to IL-8 receptors [9], that was first discovered in a screen for differentially expressed genes that conferred metastatic capability to human B-cell lymphomas [9, 10]. BLR1 expression in human was originally found restricted to mature resting B cells [9] and a subset of T-helper memory cells [11],suggesting initially that it was a lymphocyte-specific chemokine receptor family member. We identified BLR1 as an early RA (or D3)-inducible gene in HL-60 cells using differential display [15, 3], suggesting it had a broader function than lymphocyte regulation. Over expression of BLR1 in HL-60 cells enhanced ERK2 activation in both RA-untreated and treated cells and accelerated RA- and D3-induced differentiation and G0 arrest. Inhibiting ERK2 activation also inhibited the enhanced differentiation and arrest conferred by BLR1. Thus, RA-induced BLR1 expression appears to contribute to ERK2 activation and propulsion of induced differentiation and G0 arrest. Studies of the BLR1 promoter [20] identified a 5' segment approximately 1 kb from the transcriptional start that conferred RA responsiveness. Footprinting, EMSA and ChIP identified a novel 17 bp RARE consisting of two GT boxes which bound RAR and RXR in a ligand dependent fashion. This binding resulted in assembly of a complex consisting of Oct1, NFATc3 and CREB2 at this novel GT box RARE. The Oct1, NFATc3 and CREB2 bound their cognate sites 3' of the GT box RARE, but were not recruited into the complex at the GT-box RARE until addition of RA. The CREB, while necessary for assembly and transactivation, dissociates subsequently from the complex and its consensus sequence. Thus, the model is that RA causes RAR and RXR to bind the GT-boxes, and this recruits Oct1, NFATc3, and CREB2, which are tethered downstream at their consensus sequences, into the complex to form a transcriptionally active complex with subsequent CREB2 dissociation. Surprisingly two canonical RAREs found in proximity did not bind RAR or RXR and were dispensable for RA-induced transcriptional activation. RA thus directly transcriptionally activates BLR1, and the action of RA can now be followed from this receptor to downstream effects on signaling, transcription factor activation, gene expression, cell cycle arrest and differentiation.

The working hypothesis of our modeling studies is that the networks mediating the response of HL-60 to RA can be described as a System of Systems (SoS) i.e., a composite network composed of interacting subsystems. We have begun constructing and identifying a library of subnetwork models to be used for the construction of the HL-60 composite model. Every submodel in our library employs mass action kinetics to describe the rate of protein-protein and protein-DNA interactions. The mass action formulation, while expanding the dimension of the models, regularizes the mathematical structure, making it amenable to automatic generation of the model equations and analytical versions of the Jacobian and matrix of partial derivatives of the mass balances with respect to the model parameters (both of which are required for large-scale sensitivity analysis). Mass-action kinetics are also conceptually simple and reduce the unknown parameters to only three types, an association, dissociation or catalytic rate constant. Thus, we increase the dimension while simultaneously introducing mathematical structure

that ultimately reduces the complexity of model development and analysis. A second important distinction of the proposed work is that we do not have a single set of parameter values. Rather, library submodels have an ensemble of possible parameter sets selected to satisfy an error threshold between simulated and measured data. We expect that simulating over a parameter ensemble, when compared with a single best fit parameter set, will generate models (and predictions) more robust to parametric uncertainty. Our ensemble strategy was inspired by the previous work of Battogtokh *et al.* [4] and Sethna and coworkers [6].

# 2 Materials and Methods

**Formulation and solution of the model equations.** The HL-60 model was formulated as a set of coupled Ordinary Differential Equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{1}$$

The symbol $\mathbf{S}$ denotes the stoichiometric matrix ($729 \times 1356$). The quantity $\mathbf{x}$ denotes the concentration vector of proteins or protein complexes ($729 \times 1$). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates ($1356 \times 1$). Each row in $\mathbf{S}$ described a protein while each column described the stoichiometry of network interactions. Thus, the $(i, j)$ element of $\mathbf{S}$, denoted by $\sigma_{ij}$, described how protein $i$ was involved in rate $j$. If $\sigma_{ij} < 0$, then protein $i$ was consumed in $r_j$. Conversely, if $\sigma_{ij} > 0$, protein $i$ was produced by $r_j$. Lastly, if $\sigma_{ij} = 0$, there was no protein $i$ in rate $j$.

We assumed mass-action kinetics for each interaction in the network. The rate expression for protein-protein interaction or catalytic reaction $q$:

$$\sum_{j \in \{\mathbf{R}_q\}} \sigma_{jq} x_j \rightarrow \sum_{p \in \{\mathbf{P}_q\}} \sigma_{pq} x_p \tag{2}$$

was given by:

$$r_q(\mathbf{x}, k_q) = k_q \prod_{j \in \{\mathbf{R}_q\}} x_j^{-\sigma_{jq}} \tag{3}$$

The set $\{\mathbf{R}_q\}$ denotes reactants for reaction $q$. The quantity $\{\mathbf{P}_q\}$ denotes the set of products for reaction $q$. The $k_q$ term denotes the rate constant governing the qth interaction. Lastly, $\sigma_{jq}, \sigma_{pq}$ denote stoichiometric coefficients (elements of the matrix $\mathbf{S}$). We treated every interaction in the model as non-negative. All reversible interactions were split into two irreversible steps. The mass-action formulation, while expanding the dimension of the initiation model, regularized the mathematical structure. The regular structure allowed automatic generation of the model equations. In addition, an analytical Jacobian ($\mathbf{A}$) and matrix of partial derivatives of the mass balances with respect to the model parameters ($\mathbf{B}$) were also generated. Mass-action kinetics also regularized the model parameters. Unknown model parameters were one of only three types, association, dissociation or catalytic rate constants. Thus, although mass-action kinetics increased the number of parameters and species, they reduced the complexity of model analysis. In this study, we did not consider intracellular concentration gradients. However, we accounted for membrane, cytosolic and nuclear proteins by explicitly incorporating separate protein species.

**Estimation of an ensemble of model parameters.** An initial set of model parameters, $\mathbf{p}_0$ was chosen to replicate the training data taken from the literature [21]. The training data consisted of time-resolved MAPK and BLR1 western and northern blot data taken following RA addition to wild-type HL-60 cells. The difference between the training data associated with species $j$, $\hat{x}_{i,j}$, and simulation results associated with species $j$ for parameter set $k$, $x(\mathbf{p}_k)_{i,j}$, was quantified by the normalized mean squared error, $\eta$:

$$\eta = \frac{1}{N} \sum_{i,j} \frac{(\hat{x}_{i,j} - \beta_j x_{i,j})^2}{\hat{\sigma}_{i,j}^2}, \tag{4}$$

The sum was carried out over all species $j$ and observations $i$. The quantity $N$ is the total number of observations and $\hat{\sigma}$ denotes the corresponding experimental error. If no experimental error value was reported we assumed we assumed 10% of the reported observation. In cases where quantification of the stimulus or observation was unclear an augmented error of 20%-100% was applied to compensate for the added uncertainty. The scaling factor $\beta_j$ was applied to the simulation results to account for training data only known to a multiplicative constant. The value of $\beta_j$ was chosen to minimize the normalized squared error for a given experiment and species $j$ [6]:

$$\beta_j = \frac{\sum_i (\hat{x}_{i,j} x_{i,j} / \hat{\sigma}_{i,j}^2)}{\sum_i (x_{i,j} / \hat{\sigma}_{i,j})^2}. \tag{5}$$

Because of the scaling factor, the concentration units on simulation results was arbitrary (consistent with the arbitrary units associated with the training data). There was insufficient training data to properly constrain the 1356 model parameters. To account for parametric uncertainty an ensemble of parameter sets was generated using

a Monte Carlo approach similar to that of Battogtokh *et al.* [4]. Consider a set of model parameters $\mathbf{p}_i$. Let the likely-hood, $\phi(\mathbf{p}_i)$, of $\mathbf{p}_i$ be defined as:

$$\phi(\mathbf{p}_i) \equiv exp\{\frac{-\eta(\mathbf{p}_i)}{\alpha}\}, \tag{6}$$

where $\eta(\mathbf{p}_i)$ denotes the simulation error associated with parameter set $\mathbf{p}_i$. The quantity $\alpha$ is a parameter used to tune the rate of acceptance. Further let the acceptance probability, $P(\mathbf{p}_{i+1}\prime|\mathbf{p}_i)$, of a new parameter set, $\mathbf{p}_{i+1}\prime$, be $\frac{\phi(\mathbf{p}_{i+1}\prime)}{\phi(\mathbf{p}_i)}$ if $\phi(\mathbf{p}_{i+1}\prime) < \phi(\mathbf{p}_i)$ and 1 otherwise. $P$ corresponds to the probability that $\mathbf{p}_{i+1}\prime$ be accepted as $\mathbf{p}_{i+1}$ given $\mathbf{p}_i$, for consecutive Monte Carlo steps. Subsequent sets of parameters are generated by applying a small additive random perturbation in log space:

$$\log \mathbf{p}_{i+1}\prime = \log \mathbf{p}_i + \mathscr{N}(0,\nu) \tag{7}$$

where $\mathscr{N}(0,\nu)$ is a normally distributed random number with zero mean and variance $\nu$. The perturbation was applied in log space to account for the large variation between different parameters and to ensure that parameter values were always $> 0$.

Monte Carlo trajectories were generated starting from $\mathbf{p}_0$ where $\nu$ =0.05 or 0.1 and $\alpha$ =1 or 0.5. The autocorrelation function of each trajectory was calculated. The number of Monte Carlo steps between parameter sets which were added to the ensemble was taken to be the number of steps after which the autocorrelation function dropped to 95% of its initial value. This was done to ensure independence between sets in the ensemble. To compensate for noise in the autocorrelation function an exponential fit was applied. The final ensemble contained 870 parameter sets. Of the approximately 1500 parameters (1356 kinetic constants plus unspecified initial conditions), 13% (200 of 1500) had a Coefficient of Variation (CV) of 50% or less while 80% had a CV of 100% or less. Thus, the training data poorly constrained individual parameter values and emphasized the benefit of applying an ensemble approach.

**Sensitivity analysis of the HL-60 network.** Overall State Sensitivity Coefficients (OSSC) were used to estimate which structural elements of the HL-60 network were sensitive [18]. OSSC values were determined by first calculating the first-order sensitivity coefficients at time $t_k$:

$$s_{ij}(t_k) = \left.\frac{\partial x_i}{\partial p_j}\right|_{t_k} \tag{8}$$

First-order sensitivity coefficients were computed by solving the matrix differential equation:

$$\frac{d\mathbf{s}_j}{dt} = \mathbf{A}(t)\,\mathbf{s}_j + \mathbf{b}_j(t),\ \mathbf{s}_j(t_0) = \mathbf{0} \qquad j = 1,2,\ldots,P \tag{9}$$

In Eqn. 9, $j$ denotes the parameter index, $P$ denotes the number of parameters in the model, $\mathbf{A}$ denotes the Jacobian matrix, and $\mathbf{b}_j$ denotes the $j$th column of the matrix of first-derivatives of the mass balances with respect to the parameter values (denoted by $\mathbf{B}$). An analytical Jacobian and matrix of first-derivatives of the mass balances w.r.t the parameters:

$$\mathbf{A} = \left.\frac{\partial \mathbf{f}_\mathbf{x}}{\partial \mathbf{x}}\right|_{(\mathbf{x}^*,\mathbf{p}^*)} \qquad \mathbf{B} = \left.\frac{\partial \mathbf{f}_\mathbf{x}}{\partial \mathbf{p}}\right|_{(\mathbf{x}^*,\mathbf{p}^*)} \tag{10}$$

were generated from the model equations. The quantity $\mathbf{f}_\mathbf{x} = \mathbf{S}\cdot\mathbf{r}(\mathbf{x},\mathbf{p})$ and $(\mathbf{x}^*,\mathbf{p}^*)$ denotes a point along the unperturbed model solution. The sensitivity equations required that we solve the model equations to evaluate the $\mathbf{A}$ and $\mathbf{B}$ matrices. Thus, we formulated the sensitivity problem as an extended kinetic-sensitivity system of equations [8]:

$$\begin{pmatrix}\dot{\mathbf{x}}\\ \dot{\mathbf{s}}_j\end{pmatrix} = \begin{bmatrix}\mathbf{S}\cdot\mathbf{r}(\mathbf{x},\mathbf{p})\\ \mathbf{A}(t)\,\mathbf{s}_j + \mathbf{b}_j(t)\end{bmatrix} \qquad j = 1,2,\ldots,P \tag{11}$$

where $\dot{\mathbf{x}} = d\mathbf{x}/dt$ and $\dot{\mathbf{s}}_j = d\mathbf{s}_j/dt$. We solved the kinetic-sensitivity system for multiple parameters in a single calculation using the LSODE routine of OCTAVE (www.octave.org). The first-order sensitivity coefficients were then used to calculate the OSSC value for parameter $j$:

$$O_j(t) = \frac{p_j}{N_s}\left(\sum_{k=1}^{N_T}\sum_{i=1}^{N_s}\left[\frac{1}{x_i}\left.\frac{\partial x_i}{\partial p_j}\right|_{t_k}\right]^2\right)^{1/2} \tag{12}$$

The terms $N_T$, $N_s$ denote the number of time points considered and the state dimension of the model, respectively.

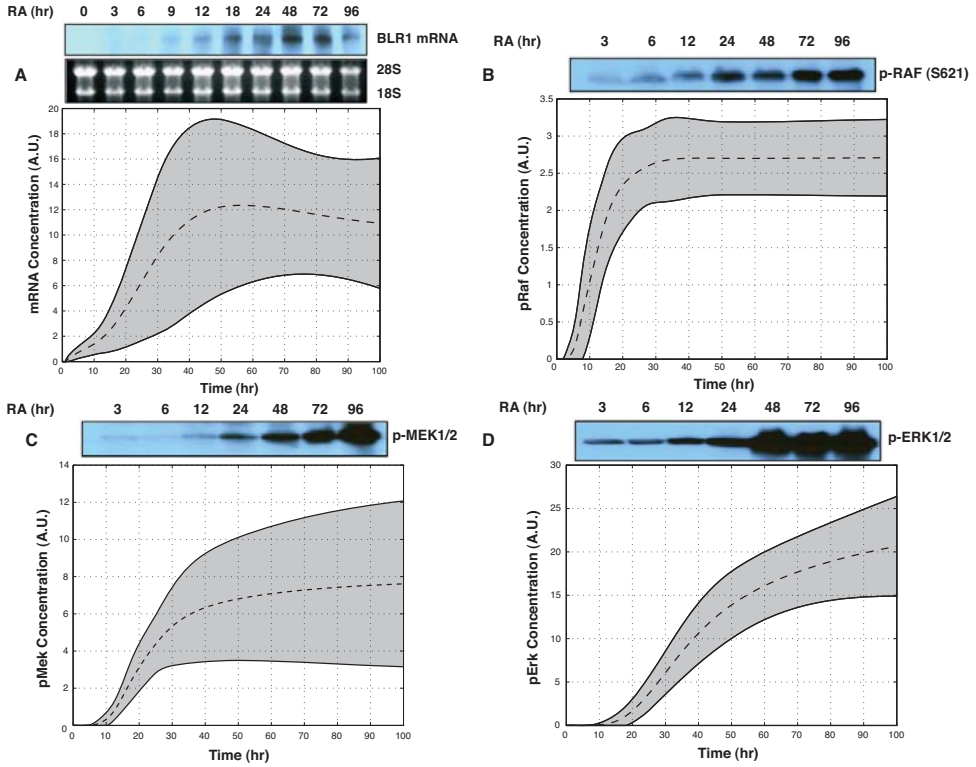


**Figure 2:** Simulations over an initial parameter ensemble (N = 876) captured the sustained activation of MAPK in HL-60 cells following RA exposure (1$\mu$M). In each case dashed lines denote the mean simulated value over the ensemble while the shaded regions denote one ensemble standard deviation.**A**: Comparison of experimental and simulated concentration profile for the BLR1 mRNA following RA exposure. No BLR1 expression was predicted before RA exposure. BLR1 expression was observed approximately 9-12 hr after RA exposure. **B**: Time profile of phosphorylated RAF1 activation following RA exposure. **C**: Simulated MEK activation following RA exposure. **D**: Simulated phosphorylated ERK following RA exposure. The northern and western blots were reproduced from Wang and Yen [21].

**Monte-carlo coupling analysis of the HL-60 architecture.** Coupling coefficients of the form:

$$\alpha\left(i,j,t_o,t_f\right) = \left(\int_{t_o}^{t_f} x_i\left(t\right)dt\right)^{-1}\left(\int_{t_o}^{t_f} x_i^{(j)}\left(t\right)dt\right) \tag{13}$$

were calculated to understand the regulatory connectedness of the HL-60 network. The coupling coefficient $\alpha\left(i,j,t_o,t_f\right)$ is the ratio of the integrated concentration of a network output in the presence (numerator) and absence (denominator) of structural or operational perturbation. Here $t_0$ and $t_f$ denote the initial and final simulation time respectively. $i$ and $j$ denote the indices for a reference species and a perturbed species respectively. If $\alpha\left(i,j,t_o,t_f\right) > 1$, then the perturbation *increases* the output concentration. Conversely, if $\alpha\left(i,j,t_o,t_f\right) \ll 1$ the perturbation *decreases* the output concentration. Lastly, if $\alpha\left(i,j,t_o,t_f\right) \sim 1$ the perturbation does not influence the output concentration.

# 3 Results and Discussion

To test the SoS hypothesis for the HL-60 network, we formulated a draft HL-60 model from literature and our identified subnetwork library. We tested the proof-of-concept model against BLR1 wild-type (wt), knock-out and knock-in cell-lines in the presence and absence of RA over an ensemble of parameters (N = 870) generated by comparing model simulations with western and northern blot data taken from the literature [21]. The initial HL-60 architecture consisted of a modified MAPK module [1, 12], G0/G1-cell-cycle module [16], a Calcium/Gq-protein cascade module and the translation initiation module [17]. The model describes the regulated expression of 17 marker proteins (BLR1, Oct1, CREB2, ETS, BRN, IRF, SRPK, pRb, RhoGDI, p47Phox, CD45, EIF2AK, SIIIp15, AAF, Stat5b, XRE and SAP). Regulated expression followed a two-operator site promoter template. Activated transcription factors were assumed to reversibly bind operator site 1 (O1) which then allowed the recruitment of RNApII to operator site 2 (O2). Transcription of BLR1 was modeled by explicitly describing the reversible assembly of the CREB2:Oct1:NFATc3 complex and dissociation of CREB2 to activate the complex at O1 and then the recruitment of RNApII at O2. The expression of the remaining un-regulated model proteins was fixed by the initial condition of the simulation. mRNA transcripts were transported from the nucleus assuming saturable transporters analogous to exportin-5. Once in the cytosol mRNA was translated using the translation module. Missing connectivity between pathway modules was estimated from a manual review of >100 publications and

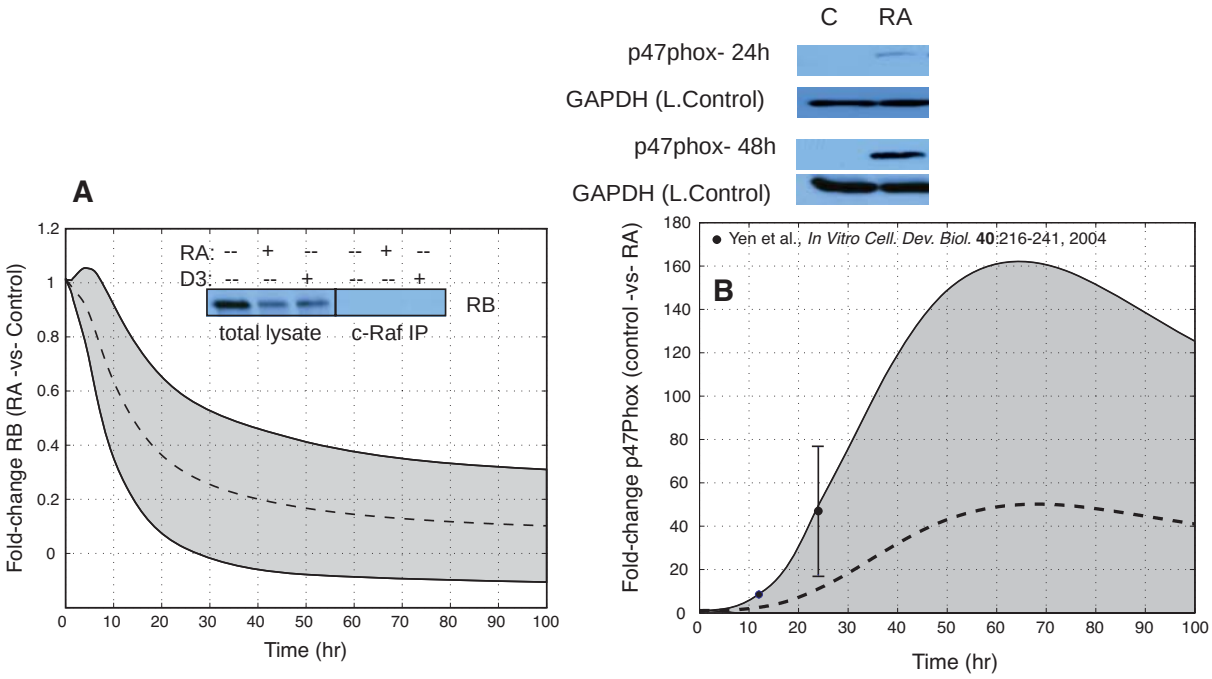


**Figure 3:** Simulations of Rb and p47Phox expression in HL-60 cells following RA exposure were qualitatively consistent with observations. In all cases the dashed line denotes the mean simulated value over the parameter ensemble while the shaded area denotes one ensemble standard deviation. **A**: Simulated profile of the Rb protein as a function of time following RA exposure (1$\mu$M). Decreased Rb protein expression was consistent with western blot analysis (inset) 24 hr following RA addition and previous studies [31]. **B:** The p47Phox protein was used as an early differentiation marker in HL-60. The model qualitatively captured the up-regulation of p47Phox following the addition of RA. Inset: Western blot analysis at 24 and 48 hr following RA exposure confirms the upregulation of p47Phox following RA treatment. These data were consistent with previously reported studies [28].

from the STRING. NETWORKIN was used to estimate missing kinase-target relationships. The 729 Ordinary Differential Equations (ODEs) of the HL60 model were machine generated using UNIVERSAL [19]. In addition to the model code, all code required to perform sensitivity analysis was also machine generated. The HL-60 model considered here had 1356 kinetic parameters and 126 unspecified initial conditions.

**Initial HL-60 model simulations revealed connectivity between MAPK and BLR1.** As a proof of concept, simulations of MAPK activation following RA treatment in BLR1 wt, knock-out and knock-in HL-60 cells were compared with previous data [21]. The initial HL-60 network model was able to show up-regulation of BLR1 expression following RA exposure along with generation of RAFp, ERKp and MEKp (Figure 2). The model was also able to qualitatively describe the linkage between BLR1 expression following RA exposure and MAPK activation in wt, blr1-/- and blr1+/+ HL-60 cells (data not shown). Both the qualitative trends and relative time scales of the initial simulations were consistent with observations over the parameter ensemble. Assembling the modular model components led to several testable linkages between BLR1 expression and MAPK. The route active in the initial simulations was Gq-protein activation of PKC via PLC$\gamma$ and DAG. BLR1 expression was predicted to drive IP3 and DAG formation through coupling with Gq-protein and PLC$\gamma$ activation. In the model DAG was linked with the activation of Protein Kinase C (PKC) by helping recruit it to the plasma membrane. PKC has been shown to phosphorylate both RAS and RAF leading to MAPK activation [14]. The BLR1-PKC connection could explain the activation of ERK following RA exposure but not the dependence of BLR1 expression upon MAPK activity. To this question, ERKp was found to activate the NFATc3 protein, a component in the transcription factor complex driving BLR1 expression. This connection could explain the dependence of BLR1 expression on RAF (data not shown).

The HL-60 model only partially captured expression shifts driving the arrest and differentiation program following RA exposure (Figure 3). The Rb and p47Phox proteins were used as markers for the arrest and differentiation programs, respectively. Simulations of the expression of these key markers following the addition of RA showed that the model qualitatively captured the expected shifts. Rb expression decreased following RA exposure consistent (Figure 3A). However, the onset of the differentiation program indicated by p47Phox expression was not well constrained (Figure 3B). The cell-cycle arrest program was directly driven by MAPK activation which was well described by the model. Thus, the tight distribution on the Rb simulations over the parameter ensemble indicated that parameters controlling Rb expression were likely well constrained. Conversely, the wide distribution of the p47Phox trajectory was likely caused by a lack of sufficient training data.

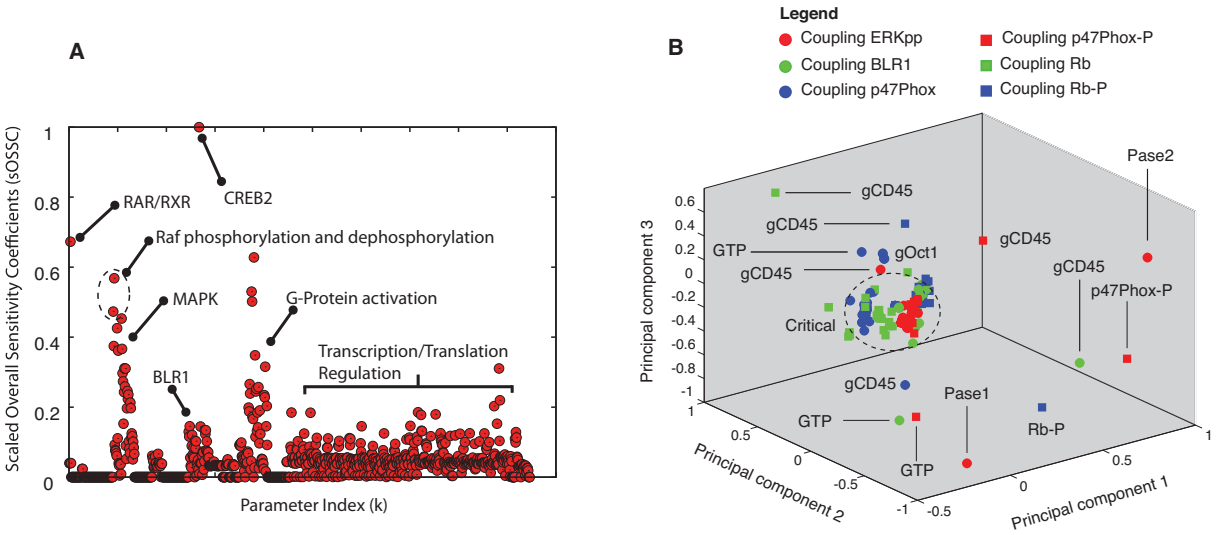


**Figure 4:** Initial sensitivity and coupling analysis predicts critical linkages and proteins. **A:** Scaled Overall State Sensitivity Coefficients (OSSCs) calculated using the initial best fit parameter set in the presence of RA. Formation of the BLR1 transcription factor complex (CREB-NFATc3-Oct1) and MAPK activation were estimated to be critical components. The OSSC value for each of the network parameters was scaled by the maximum OSSC value (CREB2 binding). **B:** Coupling analysis for six marker proteins computed over a sparse sampling of the parameter ensemble (N = 10). The family of mean coupling coefficients were analyzed using Singular Value Decomposition (SVD).

**Sensitivity and coupling analysis of the HL-60 model revealed critical model parameters.** Sensitivity analysis was conducted to better understand which parameters and interactions in the HL-60 model were critical (Figure 4). As a proof-of-concept, we preformed sensitivity and coupling analysis of wt HL-60 in the presence of RA. Overall State Sensitivity Coefficients (OSSC) for each of the model parameters were calculated using the best initial parameter set (parent set used in the ensemble calculations). The model and sensitivity equations were solved simultaneously using the LSODE routine of Octave (www.octave.org). The most sensitive model parameters involved CREB2 binding in the BLR1 transcription factor complex and components of MAPK activation. In addition, several elements of the Gq-protein activation cascade were found to be sensitive. Interestingly, of the sensitive MAPK cascade interactions, Raf activation and deactivation were found to be the most important (Figure 4A). Coupling coefficients computed between six model proteins (ERKpp, BLR1, p47Phox, p47Phox-P, Rb and Rb-P) and the knockdown of 97 model proteins were computed using a sparse sampling of the parameter ensemble (N = 10). The mean coupling for each of the marker proteins was used to construct a $97 \times 6$ coupling array which was then analyzed using the Singular Value Decomposition (SVD) function of Matlab (The Mathworks, Natick MA). Coupling analysis identified critical hubs in the network that amplified or destroyed the marker trajectory. For example, ERKpp was found to be positively coupled to deletion of phosphates acting on Raf and Mek (Pase 1 and Pase 2, respectively). This finding was consistent with the sensitivity results. Conversely, all markers were found to be critically coupled to the deletion of eI4FE and other translation components (near the origin of Figure 4B). When taken together, the results support the application of sensitivity and coupling analysis using ensembles of mechanistic models to robustly rank-order the importance of proteins and interaction in molecular networks. However, the coupling and sensitivity results were preliminary. More parameter sets should be sampled to make sure our conclusions were robust to the choice of parameters.

Honor, Loyalty and Commitment

# 4 References


[1] Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nat Biotechnol*, 20(4):370–375, Apr 2002.

[2] T. E. Battle, R. A. Levine, and A. Yen. Retinoic acid induced blr1 expression promotes erk2 activation and cell differentiation. *In Vitro Cell Dev. Biol.*, 254:287 – 298, 2000.

[3] T. E. Battle, M. S. Roberson, T. Zhang, S. Varvayanis, and A. Yen. Retinoic acid induced blr1 expression requires RARa, RXR and MAPK activation and uses ERK2 but not JNK/SAPK to accelerate cell differentiation. *Eur. J. Cell Biol.*, 80:59 – 67, 2001.

[4] D. Battogtokh, D. K. Asch, M. E. Case, J. Arnold, and H.-B. Schuttler. An ensemble method for identifying regulatory circuits with special reference to the qa gene cluster of neurospora crassa. *Proc Natl Acad Sci U S A*, 99(26):16904–16909, Dec 2002.

[5] T. R. Breitman, S. E. Selonick, and S. J. Collins. Introduction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. *Proc. Natl. Acad. Sci. USA*, 77:2936 – 2940, 1980.

[6] K. S. Brown and J. P. Sethna. Statistical mechanical approaches to models with many poorly known parameters. *Phys Rev E Stat Nonlin Soft Matter Phys*, 68(2 Pt 1):021904, Aug 2003.

[7] S. J. Collins, R. C. Gallo, and R. E. Gallagher. Continuous growth and differentiation of human myeloid leukemic cells in suspension culture. *Nature*, 270:347 – 349, 1997.

[8] R. P. Dickinson and R. J. Gelinas. Sensitivity analysis of ordinary differential equation systems - a direct method. *J Comp Phys*, 21:123 – 143, 1976.

[9] T. Dobner, I. Wolf, T. Emrich, and M. Lipp. Differentiation-specific expression of a novel G protein-coupled receptor from Burkitt's lymphoma. *Eur. J. Immunol.*, 22:2795 – 2799, 1992.

[10] T. Emrich, R. Forster, and M. Lipp. Transmembrane topology of the lymphocyte-specific G protein-coupled receptor BLR1: analysis by flow cytometry and immunocytochemistry. *Cell Mol. Biol.*, 40:413 – 419, 1994.

[11] R. Forster, I. Wolf, E. Kaiser, and M. Lipp. Selective expression of the murine homologue of the G protein-coupled receptor BLR1 in B cell differentiation, B cell neoplasia and defined areas of the cerebellum. *Cell Mol. Biol.*, 40:381 – 387, 1994.

[12] 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 signalling. *Biochem J*, 373(Pt 2):451–463, 2003 Jul 15.

[13] H. Y. Hong, S. Varvayanis, and A. Yen. Retinoic acid causes MEK-dependent RAF phosphorylation through RAR plus RXR activation in HL-60 cells. *Differentation*, 68:55 – 66, 2001.

[14] W. Kolch, G. Heidecker, G. Kochs, R. Hummel, H. Vahidi, H. Mischak, G. Finkenzeller, D. Marme, and U. R. Rapp. Protein kinase c alpha activates raf-1 by direct phosphorylation. *Nature*, 364(6434):249–252, 1993 Jul 15.

[15] D. J. Mangelsdorf, U. Kazaukiko, and R. M. Evans. *The retinoid receptors.* The Retinoids. Raven Press, 2 edition, 1994.

[16] S. Nayak, S. Salim, D. Luan, M. Zai, and J. D. Varner. A test of highly optimized tolerance reveals fragile cell-cycle mechanisms are molecular targets in clinical cancer trials. *PLoS ONE*, 3(4):e2016, 2008.

[17] S. Nayak, J. K. Siddiqui, and J. Varner. Modeling and Analysis of an Ensemble of Eukaryotic Translation Initiation Models. *Mol. Sys. Biol.*, submitted, 2008.

[18] J. Stelling, E. D. Gilles, and F. J. r. Doyle. Robustness properties of circadian clock architectures. *Proc Natl Acad Sci U S A*, 101(36):13210–13215, 2004.

[19] R. Tasseff, S. Nayak, D. Luan, T. Mansell, and J. Varner. UNIVERSAL: A web-based extensible architecture for automated model code generation. *Bioinformatics*, To be submitted, 2009.

[20] J. Wang and A. Yen. A novel retinoic acid-responsive element regulates retinoic acid induced BLR1 expression. *Mol. Cell. Biol.*, 24:2423 – 2443, 2004.

[21] J. Wang and A. Yen. A mapk-positive feedback mechanism for blr1 signaling propels retinoic acid-triggered differentiation and cell cycle arrest. *J Biol Chem*, 283(7):4375–4386, 2008 Feb 15.

[22] X. Wang and G. P. Studzinski. Activation of extracellular signal-regulated kinases (ERKs) defines the first phase of 1,25-dihydroxyvitamin D3-induced differentiation of HL-60 cells. *J. Cell Biochem.*, 80:471 – 482, 2001.

[23] A. Yen. HL-60 cells as a model of growth control and differentiation - The significance of variant cells. *Hemat. Rev.*, 4:5 – 46, 1990.

[24] A. Yen, D. Brown, and J. Fishbaugh. Precommitment states induced during HL-60 myeloid differentiation - Possible similarities of retinoic acid and DMSO induced early events. *Exp. Cell Res.*, 173:80 – 84, 1987.

[25] A. Yen, V. Cherington, B. Schaffhausen, K. Markes, and S. Varvayanis. Transformation defective polyoma middle t antigen mutants defective in plc-gamma, pi-3 or src kinase activation enhance erk2 activation and promote retinoic acid induced cell differentiation like wild type middle t. *Exp. Cell Res.*, 248:538 – 551, 1991.

[26] A. Yen, M. Forbes, G. deGala, and J. Fishbaugh. Control of HL-60 cell differentiation lineage specificity - A late event occurring after precommitment. *Cancer Res.*, 47:129 – 134, 1987.

[27] A. Yen and M. E. Forbes. C-myc down regulation and precommitment in HL-60 cells due to bromodeoxyuridine. *Cancer Res.*, 50:1411 – 1420, 1990.

[28] A. Yen, D. M. Lin, T. J. Lamkin, and S. Varvayanis. Retinoic Acid, Bromodeoxyuridine and the Delta-205 mutant polyoma virus middle T antigen regulate expression levels of a common ensemble of proteins associated with early stages of inducing HL-60 leukemic cell differentiation. *In Vitro Cell Dev. Biol.*, 40:216 – 241, 2004.

[29] A. Yen, A. W. Norman, and S. Varvayanis. Nongenomic vitamin d3 analogs activating erk2 in hl-60 cells show that retinoic acid induced differentiation and cell-cycle arrest require early concurrent mapk and rar and rxr activation. *In Vitro Cell Dev. Biol-Animal*, 37:93 – 99, 2001.

[30] A. Yen, S. L. Reece, and K. L. Albright. Dependence if HL-60 myeloid cell differentiation on continuous and split retinoic acid exposures: Pre-commitment memory associated with altered nuclear structure. *J. Cell Physiol.*, 118:227 – 286, 1984.

[31] A. Yen, M. S. Roberson, S. Varvayanis, and A. T. Lee. Retinoic Acid Induced Mitogen-activated Protein (MAP)/Extracellular Signal-regulated Kinase (ERK) Kinase-dependent MAP Kinase Activation Needed to

Elicit HL-60 Cell Differentiation and Growth Arrest. *Cancer Res*, 58(14):3163–3172, 1998.

[32] A. Yen, M. S. Roberson, S. Varvayanis, and A. T. Lee. Retinoic acid induces mitogen-activated protein (MAP)/extracellular signal regulated kinase (ERK) kinase-dependent MAP kinase activation needed to elicit HL-60 cell differentiation and growth arrest. *Cancer Res.*, 58:3163 – 3172, 1998.

[33] A. Yen, R. Sturgill, and S. Varvayanis. Retinoic acid increases amount of phosphorylated RAF. Ectopic expression of cFMS reveals that retinoic acid induced differentiation is more strongly dependent on ERK2 signaling that induced G0 arrest is. *In Vitro Cell Dev. Biol.*, 36:249 – 255, 2000.