Published in final edited form as:

Biotechnol Prog. 2014 November ; 30(6): 1247–1261. doi:10.1002/btpr.1982.

In silico model-based inference: a contemporary approach for hypothesis testing in network biology

David J. Klink II*

*Department of Chemical Engineering; Department of Microbiology, Immunology and Cell Biology; and Mary Babb Randolph Cancer Center, West Virginia University, Morgantown, WV 26506, USA

Abstract

Inductive inference plays a central role in the study of biological systems where one aims to increase their understanding of the system by reasoning backwards from uncertain observations to identify causal relationships among components of the system. These causal relationships are postulated from prior knowledge as a hypothesis or simply a model. Experiments are designed to test the model. Inferential statistics are used to establish a level of confidence in how well our postulated model explains the acquired data. This iterative process, commonly referred to as the scientific method, either improves our confidence in a model or suggests that we revisit our prior knowledge to develop a new model. Advances in technology impact how we use prior knowledge and data to formulate models of biological networks and how we observe cellular behavior. However, the approach for model-based inference has remained largely unchanged since Fisher, Neyman and Pearson developed the ideas in the early 1900's that gave rise to what is now known as classical statistical hypothesis (model) testing. Here, I will summarize conventional methods for model-based inference and suggest a contemporary approach to aid in our quest to discover how cells dynamically interpret and transmit information for therapeutic aims that integrates ideas drawn from high performance computing, Bayesian statistics, and chemical kinetics.

Keywords

computational Bayesian statistics; scientific method; inductive inference; dynamic systems; Markov Chain Monte Carlo

Supplementary Materials

Movie S1. Animation of Markov chains for simple enzyme kinetics example using an adaptive and non-adaptive MCMC algorithms.

A PDF file that includes:

- 1 Markov Chain Monte Carlo integration
- 2 A simple enzyme kinetic model.

References

Figure S1. General flowchart for MCMC-based integration to estimate $P(Y_M|M)$.

Introduction

As the fundamental unit of living systems, understanding how a cell interprets and transmits cellular information is essential for revealing mechanisms of cellular adaptation and for understanding how a collective response to a changing environment is organized.

Understanding causal mechanisms for information transfer in biological systems in health and disease is central for the rational design of therapies that aim to restore health.^{1,2} Yet, the study of mechanisms for intra- and intercellular communication is a frontier field, as evidenced by the recent discovery of RNA interference.³ Enabled by the molecular biology revolution, many new proteins and new cellular subtypes have been qualitatively associated with particular biological processes in the last 20 years. Despite this explosion in knowledge, our quantitative understanding, which is expressed in terms of causal biological networks and relates how specific components elicit a cell response, remains unclear.^{4,5} These causal biological networks may focus at the cellular scale, such as identifying how signaling proteins collectively organize to transmit information or how gene regulatory networks control phenotype, or focus at the cell population scale, such as identifying the biological role that a specific cell type plays in the tumor microenvironment. Moreover, the data that we can acquire from biological systems exhibit inherent uncertainty. The uncertainty in observing cellular information processing can be attributed to multiple sources, including the signal-to-noise characteristics of a biological assay⁶, the skill of the experimentalist⁷, ethical limitations, extrinsic variability due to slight spatial variation in experimental conditions⁸, or intrinsic variability due to sampling a finite number of states (i.e., stochasticity).^{9,10} Given the body of knowledge currently assembled and the inherent uncertainty in the data, one of the key challenges is identifying the strengths and weaknesses of our current understanding of cellular information processing.

Since the days of Francis Bacon, the scientific method has been used as a structured way to improve our understanding about natural systems.^{11,12} It is an iterative process that involves three main stages: 1) forming a hypothesis based upon prior knowledge and available data (i.e., a model), 2) designing and executing experiments that are intended to test the model, and 3) evaluating the model in light of the new data to decide on the validity of the model (Figure 1). Here, models are meant to represent causal relationships derived from either a cognitive synthesis of prior information, conventionally expressed as a hypothesis, or a mathematical synthesis of interactions thought to be important in a system.¹³ The last step, called model-based inference, either improves our confidence in a model or suggests that one revisit their prior knowledge to develop a new model that is consistent with both the prior and new data. Revealing inadequacies in our understanding of systems is essential for meaningful learning to occur.¹⁴ While the scientific method has remained unchanged through the ages, advances in technology shape how this iterative process is applied in practice.

Intense efforts have focused on improving the computational infrastructure associated with analyzing biological data to gain insight into underlying causes of observed signaling events. These advances in computing include improving the operability among storage, annotation, and analysis tools¹⁵; creating standards for converting cartoons into mathematical models¹⁶; creating platforms for crowd-sourcing model annotation^{17–19} and for simulating biochemical

reaction networks^{20–22}; and improving the speed of biochemical network simulation using graphical processing units²³ or field programmable gate arrays²⁴. In parallel, technological advances have improved the ability to observe induced changes in biological systems. These experimental advances include observing global protein expression within a population of cells^{25,26}, identifying new regulatory mechanisms using genome-wide RNA inference²⁷, and observing selective gene or protein expression at single-cell resolution.^{28,29}

Despite these advances in formulating a hypothesis (i.e., a model) and in generating data to test the model, how one interprets data in light of the model is based largely upon the ideas of statistical hypothesis testing developed in the early 1900's.^{30,31} Classical hypothesis testing techniques are conventionally used to protect against the possibility that a proposed mechanism is based upon chance alone. While such methods have been the subject of much philosophical debate^{32,33}, advances in technology motivate revisiting the ideas associated with testing models from a practical viewpoint. To provide context, I will first revisit the fundamental assumptions that underpin conventional methods used in the field. Next, I will suggest a revised approach that integrates concepts related to modeling causal dynamical systems using mathematics from chemical kinetics, to probability and inference from statistics, and to numerical integration using high performance computing from applied mathematics. Two examples, a simple enzyme kinetics example and a more complicated cytokine crosstalk example, will be used to illustrate aspects of this contemporary approach for testing hypotheses. Results for the simple enzyme kinetics example are presented in entirety to illustrate quality control metrics for these *in silico* inference methods. The cytokine crosstalk example is described more selectively as to illustrate how these tools can be used in practice and the details can be found in citations. To highlight the concepts rather than the algorithmic details, the objective here is to provide an overview of the computational workflow and how this workflow can be used in practice. Citations to specific algorithms and example applications are provided for readers interested in implementing these methods.

What is inference?

In the context of network biology, inference is the logical reasoning about how information is passed dynamically among elements of a system, expressed in terms of a model (M), using observations of system behavior. Logical reasoning can be extended using probability, as observations involve some uncertainty.³³ There are two main viewpoints in the use of probability as a tool for inference:

- **Frequentist Viewpoint:** Probability is defined as the long-run frequency of an event to occur. Observations of these events, Y , are assumed to be a random variable but the mechanistic cause of these events to occur, as one may understand it (i.e., the model M), is not random but a discrete fixed value (i.e., there is no uncertainty in M : M_F). The probability of observing events Y is then a conditional probability, which means that observing particular events depends on the model M and is denoted as $P(Y|M_F)$.
- **Bayesian Viewpoint:** Probability is defined as a level of belief in observing an event, Y , given our current understanding of the cause (M). The data, Y , and the

cause, M_R , are both considered to be random variables, in contrast to the Frequentist viewpoint. This viewpoint is also expressed as a conditional probability, $P(Y|M_R)$.

In practice, the Frequentist viewpoint is used more widely in contemporary biological network research. However to summarize how inference methods are used in practice, I will consider the conditional probability, $P(Y|M)$, irrespective of the viewpoint and then circle back to explore the implications of the different assumptions about M .

Logical reasoning in the context of signaling primarily involves inductive inference. In other words, we want to increase our understanding of the system by using uncertain observations to reason “backwards” as to what mechanism could cause those observed events. The desired quantity in this context is the probability of a given model that is conditioned on the observed data, which is denoted as $P(M|Y)$. To relate $P(Y|M)$ with $P(M|Y)$, consider a universe with two orthogonal dimensions (Figure 2). The first dimension is an observable dimension that contains a finite set of plausible events, Y_i . The second dimension is an unobservable dimension that contains a finite set of plausible causes, M_i . Not all combinations of causes and events are feasible such that observing an event, Y , limits the possibilities for M (that is it conditions our belief in plausible causes). For instance, the observation of Y_2 conditions our belief such that M_1 and M_2 are more likely than M_3 (Figure 3A).

In the late 1700's, a general relationship between conditional probabilities was proposed that can be interpreted in terms of our universe as^{34,35}:

$$P(M_i|Y) \cdot P(Y) = P(Y|M_i) \cdot P(M_i). \quad (1)$$

This is illustrated in Figure 3, where the hashed area represents the joint probability for both Y_2 to be observed and M_2 to be the cause. This joint probability is the same irrespective of whether one conditions on the observation Y_2 (Figure 3A) or the cause M_2 (Figure 3B). This relationship among conditional probabilities is expressed generally as:

$$P(M|Y) = \frac{P(Y|M) \cdot P(M)}{P(Y)}, \quad (2)$$

where $P(Y)$ and $P(Y|M)$ are called the evidence and the likelihood, respectively. The evidence can be expressed as a conditional sum over all possible causes:

$$P(Y) = P(Y|M) \cdot P(M) + P(Y|\bar{M}) \cdot P(\bar{M}), \quad (3)$$

where, here, the set of all possible causes includes just two alternatives: the model M and not the model M . The likelihood, $P(Y|M)$, corresponds to the likelihood of observing a particular event, given an underlying mechanistic cause. Using our model, one can make a prediction, y_M . The difference between the model prediction, y_M , and an observed event, y , is defined as the error, ε (i.e., $\varepsilon = y_M - y$). Assuming that the error is distributed normally and quantified by the experimental variance (σ^2), an approximate expression for the likelihood can be derived:

$$P(Y|M) \approx \frac{1}{\sum_{i=1}^n (y_i - y_{M_i})^2}, \quad (4)$$

where the sum is over n independent, identically distributed observations and their corresponding predictions.

The approximate relationship shown in Equation 4 is helpful as it illustrates how one can reason logically about our understanding of a system, expressed as a model, using a series of comparisons between what one expects to see (i.e., Y_M) and the actual observations, Y . To illustrate this point, consider a scenario where you are walking down the street and are approached by a stranger. As this stranger approaches you, you want to quickly determine, given the data available, whether this stranger comes in friendship or has other intentions. One way to assess their intentions is to look at their facial expression.³⁶ While you have never met this person before, you have mental models for what you would expect to see. If you are meeting a friend, you would expect to see the person smiling in joy. You would expect to see an expression of anger, if the stranger has ill intentions. As this is a stranger, you also have no prior data to inform your knowledge of their intentions. Thinking in terms of the relationships described in Equations 2 – 4, you could then assess the probability that the stranger comes in friendship based on how well the stranger's facial expression matches with these two mental models. This simple example serves to illustrate how Bayesian relationships can be used to frame inductive inference problems that we encounter every day. In the next section, I will describe how these relationships are used in the context of network biology for conventional inductive inference.

Conventional inductive inference in studying dynamic biological systems

To illustrate how inductive inference is used in practice, consider the following hypothetical cell signaling example. Based upon prior knowledge, a new model of a signaling pathway (i.e., M) is proposed and is presented visually as a signaling cartoon (Figure 4A). In this model, protein A is thought to be a key signaling transducer in signaling pathway B . Binding of a soluble ligand (L_B) to the B receptor (R_B) originates signals within the B pathway. The ligand-receptor complex then phosphorylates protein A . To test this model, a drug inhibitor (D) of the kinase activity of the ligand-receptor complex will be added to the cell culture media prior to stimulating the cells with ligand (Figure 4B). The phosphorylation state of protein A will be measured at a single time point in biological triplicate ($n = 3$). The extent of phosphorylation of A upon ligand-stimulation will be compared with and without the addition of the drug. We can also assume that the variance of the assay used to measure the phosphorylation status of A is known (σ), normally distributed, and similar across conditions. As the observed data, Y , are samples drawn from a distribution, one must integrate over a range of values to obtain a probability. Conventionally, the range of values used in the integration is from the observed events to more extreme observations. The likelihood is also evaluated under a “null hypothesis” (i.e., that the topology of the signaling pathway is incorrect or simply M):

$$P(Y|\overline{M}) = \int_z^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp^{-\frac{x^2}{2}} dx, \quad (5)$$

where z is defined as

$$z = \frac{\mu_I - \mu_C}{\sqrt{2}\sigma / \sqrt{n}} \quad (6)$$

and μ_I and μ_C are the mean responses following stimulation with the ligand with the small molecular inhibitor added and without the inhibitor (control), respectively. If $P(Y|M)$ is less than a defined probability (e.g., less than 0.05), then the “null hypothesis” (i.e., M) is rejected (as depicted by the * in Figure 4B).¹

To make statements about the model M , there are a few additional points that are considered. As these points are not explicitly stated, this decision process can be considered *in cerebro* model-based inference. *A priori*, the observed evidence, Y , may be attributed to multiple causes that include the model, the null hypothesis, or technical error:

$$P(Y) = P(Y|M) \cdot P(M) + P(Y|\overline{M}) \cdot P(\overline{M}) + P(Y|\text{technical error}) \cdot P(\text{technical error}). \quad (7)$$

The phosphorylation status of A following addition of the soluble ligand in the absence of the drug is used as a positive control. The phosphorylation status of A without the addition of the soluble ligand is used as a negative control. The positive and negative controls plus the technical expertise of the observer help ensure that $P(Y|\text{technical error}) \cdot P(\text{technical error})$ is negligible. The belief in the model is then:

$$P(M|Y) = 1 - \frac{P(Y|\overline{M}) \cdot P(\overline{M})}{P(Y|M) \cdot P(M) + P(Y|\overline{M}) \cdot P(\overline{M})}. \quad (8)$$

Assuming that the model priors are discrete and unbiased (i.e., $P(M)$ and $P(\overline{M}) = 0.5$), the likelihood $P(Y|M)$ is interpreted as one is 95% confident that M is true (i.e., $P(M|Y) > 0.95$) when $P(Y|M)$ is less than 0.05. While this interpretation is common practice, a number of concerns have been expressed regarding the application of statistical inference in this way.³⁷

Strong inference, as proposed by Platt³⁸, involves generating a series of competing hypotheses based upon prior knowledge and devising a minimal set of experiments that crisply differentiates among the competing hypotheses. Each of the alternative hypotheses are evaluated in a similar manner as described for this simple example, a process called “hypothesis testing.” In other words, the accepted cause for the behavior is obtained by rejecting a series of alternative “null hypotheses” rather than actually testing the proposed model. There are additional assumptions that underpin this interpretation. First, probability

¹In practice, the integral is pre-calculated with different values of z to give the corresponding probability associated with a one-sided or two-sided tail. This probability is defined as α or a “p-value.” For a given α , the corresponding value of z is set as the critical z value (z_{CRIT}). The “null hypothesis” is rejected when z is greater than z_{CRIT} .

is obtained using the available data and unobserved extreme data. Second, equal weight is assigned to the model priors. Third, lurking mechanisms do not exist (i.e., all possible causes are formulated *a priori*). This assumption is problematic in frontier areas of science. Reverse transcriptase^{39,40} and RNA interference³ constitute lurking mechanisms that, upon discovery, fundamentally shifted our understanding of the flow of genetic information. Fourth, models are crisp descriptions of system behavior that can be elucidated via experimentation (i.e., there is no intrinsic uncertainty in the model). With the exception of the prior weights, these assumptions are tenuous in contemporary biological network research.

Dynamic nonlinear feedback is a common regulatory motif in signaling systems.⁴¹ In the previous hypothetical example, events were observed at a single time point. If feedback regulation is a lurking mechanism, the belief in a particular model may depend on the time point selected. Mathematical modeling has been proposed as a tool to assist the decision process by representing the dynamics of the system within the model (e.g., *in silico* model-based inference⁴²). In such case, $P(Y|M)$ could be calculated directly and compared against data obtained at multiple time points. Conventionally, an estimate of $P(Y|M)$ is obtained by comparing the model predictions against the observed events as a function of the independent variable, time (Figure 4C). The comparison can be quantified using a Chi-squared statistic or correlation coefficient.^{43,44} If systematic differences between the model and the observed events are identified, the model is rejected. Compared to the indirect calculation of $P(Y|M)$ using $P(Y|M)$, the use of a mathematical model enables direct calculation of $P(Y|M)$ using just the observed data. However, three assumptions still apply: 1) equal weight on priors, 2) no lurking mechanisms, and 3) model uncertainty does not exist. Critics of the use of mathematical models have argued that the model predictions (Y_M) depend on the values of the model parameters, Θ (i.e., $P(Y_M|\Theta, M)$). In biological systems, the model parameters quantify the influence on the model predictions of protein-protein interaction energies, local protein concentrations, intracellular transport properties, rates of degradation and synthesis, or nonphysical empirical quantities that are based upon the form of the model. Values for many of these parameters are unknown.

Within a Frequentist view of probability, model parameters have a “true” value - a singular or discrete value. The process of observing the system convolutes this “true” value with experimental uncertainty. The magnitude of experimental uncertainty may depend on the signal-to-noise characteristics of a biological assay⁶ or the skill of the experimentalist⁷. Many of the current methods for parameter optimization aim to deconvolute this “true” value from the observed data using a model (e.g.,⁴⁵). As a first step in deconvoluting parameter values, numerical tools for parameter identification are commonly used to determine whether the parameters in the model can be uniquely determined (i.e., the problem is well posed). Methods for *a priori* or structural identifiability select identifiable parameters given unlimited information about the modeled system.^{46–48} Historically, this has been motivated by technical considerations as optimization methods that use gradients of the likelihood with respect to the parameters run into problems for parameters that are not identifiable. However, there are a number of biological and practical reasons why “true” values for model parameters can not be obtained in practice. Biologically, values for many

of these parameters may vary within cells of a population due to differences in copy numbers^{9,10}, to slight spatial variation in experimental conditions⁸, or to different point mutations in different cells of a population that alter the affinity between proteins.⁴⁹ This would suggest that there is inherent biological uncertainty associated with the model parameters. In addition, the specific observed data also limit the ability constrain possible values for the model parameters, a limitation linked to the slaving principle.

The slaving principle describes how the dynamic response of a system comprised of many elements can be described using a smaller collection of elements and was developed based upon observing physical and chemical kinetic systems.⁵⁰ In biology, identifying key elements that regulate a dynamic response from all of the elements associated with a particular biological system is a common challenge. In developing models to describe the dynamics of biological networks, one may construct the model using elements known to be involved in individual pathways within a biological network and desire to understand how information is transferred among these elements in response to a stimulus. The flow of information among the elements of a pathway is governed by biochemical transformations. Biochemical transformations include the post-translational modification of a protein, the induction of gene expression, the extracellular release of a biochemical cue, or the formation of protein-protein interactions. The challenge in model-based inference is that we may not know the dynamics associated with these transformations *a priori*. The slaving principle implies that the flow of information within biochemical networks initiated by a stimulus is governed by the slowest transformation, that is the rate-limiting step.² In contrast, other transformations act much faster than the rate-limiting step or appear stationary, as they occur much slower than the time frame over which we observe the system. Identifying transformations that are not kinetically important can be used to simplify the model⁵¹.

Similar types of transformations may also exhibit very different dynamics. For instance, two proteins may exist as a preformed complex while other proteins only associate following the activation of an upstream element. The characteristic time frame in which a transformation occurs is referred to as a time scale. Given that the time scales associated with the transformations contained within a pathway are unknown *a priori*, determining unique parameter values associated with each transformation based upon the observed data presents a challenge.³ Parameters associated with transformations that are fast, such as pre-formed multi-protein complexes, and that are kinetically unimportant, such as stationary reactions, exhibit one-sided distributions.⁵¹ This means that parameters associated with fast (or stationary) transformations are only constrained such that the value has to provide a time scale that is faster (or slower) than the observable time scales associated with the rate-limiting steps.⁴ This is a subtle but important point as many studies that apply statistical

²The concept of a rate-limiting step implies that the control of the flow through a biological network is regulated by a single step (that is, a local control mechanism). However, a contemporary view is that information flow is regulated by a distributed control mechanism where there are multiple transformations that are distributed across a network that have similar slow time scales. A distributed control mechanism provides more consistent behavior in the presence of changes that may alter the behavior of an individual transformation, such as a mutation.

³To get around this challenge, many model-based inference studies assume *a priori* that all of the postulated mechanistic steps encoded in the model are kinetically important, which means that there are no fast or extremely slow reactions (e.g., ⁵²). In addition, parameter identifiability methods are employed *a priori* to simplify the model such that it contains only identifiable parameters (e.g., ^{53,54}).

inference methods to biological network questions provide distributions in the model parameters that have bounds all supposedly informed by data (i.e., a posterior distribution) and that are in the form of a multivariate Gaussian distribution (e.g., Sethna and coworkers (e.g.,^{55–57}), Stumpf and coworkers (e.g.,^{58–60}), and Girolami and coworkers (e.g.,^{52,61}). However when “posterior” distributions for dynamic systems are inconsistent with the slaving principle, this raises the question as to whether these “posterior” distributions really reflect the data or whether they reflect an arbitrary selection of a prior for the parameters or the model has been overly simplified to achieve an empirical fit rather than test a mechanistic hypothesis. In short, the observable time scales reflect a combination of the signal-to-noise characteristics of the experimental assay⁶, how frequently the system is observed (similar to the Nyquist-Shannon sampling theorem in signal processing), and what states are observed. Collectively, the presence of inherent uncertainty and unknown dynamics within a biological network imply that model parameters are in practice not discrete values but also exhibit uncertainty that is conditioned both on the available data and the form of the model used to interpret the data ($P(\Theta|M, Y)$).

Our confidence in a mechanistic explanation for a biological observation rests in the uncertainty associated with the model predictions (i.e., model uncertainty). The uncertainty in model predictions can be due to one of two factors: (1) the model does not represent the right mechanism (i.e., topological uncertainty) or (2) the model predictions reflect a biased selection of the model parameters (i.e., parametric uncertainty). Without having a way to select all plausible combinations of parameter values (or a sufficiently large sample) that are consistent with the data, one can not rule out the possibility that the lack of agreement between a model and data is due to inappropriate parameter values nor can one specify the importance of network elements in dynamically regulating the flow of information. Therefore, discriminating between different causal mechanisms using mathematical models and parameter uncertainty are linked. In light of Platt’s comments on scientific inquiry^{38,42}, this suggests that the conventional use of mathematical models is a weak inference approach for inferring the importance of regulatory elements within networks or identifying lurking mechanisms. In the next section, I will describe an alternative approach that addresses some of the weaknesses associated with conventional model-based inference.

A Bayesian approach to in silico model-based inductive inference

In order to include model uncertainty in our inductive inference, we need to estimate the uncertainty associated with the model predictions that accounts for the specific data at hand, the uncertainty in the model parameters, and our prior causal knowledge of the system. In this way, we can improve the strength of the inference in whether the causal structure of the model is unable to describe our system. In the following paragraphs, I will briefly describe the computational approach, summarize quality-control measures in the estimation of the

⁴In the examples presented here, the possible parameter values were limited to a broad range (i.e., 10^{-8} to 10^8) to ease the computational burden that is associated with solving numerically problems with broad timescales (i.e., stiff systems). Alternatively, parameter values could be limited to ranges that are biologically plausible. While these practical limits imply that all parameter distributions are bounded, it is important to distinguish between a bound in a parameter distribution that is informed by data or that is specified for computational convenience or biological plausibility. The wide bounds specified for computational convenience, as illustrated here, makes it easier to distinguish between these two scenarios.

posterior distribution in the model predictions (i.e., $P(Y_M|M)$), and provide examples of how model-based inductive inference can be used to inform biological network research.

Generally, causal mathematical models, such as ordinary differential equation-based models⁶², define the causal relationships between k interacting species. These causal models, in conjunction with a particular set of parameter values and initial conditions, define a dynamical system that evolves in time along a deterministic trajectory within k -dimensional Euclidian species space, $\mathbb{R}^k$. Despite our uncertainty in the values of the parameters and initial conditions, the causal structure of the model constrains the possible species space through which the model predictions can traverse. We want to determine the subspace in $\mathbb{R}^k$ defined by a collection of trajectories, Y_M , that are constrained by the available data Y . As the available data only constrains particular regions in $\mathbb{R}^k$, we also want to learn what else is implied by this causal model using the collection of trajectories. Conventionally, a single trajectory or a small random sample (that is, an ensemble) of trajectories is used to make causal statements about how well the model explains the data. However, these causal statements must be made in a probabilistic framework in order to establish a level of confidence in the model predictions, given the available data.

A probabilistic collection of trajectories in $\mathbb{R}^k$ is defined by the conditional relation: $P(Y_M|M)$, where M is the model. The particular collection of interest is the subspace that accounts for (i.e., integrates over) the available data and the uncertainty in the model parameters:

$$P(Y_M|M) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(Y_M|\Theta, M) \cdot P(\Theta|M, Y) \cdot P(Y) d\Theta dY. \quad (9)$$

As $P(\Theta|M, Y)$ is difficult to obtain directly, Bayes theorem is used to re-express this in terms of quantities that we can calculate:

$$P(\Theta|M, Y) \cdot P(Y) = P(Y|\Theta, M) \cdot P(\Theta|M). \quad (10)$$

In equation 10, $P(\Theta|M)$ is the probability of sampling a point (Θ_i) in parameter space Θ prior to any knowledge about data Y (that is, the prior for Θ), $P(Y)$ is called conventionally the evidence for a model, and $P(Y|\Theta_i, M)$ is the conditional probability of simulating data Y given Θ_i and M (i.e., the likelihood of Y given Θ_i and M).

In contrast to the approximate relationship described in Equation 4, a more accurate relationship for the likelihood is used for *in silico* model-based inference. A statistical model for how the errors associated with the experimental observations are distributed, such as a normal distribution, is used to generate the specific relationship used for the likelihood. One approach, summarized in the supplemental text, is to assume that the experimental error is distributed normally and the error variance may differ among data sets.^{63,64} The resulting likelihood is expressed as:

$$P(Y|\Theta_i, M) \propto |v|^{-N_{obs}/2}. \quad (11)$$

where $|v|$ is the determinant of the dispersion matrix. For k experimental variables observed N_{obs} times, the dispersion matrix is a $k \times k$ positive-definite matrix where each element,

$$v_{jl} = \sum_{u=1}^{N_{obs}} \{Y_{ju} - Y_{Mj}(\Theta_i)\} \cdot \{Y_{lu} - Y_{Ml}(\Theta_i)\}, \quad (12)$$

corresponds to the normalized sum of the product of the deviation between specific observations, Y_j , and their respective model prediction, $Y_{Mj}(\Theta_i)$. This likelihood relationship is equivalent to a normalized summed squares of the error, where the number of observations appears as an exponential factor. Another approach is to use a likelihood function that assumes that the experimental error is the same for all experimentally observed quantities.⁶¹ In describing an approximate Bayesian computation (ABC) method, Toni et al. use an approximate likelihood function that is described as a Euclidian distance metric, which is equal to the determinant of the dispersion matrix alone and neglects the number of observations associated with a data set.⁵⁸

Upon combining equations 9 and 10 together, the desired integral is:

$$P(Y_M|M) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(Y_M|\Theta, M) \cdot P(Y|\Theta, M) \cdot P(\Theta|M) d\Theta dY. \quad (13)$$

In contrast to pre-calculated quantities used in Frequentist inductive inference (e.g., Equation 5), Bayesian model-based inductive inference requires ad hoc calculation of this integral. As the available data is a finite discrete set, integration with respect to Y is equivalent to a sum over the comparisons between each model prediction, Y_M , and the corresponding observation, Y , as represented by summation in the likelihood, $P(Y|\Theta, M)$, described in Equations 11 and 12. Integrating with respect to the parameters, Θ , is more difficult. Advances in computational power have enabled Markov Chain Monte Carlo (MCMC) methods to integrate Equation 13, as summarized in the Supplemental Text. Given the explosion of MCMC methods in general, resources for MCMC integration are abundant, including stand-alone software^{65–68}, R packages⁶⁹, and reference texts^{70,71}.

In short, a MCMC method is a computationally intensive approach that randomly walks within parameter space (i.e., a Monte Carlo method). New steps in parameter space extend the walk. While there are different approaches for determining how to extend the walk, here the emphasis is on a Metropolis-Hastings algorithm.^{72,73} Using the Metropolis-Hastings algorithm and assuming uniform priors for the model parameters and a symmetric proposal distribution, the similarity between the model predictions and the available data, as quantified by $P(Y|\theta_i, M)$, is used to calculate the probability for including the potential new step in the on-going walk. As the models described here are deterministic models, $P(Y|\theta_i, M)$ is a singular value that is calculated using Equation 11. Consistent agreement between model predictions and the available data has a high probability for including the new step in the on-going walk while disagreement has a low probability for inclusion. Inclusion of the potential new step is based solely on the comparison against the current step in the random walk (that is, it is a Markov Chain). Collectively, the Markov chain is intended to represent samples obtained from $P(Y_M|M)$. Although algorithms for MCMC integration are relatively

simple to program (see Supplemental File), they can be challenging to implement correctly such that the results provide an estimate of $P(Y_M|M)$. Similar to experimental controls, metrics must be used to control for the quality of the computational results.

Controlling the quality of $P(Y_M|M)$

In this section, a simple enzyme kinetics example illustrates how a posterior distribution in model predictions can be estimated using an adaptive MCMC approach based on the Metropolis-Hastings (M-H) algorithm. Specific details regarding the equations and the M-H algorithm used are described in the Supplemental File. In short, the kinetics of the enzymatic reaction can be described in terms of three rate equations (Figure 5A). Upon converting these rate equations into a system of two coupled ordinary differential equations, synthetic data were generated from the model predictions that mimic typical enzyme kinetics experiments for two different initial substrate concentrations (Figure 5, Panels B and C). An adaptive MCMC approach was then used to re-estimate the model parameters and the two initial substrate concentrations using the synthetic data by randomly walking in logarithmic parameter space. This adaptive MCMC approach uses the correlation of the parameters obtained from the prior steps of an evolving Markov chain to improve the efficiency of MCMC integration.⁷⁴ A log transformation was used to ensure that a parameter remains non-zero and the sign of the parameter remains constant. A combination of graphical plots and metrics are used as quality control tests when estimating $P(Y_M|M)$, as summarized in Gelman et al.⁷⁰ and illustrated in Figures 6–9.

First, three or more independent Markov chains that originate at different (that is, over-disperse) starting points are used to sample parameter space. A trace of the Markov chain, shown for each parameter, is used to assess the random nature of the walk within parameter space (Figure 6). A comparison of the traces for all three chains is used to assess the independence of the chains. The traces of the likelihoods for each of the three chains suggest that the model predictions provide similar agreement with the synthetic data. The traces suggest that the variation in the initial substrate concentrations among the Markov chains is much less than variation in the rate parameters. The wide variation in the parameter values contained within the Markov chains reflect the adaptive proposal component of the M-H algorithm. The adaptive component changes the proposal distribution to reflect the covariance structure of the cumulative Markov chain. To illustrate this point, one should recall that a k -dimensional symmetric covariance matrix can be represented by a set of k orthogonal eigenvalues and eigenvectors. As illustrated by trace of the eigenvalues of the proposal distribution (Figure 7A), the adaptive part of the M-H algorithm changes the proposal distribution from a 5-dimensional sphere, where all 5 eigenvalues are equal, to stretched hyperellipsoid, where the eigenvalues are all different. The eigenvalues of the proposal distribution correspond to the lengths of the orthogonal axes of the stretched hyperellipsoid. The magnitude of the eigenvalue corresponds to size of proposed steps in the direction of the corresponding eigenvector, which may be aligned along a parameter axis or in a direction that reflects the correlation between parameters. The volume of the proposal distribution (Figure 7B) can then be estimated by the product of the square roots of the eigenvalues times a scaling parameter, s , such that approximately 20% of the proposed steps

are accepted (Figure 7C). Having the M-H algorithm accept only 20% of the proposed steps allows the chains to exit local minima within a rugged posterior probability landscape.⁷⁵

The volume of the proposal distribution corresponds to the potential region in parameter space where proposed steps can be obtained. The trace of the eigenvalues implies that this AMCMC method proposes small steps for parameters that have a large influence on the likelihood values and large steps for parameters that do not influence the likelihood value. As suggested by the traces, the parameters for the initial substrate concentrations have small eigenvalues while k_f and k_r have large eigenvalues and, as will be shown later, an eigenvector that reflects the correlation between these two parameters. Moreover, the trace of the volume of the proposal distribution implies that the region of parameter space potentially sampled in each step is increased for the adaptive versus a non-adaptive M-H algorithm. In comparing three Markov Chains obtained using adaptive versus non-adaptive variants of the M-H algorithm, we see that the adaptive Markov chains are more efficient in tracing the contours of the posterior distribution (Movie S1).⁶

As a Markov chain may originate anywhere in parameter space, convergence is a criterion used to assess whether the Markov chain is sampling from the target distribution, $P(Y_M|M)$. Here, the Gelman-Rubin statistic (GRS) was used, as it can be readily applied to model predictions as well as model parameters.^{79,80} The GRS is based upon the concept that convergence has been achieved when the variance among chains is less than within single chains. Here, we compare the model predictions derived from the three Markov chains and plot the GRS as a function of time and AMCMC step (Figure 8A). For this simple enzyme kinetic example, the model predicts the substrate concentration as a function of time. A value for the GRS less than 1.2 suggests that the chains are converged. For both simulated experimental conditions, the chains appear to be converged after 20,000 steps. The MCMC samples obtained prior to convergence should be discarded, as they may come from the tails of the distributions.⁷

Once the random walks are sampling from the target distribution, the cumulative Markov Chains correspond to samples of parameter space that provide model predictions that are consistent with the observed data (Figure 8B). For this simple example, the posterior distributions in the model predictions for all three chains essentially overlap and exhibit a narrower distribution than the experimental data. Allowing the Markov chains to continue

⁵One particular challenge is when the posterior distribution is multi-modal, where distinct regions of parameter space provide similar posterior predictions. The inability to traverse low probability manifolds that separate the distinct regions of parameter space is an inherent risk associated with MCMC integration methods. Various strategies can be used to mitigate this risk, such as over-disperse starting points for more independent chains and a lower acceptance ratio of proposed steps. Alternatively, steps can be exchanged between chains using methods such as population-based reversible jump MCMC⁷⁶, exchange Monte Carlo⁷⁷, and Parallel Hierarchical Sampling.⁷⁸ Brute force methods that march sequentially over the entire parameter space are the alternative; yet, they are computationally intractable for many practical applications.

⁶At first glance, the adaptive M-H variant may seem to violate the Markov condition, that is acceptance of a new step only depends on the comparison between posterior probabilities of the current step and the proposed step. However, the adaptive component is implemented such that the adaptive component diminishes as the inverse of the number of steps. In the limit of a long chain, there is no adaptation. In addition, the proposal distribution, as implemented here, is always symmetric. This satisfies the detailed balance condition imposed by the Metropolis-Hasting approach and reduces the acceptance criteria to a ratio of likelihoods.

⁷In conventional statistical inference of biochemical networks using MCMC methods, the model parameters are used to assess convergence (e.g.,^{52,81,82}), which implies that the integration performed in those studies is different from Equation 13. Considering the action of the slaving principle in biochemical networks, parameters that exhibit one-sided distributions will be unable to be converged and are commonly referred to as practically non-identifiable.⁴⁶

past the point of convergence allows the chains to also illuminate the contours of the posterior distribution in the model parameters (i.e., $P(\Theta|M, Y)$) (Figure 9). The posterior distributions in the initial substrate concentrations are tightly bounded normal distributions with little correlation with the other parameters. The posterior distributions in the rate constants exhibited more complicated structure, whereby the joint posterior distribution for k_f and k_r is not a multivariate Gaussian distribution but suggests a bi-linear relationship with two distinct regions within the joint posterior distribution. In the first region, high likelihood predictions can be obtained when k_f is $10^{1.2} \text{ mM}^{-1} \text{ sec}^{-1}$ and k_r is assigned any value in the range 10^{-6} to 10^{-2} sec^{-1} . In the second region, high likelihood predictions can be obtained with values of k_r greater than 10^{-2} sec^{-1} but a corresponding increase in k_f is required, that is the values of these two parameters are highly correlated. In comparing these two regions, k_r would be considered “sloppy” in first region while k_f would be more “sloppy” in the second region.⁵⁷

These two different joint posterior regions for k_f and k_r have different biological implications. In biochemical reactions, the propensity for two proteins to associate, as represented by the association rate constant k_f , is determined by the ability of a protein to migrate to a binding site and to re-arrange itself spatially. Dissociation reactions, as parameterized by k_r , account for the thermodynamic stability of the complex. Point mutations alter the strength of interaction between two proteins. The first region implies that this enzymatic reaction is insensitive to point mutations that alter affinity between substrate and enzyme, that is, the dissociation constant can change by a factor of 10^4 . The second region implies that this enzymatic reaction is sensitive to point mutations as a change in k_r would give an enzymatic reaction rate that is different from the observed data. This joint posterior distribution highlights the challenges with using local sensitivity metrics to infer the importance of a particular mechanistic target within a biological network.^{83,84} A recommended approach would be to identify these distinct regions and assess target sensitivity within each one of the distinct posterior regions. If the target is sensitive to variation that exists within each one of the distinct posterior regions, additional experiments should be designed to eliminate some of these competing regions. If the target is insensitive, additional experiments are not necessary. Collectively, these results highlight how the available data limits the ability to identify parameter values but can still provide high confidence predictions. The posterior distribution in model predictions and their associated distributions in model parameters provide rationale for designing subsequent experiments to discriminate between competing hypotheses that are unresolved given the existing data. The second example helps to illustrate this point.

A few comments on Model Selection

Uncertainty in interpreting observations of the behavior of biological networks naturally leads to alternative hypotheses proposed for the same biological network. Model selection describes a number of qualitative and quantitative criteria used to discriminate among these competing hypotheses. The qualitative criteria speak to whether underlying assumptions in the approach remain valid or to whether the overall behavior of the model makes sense.⁶⁴ For instance, the individual fits of the candidate models to the observed responses should be checked by analyzing the residuals, which should be normally distributed. The predicted

individual responses should also be mutually consistent when a mathematical model is used to synthesize data obtained from different studies together. Quantitative model selection criteria have been developed that enable one to select the model that most closely describes the observed data using concepts drawn from statistics and information theory, as reviewed in⁸⁵. The more simple model selection criteria, like the Akaike Information Criterion (AIC)^{86,87} or Bayesian Information Criterion (BIC)⁸⁸, consider the trade-off between how well the model captures the data, as commonly represented by the logarithm of the maximum likelihood or simply a summed squared error term (C_i), and a penalty associated with the number of parameters (N_i) and the number of experimental observations (N_R):

$$AIC = C_i + 2N_i \quad (14)$$

$$BIC = C_i + N_i \cdot \log(N_R) \quad (15)$$

The data favor the model with the lowest value for either criteria. While the AIC and BIC criteria have sound foundations in statistics and information theory and are easily calculated, the penalty terms are considered ad hoc.⁸⁹ There is also no way to account for model uncertainty, which makes it difficult to select among similarly scoring models. To address these concerns, a Bayes factor has been proposed as a more general approach for model selection.⁸⁹ The Bayes Factor (B_{ij}) quantifies the strength of evidence that favors model i over model j :

$$B_{ij} = \frac{P(Y|M_j)}{P(Y|M_i)} \quad (16)$$

Values for the B_{ij} between 1 to 3 are considered weak, between 3 to 20 are positive, between 20 to 150 are strong, and greater than 150 are very strong evidence favoring model i over model j . While these are a sampling of the quantitative criteria proposed for model selection, there is substantial danger in employing these methods in isolation to more complicated problems associated with inference of biological networks.

To paraphrase George E. P. Box⁹⁰, it is important to remember that, in some regards, all models are wrong but some may be useful for enabling one to think more clearly about the dynamic relationships among components of biological networks. Inevitably, developing a model of a system involves abstraction, where key elements that are thought to be important in governing system behavior and their interactions are included while other components are left out to minimize confounding influences. The data inform the strength of the interactions included in the model. For instance, a common framework for modeling cell signaling networks is to assume that signaling events occur in the context of an average cell of constant volume. This framework neglects the impact of cell proliferation, which can reduce the concentration of activated species in the system through dilution alone. In such case, modeling interactions that regulate the activity of signaling proteins under conditions where cell proliferation may be important may lead to incorrect conclusions. Whether a component and its associated interactions are included or left out places conceptual boundaries on the

specific questions that can be asked of a model. Despite the fact that a model may pass these model selection criteria, the value of model-based inference depends clearly on understanding these conceptual boundaries.

Practical use of $P(Y_M|M)$

Using the set of parameter values that correspond to the converged segment of the Markov chain, the expected value for some property of a model can be calculated from the following integral:

$$P(f(\Theta)|M) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\Theta) \cdot P(\Theta|M, Y) \cdot P(Y) d\Theta dY, \quad (17)$$

where $f(\Theta)$ is a generic function of the model parameters.⁸ The property of the model may be a distribution in a rate parameter, the sensitivity of a variable to changes in a parameter value, or the flow of information down a specific branch of a signal network, which can be thought of as a pathway flux. Generally, these results can be used for selecting among competing models or targeted experimental design.⁹¹ As illustrated in Klinke et al., a posterior distribution in the pathway flux can be used to answer the question: what fraction of the observed gene expression is caused by one of two parallel pathways.⁹² In this study, a mechanistic model was constructed with two competing pathways for both the regulation of TNF α production and the regulation of IL12R β 2 expression (see Figure 10). A high content assay was used to quantify dynamic changes in Interleukin-12 (IL-12) receptor abundance, transcription factor activation, cytokine production, and the fate of cells in the system to account for cell proliferation and death. The AMCMC approach was used to obtain rate parameters that give predictions consistent with the observed experimental data. Using these rate parameters, pathway fluxes down each of the competing pathways are represented by a predicted increase in mRNA in response to activation of an upstream transcription factor. The posterior distributions in pathway flux for these competing pathways can be obtained using parameter sets sampled from the converged segment of the Markov chains. In the case of TNF α production, the *in silico* model-based inference suggests that the autocrine feedback loop regulates TNF α production within this cell system. As summarized in Figure 10, the joint posterior distribution in pathway flux between the RP5 and RP6 branches favored the autocrine feedback pathway as the distribution was primarily below the diagonal. Validation experiments where an antibody that inhibited the autocrine feedback loop was added confirmed the TNF α pathway flux predictions.

The most exciting phrase to hear in science, the one that heralds the most discoveries, is not “Eureka!” (I found it!) but “That’s funny...”

- Isaac Asimov

As illustrated by this quote, the engine for scientific discovery is not when data conforms to our preconceived model (i.e., “I found it!”) but when it disagrees with our model (i.e., “That’s funny...”). A discrepancy between our model and an observation identifies flaws in our current understanding, the first step in meaningful learning. The second example of

⁸As not all of the parameters may exhibit bounded distributions, the evaluation of this integral may not be proper. While this is a risk, the sampled values may be sufficient to determine whether a competing hypothesis can still explain the available data.

evaluating these competing hypotheses illustrates how model-based inference can be used to identify flaws in our current understanding. In the case of IL12R β 2 expression, the data obtained were insufficient to discriminate between the indirect IFN γ -mediated autocrine feedback loop or the direct intracellular STAT4-mediated feedback loop. This is highlighted by the fact that the posterior distributions in the model predictions exhibit narrow distributions while the joint posterior distribution in pathway flux between the RP1 and RP2 branches exhibits a complicated structure. In this joint posterior distribution in pathway flux, point A corresponds to a model where the regulation of the IL-12 receptor occurs through an indirect IFN γ -mediated feedback loop and flux through the direct intracellular STAT4-mediated feedback loop is negligible. In contrast, point B corresponds to a model where the regulation occurs through the direct intracellular STAT4-mediated feedback loop and flux through the indirect IFN γ -mediated feedback loop is negligible. The break in the joint posterior distribution implies that there is an intermediate point where RP1 has a slight preference over RP2 that provides predictions with a lower posterior probability and indicates that the adaptive M-H algorithm was able to traverse this lower posterior probability manifold. Thinking in terms of a Bayes ratio, these results suggest that there is not strong evidence to support either RP1 or RP2 as the predominant pathway as the posterior distributions in the model predictions essentially overlap. By finding that the evidence can not discriminate between these competing hypotheses, we were forced to revisit our experimental design to consider biases, such as the timing of observation or the particular cell model, that may have influenced prior work.

To rule out one of these mechanisms, we designed additional experiments. In particular, we changed three aspects. First, we measured the phosphorylation of STAT1, which was originally included in the model but not measured, in addition to STAT4 phosphorylation. We originally assumed that STAT1 is phosphorylated by an autocrine or paracrine positive feedback loop through secreted IFN- γ . Second, we designed an experiment to test for the presence of an autocrine or paracrine signaling loop by stimulating the cells with IFN- γ and changing the cell density. In our original experimental design, we measured STAT4 phosphorylation at 2 hours following adding IL-12 but the phosphorylation of STAT1 is measured typically within minutes of adding a stimulus. Our third aspect was that we changed the experimental time points to sample faster dynamics within the system. Collectively, these additional data suggest that IL12R β 2 expression in both this T helper cell model and primary splenocytes is regulated by an intracellular mechanism mediated by STAT1 and that STAT1 and STAT4 are both directly activated but are differentially regulated upon IL-12 stimulation. IL-12 mediated activation of STAT1 to induce production of more IL12R β 2 is a positive feedback loop. The observed dynamics suggest that IL-12 mediated activation of STAT4 induces the production of a protein that negatively regulates the activity of STAT1, thereby creating a negative feedback loop. An interlocked positive and negative feedback loop is a common network motif in biology.⁴¹ Conceptually, the competition between positive and negative feedback loops can give rise to a variety of dynamic behaviors that elicit different cellular responses, including oscillations, switching behavior, and single transient bursts of signaling activity.⁵ Work is on-going to identify the mechanistic details of this interlocked positive and negative feedback loop. Based upon the prior knowledge, the direct and differential activation of STAT1 and STAT4 by IL-12 is an

example of a lurking mechanism, as it was not part of the original model. This observation provides a more nuanced view of the signal transduction by these two cytokines relative to the canonical view of STAT1 as associated with IFN- γ signaling and STAT4 as associated with IL-12 signaling. This discovery would have been missed using a conventional model-based inference approach, where rejection of the model is based on whether the model fits the data. These results highlight that a preconceived model encoded with either one of the alternatives can be fit to the observed data. However, evaluating these competing hypotheses using *in silico* model-based inference identified flaws in our current understanding of this signaling network.

Conclusion

While Frequentist and Bayesian methods are frequently presented as mutually exclusive, a pragmatic approach for biological network research is suggested for contemporary practice, as echoed in related work.⁹³ Conventional Frequentist methods are particularly useful for exploratory analysis of large data sets, where the objective is to identify a short list of mechanistic candidates for in depth study. While the ease of using conventional methods is attractive, one must not forget the underlying assumptions when drawing inductive inferences. Given the weaknesses associated with conventional model-based inference, a Bayesian *in silico* model-based inference framework is proposed as a contemporary alternative to the classical hypothesis testing originally developed in the mid-1900's and to improve discovery at the frontier of biological network research. Biological network research - such as understanding the dynamics and crosstalk associated with intracellular signaling pathways, gene regulatory networks (e.g.,⁹⁴), or intercellular cues that regulate multi-cellular behavior (e.g.,⁹⁵) - involves testing postulated causal mechanisms against observations of dynamic system response. Relative to the mid-1900's, contemporary knowledge of the elements involved in biological networks has exploded and the tools available to observe changes in biological response may be unimaginable. The strength of Bayesian *in silico* model-based inference is the ability to encode prior knowledge of the biological networks in the form of a causal model and interrogate larger data sets using a common perspective. This common perspective aids in identifying conditions where the biological response is inconsistent with our postulated model, that is our best ideas fail to explain the behavior. Identifying failures in our understanding leads to the discovery of lurking mechanisms. Moreover, these *in silico* model-based inference methods enable one to reason about signaling networks within a probabilistic framework despite the fact that estimates for many of the parameter values are unknown. The computational burden of the approach and the difficulty in traversing parameter space to obtain a representative sample of the posterior distribution in the model predictions are the main challenges associated with *in silico* model-based inference. The quality control metrics summarized here aim to improve confidence in the *in silico* model-based inference results. Two examples drawn from studies related to cellular signaling were used to illustrate the approach. Future efforts should focus on integrating these *in silico* model-based inference methods with existing software infrastructure for data-storage and model-generation and on improving the computational efficiency and transparency to enable wider adoption of the methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The author thanks A. Yates and J. Kaiser for comments on the manuscript. This work was supported by grants from the National Science Foundation (CAREER 1053490) and the National Cancer Institute (R15CA132124). The content is solely the responsibility of the author and does not necessarily represent the official views of the NSF, the National Cancer Institute, or the NIH.

Literature Cited

1. Hansen J, Iyengar R. Computation as the mechanistic bridge between precision medicine and systems therapeutics. *Clin Pharmacol Ther.* Jan; 2013 93(1):117–128. [PubMed: 23212109]
2. Klipp E, Wade RC, Kummer U. Biochemical network-based drug-target prediction. *Curr Opin Biotechnol.* Aug; 2010 21(4):511–516. [PubMed: 20554441]
3. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature.* 1998; 391:806–811. [PubMed: 9486653]
4. O'Garra A, Gabrysova L, Spits H. Quantitative events determine the differentiation and function of helper t cells. *Nat Immunol.* 2011; 12(4):288–294. [PubMed: 21423225]
5. Purvis JE, Lahav G. Encoding and decoding cellular information through signaling dynamics. *Cell.* Feb; 2013 152(5):945–956. [PubMed: 23452846]
6. Klinke DJ, Ustyugova IV, Brundage KM, Barnett JB. Modulating temporal control of nf-kappab activation: implications for therapeutic and assay selection. *Biophys J.* 2008; 94(11):4249–4259. [PubMed: 18281385]
7. Sorger PK. A reductionist's systems biology: opinion. *Curr Opin Cell Biol.* 2005; 17(1):9–11. [PubMed: 15661513]
8. Wang CC, Bajikar SS, Jamal L, Atkins KA, Janes KA. A time- and matrix-dependent TGFBR3-JUND-KRT5 regulatory circuit in single breast epithelial cells and basal-like premalignancies. *Nat Cell Biol.* Apr; 2014 16(4):345–356. [PubMed: 24658685]
9. Fiering S, Northrop JP, Nolan GP, Mattila PS, Crabtree GR, Herzenberg LA. Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the t-cell antigen receptor. *Genes and Development.* 1990; 4(10):1823–1834. [PubMed: 2123468]
10. Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK. Non-genetic origins of cell-to-cell variability in trail-induced apoptosis. *Nature.* 2009; 459(7245):428–432. [PubMed: 19363473]
11. Popper, KR. *The Logic of Scientific Discovery.* Basic Books; New York: 1959.
12. Kuhn, TS. *The Structure of Scientific Revolutions.* Univ of Chicago Press; Chicago: 1962.
13. Lander AD. The edges of understanding. *BMC Biology.* 2010; 8:40. [PubMed: 20385033]
14. National Research Council (U.S.). Committee on Learning, Research, Practice and Education. *How People Learn: brain, mind, experience, and school.* National Academies Press; Washington, DC: 2000.
15. Saez-Rodriguez J, Goldsipe A, Muhlich J, Alexopoulos LG, Millard B, Lauffenburger DA, Sorger PK. Flexible informatics for linking experimental data to mathematical models via DataRail. *Bioinformatics.* 2008; 24:840–847. [PubMed: 18218655]
16. Kitano H, Funahashi A, Matsuoka Y, Oda K. Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol.* 2005; 23:961–966. [PubMed: 16082367]
17. Matsuoka Y, Ghosh S, Kikuchi N, Kitano H. Payao: a community platform for SBML pathway model curation. *Bioinformatics.* 2010; 26:1381–1383. [PubMed: 20371497]
18. Patil S, Pincas H, Seto J, Nudelman G, Nudelman I, Sealfon SC. Signaling network of dendritic cells in response to pathogens: a community-input supported knowledgebase. *BMC Syst Biol.* 2010; 4:137. [PubMed: 20929569]

19. Fink MY, Pincas H, Choi SG, Nudelman G, Sealfon SC. Research resource: Gonadotropin-releasing hormone receptor-mediated signaling network in LbetaT2 cells: a pathway-based web-accessible knowledgebase. *Mol Endocrinol*. 2010; 24:1863–1871. [PubMed: 20592162]
20. Funahashi A, Jouraku A, Matsuoka Y, Kitano H. Integration of CellDesigner and SABIORK. In *Silico Biol (Gedrukt)*. 2007; 7:81–90.
21. Zinovyev A, Viara E, Calzone L, Barillot E. BiNoM: a Cytoscape plugin for manipulating and analyzing biological networks. *Bioinformatics*. 2008; 24:876–877. [PubMed: 18024474]
22. Drager A, Hassis N, Supper J, Schroder A, Zell A. SBMLsqueezer: a CellDesigner plug-in to generate kinetic rate equations for biochemical networks. *BMC Syst Biol*. 2008; 2:39. [PubMed: 18447902]
23. Zhou Y, Liepe J, Sheng X, Stumpf MP, Barnes C. GPU accelerated biochemical network simulation. *Bioinformatics*. 2011; 27:874–876. [PubMed: 21224286]
24. Salwinski L, Eisenberg D. In silico simulation of biological network dynamics. *Nat Biotechnol*. 2004; 22:1017–1019. [PubMed: 15235611]
25. Bodenmiller B, Wanka S, Kraft C, Urban J, Campbell D, Pedrioli PG, Gerrits B, Picotti P, Lam H, Vitek O, Brusniak MY, Roschitzki B, Zhang C, Shokat KM, Schlapbach R, Colman-Lerner A, Nolan GP, Nesvizhskii AI, Peter M, Loewith R, von Mering C, Aebersold R. Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci Signal*. 2010; 3:rs4. [PubMed: 21177495]
26. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA, Brunak S, Mann M. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal*. 2010; 3:ra3. [PubMed: 20068231]
27. Jacob LS, Wu X, Dodge ME, Fan CW, Kulak O, Chen B, Tang W, Wang B, Amatruda JF, Lum L. Genome-wide RNAi screen reveals disease-associated genes that are common to Hedgehog and Wnt signaling. *Sci Signal*. 2011; 4:ra4. [PubMed: 21266715]
28. Gong Y, Ogunniyi AO, Love JC. Massively parallel detection of gene expression in single cells using subnanolitre wells. *Lab Chip*. 2010; 10:2334–2337. [PubMed: 20686711]
29. Bendall SC, Simonds EF, Qiu P, Amir elAD, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sachs K, Pe'er D, Tanner SD, Nolan GP. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science*. 2011; 332:687–696. [PubMed: 21551058]
30. Neyman J, Pearson ES. On the problem of the most efficient tests of statistical hypotheses. *Phil Trans Royal Soc A*. 1933; 231:289–337.
31. Fisher RA. The logic of inductive inference. *J Roy Stat Soc*. 1935; 98:39–82.
32. Pratt JW. Book review: Testing statistical hypotheses by lehmann. *J Am Stat Assoc*. 1961; 56:163–166.
33. Jaynes, ET. Probability theory: the logic of science. 1. Cambridge University Press; Cambridge, UK: 2003.
34. Bayes T. An essay towards solving a problem in the doctrine of chances. 1763. *MD Comput*. 1991; 8:157–171. [PubMed: 1857193]
35. Stigler SM. Laplace's 1774 memoir on inverse probability. *Stat Sci*. 1986; 1:359–378.
36. Elfenbein HA, Ambady N. On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychol Bull*. Mar; 2002 128(2):203–235. [PubMed: 11931516]
37. Nuzzo R. Scientific method: statistical errors. *Nature*. Feb; 2014 506(7487):150–152. [PubMed: 24522584]
38. Platt JR. Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*. 1964; 146:347–353. [PubMed: 17739513]
39. Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*. 1970; 226:1209. [PubMed: 4316300]
40. Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*. 1970; 226:1211. [PubMed: 4316301]
41. Alon, U. Introduction to systems biology: design principles of biological circuits. 1. Chapman & Hall/CRC; Boca Raton, FL: 2007.

42. Beard DA, Kushmerick MJ. Strong inference for systems biology. *PLoS Comput Biol.* 2009; 5:e1000459. [PubMed: 19714210]
43. Costa KD, Kleinstein SH, Hershberg U. Biomedical model fitting and error analysis. *Sci Signal.* 2011; 4:tr9. [PubMed: 21954296]
44. Cedersund G, Roll J. Systems biology: model based evaluation and comparison of potential explanations for given biological data. *FEBS J.* 2009; 276:903–922. [PubMed: 19215297]
45. Banga JR. Optimization in computational systems biology. *BMC Syst Biol.* 2008; 2:47. [PubMed: 18507829]
46. Jacquez JA, Perry T. Parameter estimation: local identifiability of parameters. *American Journal of Physiology.* 1990; 258:E727–E736. [PubMed: 2333964]
47. Audoly S, Bellu G, D'Angio L, Saccomani MP, Cobelli C. Global identifiability of nonlinear models of biological systems. *IEEE Transactions on Biomedical Engineering.* 2001; 48:55–65. [PubMed: 11235592]
48. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmuller U, Timmer J. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics.* Aug; 2009 25(15):1923–1929. [PubMed: 19505944]
49. Shah NP, Nicoll JM, Nagar B, Gorre ME, Paquette RL, Kuriyan J, Sawyers CL. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell.* 2002; 2:117–125. [PubMed: 12204532]
50. Haken, H. *Advanced Synergetics.* 1. Springer-Verlag; New York, NY: 1983.
51. Klinke DJ, Finley SD. Timescale analysis of rule-based biochemical reaction networks. *Biotechnol Prog.* 2012; 28:33–44. [PubMed: 21954150]
52. Calderhead B, Girolami MA. Statistical analysis of nonlinear dynamical systems using differential geometric sampling methods. *Interface Focus.* 2011; 1(6):821–835. [PubMed: 23226584]
53. Birtwistle MR, Hatakeyama M, Yumoto N, Ogunnaike BA, Hoek JB, Kholodenko BN. Ligand-dependent responses of the erbb signaling network: experimental and modeling analyses. *Mol Syst Biol.* 2007; 3(144):1–16.
54. Klinke DJ. An age-structured model of dendritic cell trafficking in the lung. *Am J Physiol Lung Cell Mol Physiol.* Nov; 2006 291(5):L1038–1049. [PubMed: 17030902]
55. Brown KS, Sethna JP. Statistical mechanical approaches to models with many poorly known parameters. *Phys Rev E.* 2003; 68(2)
56. Brown KS, Hill CC, Calero GA, Myers CR, Lee KH, Sethna JP, Cerione RA. The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Phys Biol.* 2004; 1:184–195. [PubMed: 16204838]
57. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol.* Oct; 2007 3(10):1871–1878. [PubMed: 17922568]
58. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface.* 2009; 6(31):187–202.
59. Toni T, Stumpf MP. Simulation-based model selection for dynamical systems in systems and population biology. *Bioinformatics.* 2010; 26:104–110. [PubMed: 19880371]
60. Erguler K, Stumpf MP. Practical limits for reverse engineering of dynamical systems: a statistical analysis of sensitivity and parameter inferability in systems biology models. *Mol Biosyst.* 2011; 7:1593–1602. [PubMed: 21380410]
61. Vysheirsky V, Girolami MA. Bayesian ranking of biochemical system models. *Bioinformatics.* 2008; 24(6):833–839. [PubMed: 18057018]
62. Pearl, J. *Causality: models, reasoning, and inference.* 1. Cambridge University Press; Cambridge, UK: 2000.
63. Klinke DJ. An empirical bayesian approach for model-based inference of cellular signaling networks. *BMC Bioinform.* 2009; 10:371.

64. Box GEP, Draper NR. The bayesian estimation of common parameters from several responses. *Biometrika*. 1965; 52:355–365.
65. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions (with discussion). *Stat Med*. Nov; 2009 28(25):3049–3082. [PubMed: 19630097]
66. Vysheirsky V, Girolami M. BioBayes: a software package for Bayesian inference in systems biology. *Bioinformatics*. Sep; 2008 24(17):1933–1934. [PubMed: 18632751]
67. Chen Y, Lawless C, Gillespie CS, Wu J, Boys RJ, Wilkinson DJ. CaliBayes and BASIS: integrated tools for the calibration, simulation and storage of biological simulation models. *Brief Bioinformatics*. May; 2010 11(3):278–289. [PubMed: 20056731]
68. Eydgahi H, Chen WW, Muhlich JL, Vitkup D, Tsitsiklis JN, Sorger PK. Properties of cell death models calibrated and compared using Bayesian approaches. *Mol Syst Biol*. 2013; 9:644. [PubMed: 23385484]
69. Geyer, CJ.; Johnson, LT. Technical report. 2012. R package mcmc (Markov Chain Monte Carlo).
70. Gelman, A.; Carlin, JB.; Stern, HS.; Rubin, DB. *Texts in Statistical Science*. Chapman and Hall; Boca Raton, FL: 2004. *Bayesian Data Analysis*.
71. *Handbooks of Modern Statistical Methods*. Chapman and Hall/CRC; Boca Raton, FL: 2011. *Handbook of Markov Chain Monte Carlo*.
72. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys*. 1953; 21:1087–1092.
73. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*. 1970; 57:97–109.
74. Haario H, Saksman E, Tamminen J. An Adaptive Metropolis Algorithm. *Bernoulli*. 2001; 7:223–242.
75. Slezak DF, Suarez C, Cecchi GA, Marshall G, Stolovitzky G. When the optimal is not the best: Parameter estimation in complex biological models. *PLoS ONE*. 2010; 5(10):e13283. [PubMed: 21049094]
76. Jasra A, Stephens DA, Holmes CC. Population-based reversible jump Markov chain Monte Carlo. *Biometrika*. 2007; 94:787–807.
77. Hukushima K, Nemoto K. Exchange Monte Carlo method and application to spin glass simulations. *J Phys Soc Japan*. 1996; 65:1604–1608.
78. Rigat F, Mira A. Parallel hierarchical sampling: A general-purpose interacting Markov chains Monte Carlo algorithm. *Comput Stat Data Anal*. 2012; 56:1450–1467.
79. Gelman A, Rubin DB. Inference form iterative simulation using multiple sequences. *Statistical Science*. 1992; 7(4):457–474.
80. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. 1998; 7(4):434–455.
81. Vanlier J, Tiemann CA, Hilbers PA, van Riel NA. An integrated strategy for prediction uncertainty analysis. *Bioinformatics*. Apr; 2012 28(8):1130–1135. [PubMed: 22355081]
82. Hug S, Raue A, Hasenauer J, Bachmann J, Klingmuller U, Timmer J, Theis FJ. High-dimensional Bayesian parameter estimation: Case study for a model of JAK2/STAT5 signaling. *Math Biosci*. Dec; 2013 246(2):293–304. [PubMed: 23602931]
83. Schoeberl B, Pace EA, Fitzgerald JB, Harms BD, Xu L, Nie L, Linggi B, Kalra A, Paragas V, Bukhalid R, Grantcharova V, Kohli N, West KA, Leszczyniecka M, Feldhaus MJ, Kudla AJ, Nielsen UB. Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Sci Signal*. 2009; 2(77):ra31. [PubMed: 19567914]
84. Faratian D, Goltsov A, Lebedeva G, Sorokin A, Moodie S, Mullen P, Kay C, Um IH, Langdon S, Goryanin I, Harrison DJ. Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab. *Cancer Res*. Aug; 2009 69(16):6713–6720. [PubMed: 19638581]
85. Burnham, KP.; Anderson, DR. *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach*. 2. Springer; New York: 2002.
86. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Control*. 1974; 19:716–723.

87. Yamaoka K, Nakagawa T, Uno T. Application of akaike's information criterion (aic) in the evaluation of linear pharmacokinetic equations. *J Pharmacokinet Biopharm.* 1978; 6:165–175. [PubMed: 671222]
88. Schwarz G. Estimating the dimension of a model. *Ann Stats.* 1978; 6:461–464.
89. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* 1995; 90:773–795.
90. Box, GEP.; Draper, NR. *Empirical Model Building and Response Surfaces.* John Wiley and Sons, Inc; New York: 1987.
91. Vanlier J, Tiemann CA, Hilbers PA, van Riel NA. A Bayesian approach to targeted experiment design. *Bioinformatics.* Apr; 2012 28(8):1136–1142. [PubMed: 22368245]
92. Klinke DJ, Cheng N, Chambers E. Quantifying crosstalk among Interferon-gamma, Interleukin-12 and Tumor Necrosis Factor signaling pathways within a Th1 cell model. *Sci Signal.* 2012; 5:ra32. [PubMed: 22510470]
93. Raue A, Kreutz C, Theis FJ, Timmer J. Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. *Philos Trans A Math Phys Eng Sci.* Feb. 2013 371(1984):20110544. [PubMed: 23277602]
94. Gonzalez J, Vujacic I, Wit E. Inferring latent gene regulatory network kinetics. *Stat Appl Genet Mol Biol.* Mar; 2013 12(1):109–127. [PubMed: 23744300]
95. Breto C, He DH, Ionides EL, King AA. Time series analysis via mechanistic models. *Ann Appl Stat.* 2009; 3(1):319–348.

Figure 1. Scientific method is an iterative process comprised of three steps: formulating a hypothesis (model) based on prior information, obtaining data that informs the model, and testing the model against the data

Formulating a hypothesis involves integrating observed events, which is expressed as existing data, with evolving knowledge of a system, which is expressed as our prior knowledge. How these observed events integrate with existing knowledge forms the basis of the hypothesis. Recent computational advances have made it easier to compare existing data against our current knowledge and to formulate new models. Experiments are designed to test the formulated model but also include implicit constraints that may bias the conclusions of the study. These implicit biases may include selecting a model system, relevant time points, and metrics of system response. Recent experimental advances have enabled obtaining more information about system response; yet, this information is still limited. The inference step involves testing the model against the data to assess within a statistical framework whether the data are consistent with the model (i.e., the model is not invalid). If the data are inconsistent with the model (i.e., the model is invalid), a new model may be proposed that reflects the prior knowledge and the available data - both pre-existing and newly acquired data.

Figure 2. A two-dimensional logical universe

The universe is comprised of an observable event dimension and an unobservable cause dimension. Feasible regions are unshaded (e.g., $P(M_1|Y_1) > 0$) while unfeasible regions are shaded (e.g., $P(M_3|Y_3) = 0$). Probability is proportional to the conditioned area of unshaded regions.

Figure 3. Conditioning the universe based upon perspective

The universe can be conditioned based upon observing an event (Y_i) or conditioned based upon cause M_i but the joint probability is the same irrespective of perspective. (A) The hashed region equals the product of the probability of M_2 , given that event Y_2 was observed, and the probability of event Y_2 to occur ($P(M_2|Y_2) \cdot P(Y_2)$). (B) The hashed region equals the product of the probability of Y_2 , given that cause M_2 is known, and the probability that M_2 is a possible cause ($P(Y_2|M_2) \cdot P(M_2)$).

Figure 4. Conventional model-based inference using a hypothetical cell signaling example
 (A) A model of the “B” signaling pathway is depicted as a cartoon. To test the model, a drug (D) was used to inhibit the activation of protein A following binding of the ligand (L_B) to the receptor (R_B). (B) In conventional *in cerebro* model-based inference, the system is observed at a single time point and a one-sided Z-test is used to assess whether the addition of the drug decreases the activation of protein A. (C) For conventional *in silico* model-based inference, the cartoon is converted into a mathematical model. The mathematical model is calibrated to data obtained at multiple time points. The calibration data is depicted as symbols while the model predictions are shown as lines. Results for the ligand stimulation only (triangles and solid line), ligand stimulation in the presence of a drug inhibitor (+ and dotted line), and drug inhibitor alone (circles and dotted line) are shown separately. A Chi-squared test is used to assess whether the mathematical models fit the data.

Figure 5. Synthetic data was generated for a simple enzymatic reaction

(A) Three elementary reactions were used to represent the enzymatic conversion of a substrate (S) into a product (P) by an enzyme (E) through an intermediate enzyme-substrate complex (C). Synthetic data were generated for two different initial concentrations of substrate (Panel B: “Low” = 10^{-3} mM and Panel C: “High” = 3×10^{-3} mM). Model predictions (solid lines) were generated using values for the total enzyme concentration and three rate constants to be 10^{-5} mM (E_o), 10^2 mM $^{-1}$ sec $^{-1}$ (k_f), 10^{-1} sec $^{-1}$ (k_r), and 10^{-2} sec $^{-1}$ (k_{cat}). Synthetic data (squares) were generated by sampling the model predictions at 7 time points and adding random noise with a root mean square deviation equal to 10% of the maximum substrate concentration in triplicate at each time point.

Figure 6. Traces of the Markov chains illustrate a random walk in parameter space that is conditioned on the data

An adaptive MCMC algorithm was used to generate three chains (shown in red, black, and blue) that traverse a 5-dimensional parameter space. Traces of the likelihood values (A), low initial substrate concentration (B), high initial substrate concentration (C) and the rate parameters (D: k_f , E: k_r , F: k_{cat}) for each step in the Markov chain are shown in different colors (red, blue, and black). Note that the Y-axis range for the traces of the initial substrate concentrations is different than the rate parameters. Results obtained during the “burn in” period are not shown.

Figure 7. The adaptive Metropolis-Hastings algorithm changed the proposal distribution based on the cumulative samples

(A) Eigenvalues quantify the magnitude of the orthogonal dimensions of a 5-dimensional proposal distribution. Each of the lines represent traces of the five eigenvalues for one of the chains. The overall volume of the proposal distribution (B) and the acceptance ratio (C) is shown as a function of MCMC step for each of the three chains, shown in red, blue, and black, respectively.

Figure 8. The Markov chains relatively quickly converge and provide samples from the posterior distribution in the model predictions

(A) Convergence of the model predictions for each of the two initial substrate concentrations (left panel - low concentration; right panel - high concentration) was estimated using the Gelman-Rubin statistic (GRS), as indicated by colorbar. (B) Using converged segments of the three Markov Chains, posterior distributions in the model predictions for each of the chains were compared against the experimental data (squares). For each of the three colored chains, the solid line indicates the median while the dotted lines enclose 95% of the posterior distribution of the model predictions.

Figure 9. Converged segments of the Markov chains provide bounded estimates for only a subset of the model parameters

The converged segments of the three Markov chains were used to estimate the posterior distributions in the initial substrate concentrations (A) and rate constants (B). In (A), the red and blue curves correspond to distributions in estimates of the low and high initial substrate concentrations, respectively. In (B), the distributions in rate parameters, which are labeled on the diagonal, are represented by bivariate scatter plots, which are shown below the diagonal, and correlation coefficients, which are shown above the diagonal. The density of points, which represent points in parameter space sampled by the Markov chain, in a scatter plot are color coded whereby blue and yellow correspond to low and high density, respectively.

Figure 10. *In silico* model-based inference applied to study crosstalk among cytokine signaling pathways in a T helper cell model

A signaling cartoon presented in the center summarizes a mathematical model that was constructed to gain insight into the signaling pathways that were activated by the cytokine IL-12 in a T helper cell model. A high content assay was used to quantify dynamic changes in IL-12 receptor abundance, signaling protein activation, cytokine production, and cell population changes. A sampling of this calibration data is shown as graphs below and to the left side of the model cartoon. The red arrows indicate how the experimental data map onto elements of the mechanistic model. The mechanistic model incorporated competing hypotheses regarding how the IL-12 receptor is regulated in T helper cells, which is either through an extracellular positive feedback loop that acts through Interferon- γ (denoted by RP1) or through a direct intracellular positive feedback loop that acts primarily through STAT4 (denoted by RP2). These two competing pathways are highlighted by the dashed blue box. The model also incorporated competing hypotheses regarding how the production of TNF- α was regulated in T helper cells, which is either by IL-12 through STAT4 (denoted by RP5) or through an autocrine feedback loop (denoted by RP6). These two competing pathways are highlighted by the dotted blue box. An AMCMC method was used to estimate the posterior distribution in the model predictions (lines in calibration graphs). The posterior distributions in model predictions are summarized by a median (solid line) and dotted lines enclose 95% of the distribution. The resulting set of parameter values, that is a random sample from the posterior distribution in the model parameters, were used to estimate the joint posterior distribution in the flow of cellular information between the two competing pathways. The blue arrows point to the corresponding pathway flux panels (right side). In comparing the predicted flow of molecular information between these two competing pathways, the diagonal lines indicate where the signaling flux is predicted to be equal between the RP1 and RP2 pathways and RP5 and RP6 pathways, respectively. Points A and B indicate parameter values that favor one pathway over the other. Point A indicates RP1 is the predominant pathway while point B indicates that RP2 is the predominant pathway.