Computational models of the JAK1/2-STAT1 signaling

Anna Gambin\textsuperscript{1,2,*}, Agata Charzyńska\textsuperscript{3}, Aleksandra Ellert-Miklaszewska\textsuperscript{4} and Mikołaj Rybiński\textsuperscript{2,5}

\textsuperscript{1} Institute of Informatics, University of Warsaw, Warsaw, Poland
\textsuperscript{2} Mossakowski Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland
\textsuperscript{3} Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland
\textsuperscript{4} Laboratory of Transcription Regulation, The Nencki Institute of Experimental Biology, Warsaw, Poland
\textsuperscript{5} Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

Keywords: IFN signaling, JAK1/2-STAT1 pathway, computational model, mathematical modeling, kinetics, network inference.

Abbreviations: Epo, Erythropoietin; GAS, Gamma Activation Site; GSA, Global Sensitivity Analysis; IFN, Interferon; IFNGR, Interferon Gamma Receptor; IL, Interleukin; IRF, Interferon Regulatory Factor; ISGF, Interferon-Stimulated Gene Factor; ISRE, Interferon Stimulated Response Element; JAK, Janus Kinase; MAPK, Mitogen-Activated Protein Kinase; MHC, Major Histocompatibility Complex; NK, Natural Killer; ODE, Ordinary Differential Equations; PIAS, Protein Inhibitor of Activated STAT; PSC, Pancreatic Stellate Cells; PTP, Protein Tyrosine Phosphatase; RRE, Reaction Rate Equations; SHP-1/2, Src Homology Region 1/2 Domain-Containing Phosphatase; SOCS, Suppressor of Cytokine Signaling; STAT, Signal Transducer and Activator of Transcription; TNF, Tumor Necrosis Factor; TYK, Tyrosine Kinase.

Abstract
Despite conceptually simple mechanism of signaling, JAK-STAT pathway exhibits considerable behavioral complexity. Computational pathway models are tools to investigate in detail signaling process. They integrate well with experimental studies, helping to explain molecular dynamics and to state new hypotheses, most often about the structure of interactions.

A relatively small amount of experimental data is available for a JAK1/2-STAT1 variant of the pathway, hence, only several computational models were developed. Here we review a dominant approach of kinetic modeling of the JAK1/2-STAT1 pathway, based on ordinary differential equations. We also give a brief overview of attempts to computationally infer topology of this pathway.

1. Introduction

1.1 Computational modeling: mathematical description integrated with experimental data.

Computational modeling of signal transduction integrates available knowledge about pathway regulation, and the general chemical and physical principles with experimental data from different biotechnology platforms. Such approach constitutes a powerful solution for formalizing and extending traditional molecular and cellular biology. However, success of the entire project depends crucially on the said integration. Figure 1 illustrates the model’s life cycle paradigm: the recurring processes of molecular experiments, measurements, bioinformatics analyses, and mathematical modeling.

*Correspondence to Anna Gambin; Email: aniag@mimuw.edu.pl
In this review we sketch existing attempts to model JAK1/2-STAT1 signaling. There are relatively few such attempts for this particular pathway. The limiting factor is the availability of experimental data. As we will see a new set of molecular data usually allows researchers to formulate more adequate model.

Formal modeling of signaling pathways most often starts with a definition of a topology of the pathway, represented by a biochemical reactions network. Construction of a structure of a model requires systematization of a diverse, incomplete, and usually inconsistent biological knowledge about the underlying phenomena. Consequently, the structure of a model itself is a valuable contribution, provided that it is validated against available experimental data. However, computational modeling of signaling pathways offers much more. It can help to understand roles of designed parts of an investigated pathway. Standard examples include comprehension of a function of cascade layers and negative feedback in the MAPK pathway or, in case of the JAK-STAT pathway, role of a nuclear export of STATs and of a post-translational negative feedback.

1.2. JAK-STAT signaling

JAK-STAT pathways are highly conserved intracellular signaling pathways of all vertebrates and even found in some more primitive metazoas. These pathways have co-evolved with multiple fundamental cellular processes, such as innate and adaptive immune response, cell growth, differentiation and apoptosis regulation, maintenance of homeostasis, organogenesis, or embryonic development.

JAK-STAT pathways provide a simple and direct route from the membrane receptors to the nucleus for mediating cellular responses to numerous cytokines. Until now, four receptor-associated JAK tyrosine kinases (JAK1/2/3, TYK2) and seven different STAT family members (STAT1/2/3/4/5A/5B/6) have been identified in mammalian cells. In short, there are three main stages involved in the transduction of cytokine receptor derived signal through any JAK-STAT pathway: activation of the receptor, direct translocation of the signal into the nucleus and expression of target genes. Binding of cytokine triggers changes to the cognate receptor that permit cross-activation of receptor-associated JAKs. Activated JAKs
phosphorylate intracellular tail of the receptor, thereby providing docking sites for latent STAT transcription factors (see Figure 2). Recruited STAT monomeric proteins become subsequently phosphorylated by the proximate JAKs, dissociate from the receptor and form homo- or heterodimers directed to enter to the nucleus. Acting as dimers or as more complex oligomers STATs bind to consensus DNA sequences in the regulatory regions of target genes and initiate or enhance the appropriate transcriptional response.\(^3\)\(^,\)\(^12\)

Each stage of JAK-STAT signaling takes different time to execute - it ranges from seconds up to hours, depending on an organism and a tissue. The magnitude and duration of the transduced signal is tightly controlled and can be attenuated at all levels. The key players here are: the protein tyrosine phosphatases (PTPs; e.g. SHP-1 or SHP-2), which

\[\text{Figure 2: Key steps of JAK-STAT signaling in response to IFNs. IFNs induce activation of JAKs, receptors and subsequently, of cytoplasmic STATs; all by phosphorylation. Active STATs essentially form dimers, which then are translocated to nucleus where they act as transcription factors in the specific gene promoter or enhancer regions. IFN signaling negative regulators include multiple cytoplasmic PTPs, nuclear PTP and PIAS proteins, as well as specific to IFN signaling SOCS1 which in principle inhibits STAT1 phosphorylation process. See text for more details.}\]
dephosphorylate receptors at the membrane and activated STATs in the nucleus;\textsuperscript{4} protein inhibitors of activated STATs (PIAS), which in principle, block STAT dimers DNA binding or transactivation capacity in the nucleus, both directly (PIAS-1/3) or by recruitment of other co–repressor molecules (PIAS-x/y);\textsuperscript{13} and expressed as a negative feedback, some members of the family of suppressors of cytokine signaling (SOCS1/2/3), which inhibit receptor signaling by directly inhibiting both JAKs and cytokine receptors.\textsuperscript{4,14} In principle, these SOCS family members block STAT docking sites on the receptor,\textsuperscript{15} but they also probably mediate STAT degradation by ubiquitin-proteasome pathway.\textsuperscript{14} In addition, the transcriptional potency of STATs can be influenced by other proteins, either through direct interaction on a promoter (e.g. NF-κB) or through post-translational modification, including phosphorylation by growth factor receptors tyrosine kinases or mitogen-activated protein kinases (MAPKs).\textsuperscript{12} In general, JAK-STAT signaling is involved in a broad spectrum of fundamental processes. However, each cytokine receptor is activated by a small set of cytokines, specific only to this receptor. Consequently, cytokine receptor activates a characteristic combination of individual JAKs and STATs that is determined by the structure of the intracellular domains of the receptor chains.\textsuperscript{11}

1.3. Interferons

Interferons – potent regulators of innate immunity – play a key role in the inflammatory response after viral infection and in host defense against microorganisms. Currently, there are three distinct classes recognized: Type I (IFN-α/β/ε/κ/ω), Type II (IFN-γ) and Type III (IFN-λ), distinguishable by their differing target receptors. Types I and II IFN receptors are present on most human cell types whilst type III IFN-λ receptors are highly expressed on hepatocytes but not on microvascular endothelium, adipocytes, fibroblasts or CNS cells.\textsuperscript{16} Types I and III IFN, despite signaling through distinct receptor complexes, stimulate very similar signaling pathways, namely they activate JAK1 and TYK2, which phosphorylate STAT1 and STAT2, respectively. In this pathway variant, the downstream heterotrimer of phosphorylated STAT1, STAT2 and IRF9 is generated to form interferon-stimulated gene factor 3 (ISGF3). ISGF3 acts as a transcription factor within interferon-stimulated response element (ISRE) sites on the promoter or enhancer regions of Type I and III IFN-responsive genes.\textsuperscript{17,18} The focus of this review is modeling of cellular response to Type II IFN (IFN-γ), which signals through the IFN-γ receptor (IFNGR), pre-associated with JAK1 and JAK2. Stimulation of the receptor leads to phosphorylation and dimerization of STAT1. Both IFN signaling schemes, including pathway negative regulators are depicted in Figure 2.

1.4. JAK1/2-STAT1 variant

Active form of IFN-γ receptor is more complex than that for Types I and III IFN. IFN-γ signal transduction engages oligomerization of two subunits IFNGR1 and IFNGR2, which triggers activation of receptor-associated kinases JAK1 and JAK2 and phosphorylation of STAT1 on critical tyrosine residue. Phosphorylated STAT1 translocate as transcriptionally active homodimers into the nucleus, bind to the promoter regions via DNA sequences termed GAS (γ-activated site) and, thereby, affect transcription of IFN-γ-induced genes.\textsuperscript{17} The mechanism that regulates active nuclear import of tyrosine phosphorylated STAT1 dimers was shown to involve a specific carrier, importin α5, and metabolic energy \textsuperscript{19}. It has been suggested that besides the carrier-dependent transport of phospho-dimers, also unphosphorylated STAT1 molecules shuttle continuously between the nucleus and cytoplasm, in a process mediated by direct interaction with nucleoporins located in the nuclear pore complexes.\textsuperscript{20} This constitutive nucleocytoplasmic translocation of unphosphorylated STAT1 occurs both in resting and IFN-γ-stimulated cells. The existence of distinct transport pathways allows the controlled nuclear
transport of phosphorylated STAT1 dimers during cytokine induction. It was demonstrated that nuclear import and nuclear retention are two separate steps leading up to nuclear accumulation of STAT1. DNA binding protects STAT1 from dephosphorylation, so only once released from DNA, the phosphorylated STAT1. Dimers become dephosphorylated by nuclear phosphatases. Dimer dephosphorylation is a critical step, which enables the return of STAT1 to the cytosol to participate in additional rounds of the activation and deactivation cycle.

Recent data have proved that regulation of STAT1 signaling is more complex. There is evidence that unphosphorylated STAT1 also exist as dimers with DNA binding activity and, although less active than phosphorylated forms, may regulate gene expression. Besides, STAT1 is known to form larger assemblies than dimers when bound to DNA, including tetramers and extended polymers, which may significantly influence in vivo binding preferences.

Similarly to other variants of JAK-STAT signal transduction, JAK1/2-STAT1 pathway is negatively regulated by a number of regulatory mechanisms, including activity of various PTPs and the PIAS family, as well as feedback inhibition by SOCSs. However, some of these mechanisms have been recognized as specific or particularly important for attenuation of IFN-γ – induced response.

Among PIAS family members, primarily PIAS1 acts as a partial physiological inhibitor of STAT1. It was originally shown to block STAT1 DNA-binding activity, but due to its E3 SUMO ligase activity it has been suggested to promote SUMO-conjugation to STAT1. Simulation of STAT1, which obstructs tyrosine phosphorylation and affects STAT1 solubility, may be a unique mechanism that has evolved to negatively regulate IFN-γ – induced signaling, however the role of PIAS1 in this process has been controversial and the identity of the E3 SUMO ligase responsible remains unclear.

The most studied inhibitors of JAK-STAT signaling are the SOCS proteins. IFN signaling-specific SOCS1 is known to interact with IFNAR1 receptor and respective TYK2 in Type I IFN signaling and with IFNGR1 receptor and JAK2 in Type II IFN signaling. SOCS1 is itself upregulated via Type II IFN signaling, but not known to be upregulated via Type I or III IFN signaling. Moreover, as revealed by gene-targeting experiments, loss of SOCS1 results in excessive response to IFN-γ and the majority of inflammatory defects exhibited by SOCS1 knockout mice are related to unbridled IFN-γ signaling. Therefore, SOCS1 is recognized as a crucial regulator of IFN-γ induced response.

Whereas Type I IFNs are primarily induced in response to viral infection, IFN-γ is produced predominantly by T lymphocytes, NK and NKT cells, following activation with immune and inflammatory stimuli. IFN-γ acts on its receptor to augment innate cellular immunity by activating NK cells and macrophages, upregulating MHC expression, promoting leucocyte migration and inducing the expression of proinflammatory cytokines (e.g. IL-12 p40 subunit, TNF-α, IL-1) and enzymes involved in elimination of ingested bacterial or protozoan pathogens, such as nitric oxide synthase. Some data also implicates role of IFN-γ as a critical immune system component, protecting the host from development of variety of tumors, e.g. in fibrosarcoma, melanoma, mammary carcinoma models. Except from effects on both the innate and adaptive antitumour immune responses and inhibition of angiogenesis, one of the major targets of IFN-γ’s anti-tumor actions is the tumor cell itself. The molecular basis of this effect is the JAK1/2-STAT1-dependent activation of genes encoding cell cycle inhibitors or implicated in promoting tumor cell apoptosis. Owing to the central role of IFNγ in immune response to microbial infections and in promoting anti-tumor protection, JAK1/2-STAT1 pathway exploited by the cytokine is an excellent target for detailed studies. Because of its relative simplicity and high evolutionally conservation it has attracted attention of systems biologists, who have contributed to mathematical models of the signaling cascade.
Dysregulation of JAK-STAT signaling is associated with various immune disorders and cancers. Identification of key components and steps in signal transduction pathway provides useful strategy for drug discovery. Computational models can greatly facilitate such predictions.

2. Computational models

Most of the experimental data for the JAK-STAT pathways available in the recent literature in vast majority concerns variants activated by ligands other than Type II IFN. This data was used in development of many formal models of JAK-STAT signaling. Most notably, for Epo-stimulated variant, a simple model which originated from work of Swameye et al. This model was re-used multiple times in mainly methodological case studies, with adjustments such as linear chain approximation of a nuclear transport delay. For other examples of computational models of JAK-STAT pathway variants activated by Epo, as well as by IL or by Type I IFN, see respectively ref. or refs. For a broader review see ref. In this review we focus on models of the JAK1/2-STAT1 signaling pathway. There are relatively few distinct computational models for this variant of the JAK-STAT pathway.

Please be aware of the fact that in principle mathematical modeling is not suitable for the very detailed description of reality. It is important to maintain an appropriate balance between the number of mechanistic details responsible for the complexity of the model, and their significance in the context of the data available for model validation process (cf. refs.). The identifiability of model parameters is one of key issues here: non-identifiable model parameters may require more appropriate experimental design, or model reduction, adapting the complexity of the model to the information content of experimental measurements (cf. refs.). Finally, focusing on specific hypotheses while designing the model should also enhance its predictive power (cf. ref.).

2.1. Kinetic modeling

Mathematical models resulting from a biochemical reactions network are complex, even for the simplest signaling pathways, such as JAK-STAT. Therefore, an in silico analysis is usually based on computationally efficient numerical solutions for a deterministic framework which despite of a coarse simplification of a spatial distribution and diffusion effects under spatial homogeneity assumption proves to be very elucidative. In this framework state of a modeled system is represented by the time dependent vector \( \tilde{S}(t) = (S_1(t), ..., S_N(t)) \) denoting concentrations of \( N \) reacting species \( S_1(t), ..., S_N(t) \). Dynamics of the system is governed by a set of a set of first-order ODE, known as the reaction rate equations (RRE), i.e.:

\[
\frac{d\tilde{S}(t)}{dt} = \tilde{f}(\tilde{S}(t)) = C \cdot \tilde{\vartheta}(\tilde{S}(t)),
\]

where the right-hand side of the differential equations \( \tilde{f} = (f_1, ..., f_N) \) is equal to \( M \) state-dependent reactions rates \( \tilde{\vartheta} = (\vartheta_1, ..., \vartheta_M) \) linearly combined according to a fixed stoichiometric matrix of the biochemical reactions network \( C \), i.e.:

\[
C \cdot \tilde{\vartheta}(\tilde{S}(t)) = \begin{pmatrix} c_{1,1} & \cdots & c_{1,M} \\ \vdots & \ddots & \vdots \\ c_{N,1} & \cdots & c_{N,M} \end{pmatrix} \begin{pmatrix} \vartheta_1(\tilde{S}(t)) \\ \vdots \\ \vartheta_M(\tilde{S}(t)) \end{pmatrix},
\]
Stoichiometric matrix directly reflects a structure of the modeled network. Reaction rates, typically, follow the mass action law\textsuperscript{45} or its enzyme kinetics approximations such as the simplest Michaelis-Menten kinetics.\textsuperscript{46} Together with the initial state the dynamics of the system is formulated as an initial value problem. Below we briefly review kinetic models for the JAK1/2-STAT1 signaling. All of them describe dynamics using RRE. Development of stochastic counterparts of these models still merits further exploration.

**Model by Yamada et al.** Arguably, the first model of the JAK1/2-STAT1 pathway was introduced by Yamada et al.\textsuperscript{47} This model describes a control mechanism and the factors influencing the kinetics of JAK/STAT pathway in IFN-\(\gamma\)-stimulated hepatocytes, recapitulating liver injury associated with increased IFN-\(\gamma\) activity. Yamada et al.\textsuperscript{47} model is relatively complex as it captures all essential elements in the JAK1/2-STAT1 signaling, together with many short-lived, intermediate species. These species most likely could be omitted or even systematically eliminated (cf. refs.\textsuperscript{38,48}). Scheme of the model is depicted in Figure 3(a). The model can be informally divided into three modules: receptor module, transcription factor module (the STAT life-cycle) and post-translational feedback module. Panels Figure 2(b-d) show the numerical solutions of RRE for output species of each module. Respective modules emit output signal for ca. 25 min, 50 min and 1.5 h, with its peak activity in ca. 25\textsuperscript{th} minute, 55\textsuperscript{th} minute and 2\textsuperscript{nd} and a half hour.
Figure 3: JAK1/2-STAT1 pathway modules denoted by colors on the scheme of a model by Yamada et al.\
(a), and numerical simulations of modules output species representing concentration of, respectively, phosphorylated receptor dimer (b), phosphorylated nuclear STAT1 dimer (c), and SOCS1 (d) with respect to time. Additional decorations of the simulations graphs represent signaling properties, such as: peak activity time $\tau$, duration $\vartheta$ and amplitude $\alpha$ at the basal, steady state level $\pi$, corresponding to a constant IFN input of 10 nM. In such in-silico experimental setup, unbound active receptor reaches peak in ca. 25 min (b) and from that moment on it is gradually used up in activation of STAT1 proteins. In turn, phosphorylated STAT1 dimers accumulate in nuclei, with maximum concentration reached in ca. 1 hour (c). Their slow concentration descent is followed up by delayed expression of SOCS1 proteins (d), which gradually overtake active receptors (b). The remaining small excess of the latter induces a second, much weaker phase of signaling, starting in ca. 5-th hour (b-d). After that the signal is completely attenuated. Duration of activity of signaling molecules $\nu$ elongates downstream of the pathway, and extent of this effect depends on the strength of signaling $\alpha$ (cf. the first and the second phase of signaling; b-d).
The model by Yamada et al.\cite{Yamada2007} is currently curated in DOQCS\cite{DOQCS} and BioModels databases\cite{BioModels,Douglas2007}, for instance, in the machine-readable SBML format.\cite{SBML} It means that this model is basically ready to be simulated and it has been assured that it replicates published behavior. This is a priceless advantage of this model because it is relatively large. Namely, it consists of over 30 variables representing species and of over 60 parameters of reaction rates. In principle, this model was calibrated and validated against the experimental data.\cite{Yamada2007} Kinetic constants and protein concentrations were set based on experimental results mainly from Brysha et al.\cite{Brysha2007}

There are noticeable quantitative differences in the kinetics of response to IFN-γ depending on cell type, i.e. mouse liver cells\cite{Brysha2007} and T cells data\cite{Bykov2008} but the overall, qualitative behavior of the JAK1/2-STAT1 pathway is consistent with published biological data. Yamada et al.\cite{Yamada2007} used their model to underline importance of nuclear phosphatase PTPs and the negative feedback created by SOCS-1 protein, i.e. importance of the JAK1/2-STAT1 pathway negative regulators. Through in silico knockout experiments nuclear PTP was identified as the most important pathway regulator with respect to attenuation of an unwanted signal which can arise from a non-significant random appearance of IFN, i.e. in case of excessive exposure of cells to the cytokine. Dephosphorylation of STAT1 dimer by nuclear phosphatase is a prerequisite for the transcription factor export to cytoplasm and thus one of the crucial factors in controlling nuclear accumulation of active STAT1 at the site of regulation of gene expression\cite{Davison1999}. On the other hand, mechanism of a post-translational production of SOCS was justified by a requirement of a delay in the activation of the negative feedback. To that end, Yamada et al.\cite{Yamada2007} analyzed hypothetical model, where SOCS was stimulated directly from a downstream of a signal, before the end of transcription and translation of target genes. Available experimental data show that SOCS-1 expression is induced by IFN-γ and overexpression of SOCS-1 inhibits IFN-γ signaling.\cite{Zi2007,Bykov2008}

The only obvious component of the JAK-STAT pathway in general, which is missing in the model, is the PIAS inhibitor. Moreover, this model does not take into consideration some currently accepted regulatory mechanisms affecting STAT1:DNA binding activity, such as STAT1 simulation and transcriptional activity of unphosphorylated STAT1. One of the reasons is obviously the fact, that these processes had not been clearly defined at the time, when the model was created.

**Sensitivity analysis.** The same model was analyzed in several following papers,\cite{Zi2007,Shiu2008,Carli2008} which, to some extent, employed a concept of biological robustness, based on a mathematical technique termed as the sensitivity analysis. Robustness is a property of a system to maintain one or more of its functions under external and internal perturbations.\cite{Krstic2000} On the other hand sensitivity analysis in principle investigates the relation between uncertain parameters of a model and the observable output. For a review of sensitivity analysis methods applicable to chemical reaction networks and signaling pathways in particular see refs.\cite{Stelling2004,Bo觸ε2007} Global sensitivity analysis (GSA) is the most suitable for highly nonlinear models, such as RRE models.\cite{Stelling2004} Based on the mathematical model developed by Yamada et al.\cite{Yamada2007}, by means of GSA of the IFN-γ induced JAK-STAT signaling pathway, Zi et al.\cite{Zi2007} confirmed that SOCS1, and nuclear phosphatase are critical components for the perturbation of the system output and additionally pointed out an importance of cytoplasmic STAT1 for a dynamic behavior of the pathway. The result that the nuclear phosphatase and SOCS1 are more critical than SHP-2 and the cytoplasmic phosphatase underscores the importance of downstream (hence, more direct upon STAT1) negative regulators compared to upstream regulators. The SOCS proteins are generally expressed at low levels in unstimulated cells and become rapidly induced by cytokines, thereby blocking continued signaling and forming a classic negative-feedback loop.\cite{Ruprecht2008} The finding by Zi et al.\cite{Zi2007}, that the nuclear phosphatase (which promotes the nuclear export) is an important signaling component, matches well with the importance of...
nucleocytoplasmic shuttling process of STAT1, described in biological systems. Moreover, according to Swameye et al., who developed a much simpler model of JAK2-STAT5 signaling pathway, STATs nucleocytoplasmic shuttling parameters are the most sensitive to perturbations with respect to the total amount of nuclear activated STAT.

The Yamada et al. model and three hypothetical models, which differ by a form in which STAT protein enters the nucleus and acts as a transcription factor, were compared in work of Shudo et al. The original model was selected as an evolutionary preferable one, by presenting the most reasonable strength and time of the response, as well as lowest local sensitivity of kinetic parameters with respect to the input stimuli. This computational reasoning was used to justify that the STAT1 dimerization is an indispensable signaling step in the generation of cellular response to IFN-γ. This is in line with published data implying that active nuclear import induced by cytokine stimulation as well as subsequent STAT1:DNA binding requires STAT1 dimer formation. However, the model version by Shudo et al. differs from the model by Yamada et al. by lack of a reaction of dimerization of receptor molecules. This reduction was based on the assumption that cytokine receptors are already assembled on the membrane. We studied consequences of this reduction in ref. The receptor activation mechanism is not easily accessible for experimental measurements, thus we carried out a comparative computational analysis of four variants of activation of the JAK1/2-STAT1 pathway. Using GSA, complemented with the profile-likelihood based identifiability analysis, we showed that the on-membrane pre-assembly of the dimers in the absence of ligand increases the overall environmental robustness of the pathway model. In a more general sense, we evaluated the usefulness of different model selection methods in a frequently encountered, but not much discussed case of a model of a considerable size, which has several variants differing at peripheries. With lack of the sufficient experimental data to test against, the generalizability parsimony principle turned out not to be a sufficient criteria for model selection. The sensitivity analysis based on the robustness concept enabled more conscious, expert-mediated choice of the preferred model.

Hierarchical subsystems identification. On the other hand, Soebiyanto et al. applied the multilevel hierarchical systems framework to study the coordination principle in the RRE model of JAK1/2-STAT1 signaling by Yamada et al. Based on the knowledge of biological functionality they divided the model into five subsystems - 4 modules in the cytoplasm and one representing the STAT1 activity in the nucleus. Compartmentalization of the pathway model coupled with in silico inhibition, knockdown/deletion and perturbation experiments identified SOCS1 as a coordinator. Coordinator’s role is to regulate the subsystems at the lower level to achieve the overall objective of the system. In fact, SOCS1 was shown to regulate functionally independent subsystems related to SHP-2 and cytoplasmic STAT1. This discovery of SOCS1 as a coordinator is in line with the recent publication that shows biological data supporting SOCS1 as a crucial signaling component regulator. The model by Soebiyanto also showed that a SOCS1 knockdown leads to high-level activation behavior of the pathway, while SHP2 knockdown results in its constitutive, though pathological, activation. These in silico experimental data agree with the results of STAT1 phosphorylation with SOCS1 knockdown shown by Yamada et al. and are consistent with Brysha et al. Moreover, it suggests that with abundance of IFN-γ, as it occurs during inflammation, STAT1 can be persistently activated if a SOCS1 mutation/knockdown exists. Without the high availability of cytokines, it is SHP2 that is regulating the constitutive activation of STAT1.

Identifiability of the model. Finally, Quaiser et al. also analyzed the model by Yamada et al., however, focusing on the mathematical problem of model identifiability. Authors proposed an iterative method for model simplification, and demonstrated its use by
simplifying the JAK1/2-STAT1 model. Their procedure numerically estimates kinetic parameters of the model, create an identifiability ranking for them, and finally simplify the model based on the identifiability analysis results. Parameter is called identifiable if it is uniquely defined by the model structure and data, simulated or experimental. The simplification steps are applied until the model is identifiable (i.e. parameter variances are small). Fully identifiable version of the original model was obtained after six in silico iterations. The resulting model has only 9 state variables and 10 parameters. Simplifications include: constitutive binding of JAKs and receptors (cf. ref.44); lumping formation of the receptor complex with STAT1 phosphorylation and, conversely, modeling step dissociation of the whole receptor complex in a single step; as well as omitting cytoplasmic STAT1 dephosphorylation (cf. ref.47).

Modeling based on PSC molecular data. A qualitatively different mathematical model of IFN-γ-mediated signaling in pancreatic stellate cells (PSC) was presented by Rateitschak et al.68 It was based on authors own quantitative experimental data. Their relatively small RRE (11 kinetic parameters, 9 species variables) model describes kinetics of chemical reactions. Delayed processes (such as transcription and translation) have been described as a distributed time delay with gamma kernel. Reduced reaction network includes main JAK1/2-STAT1 signaling features, i.e.: binding of the cytokine (IFN) to the receptor (R) combined with its immediate activation, phosphorylation of STAT1 combined with its immediate homodimerization, nuclear cycling of STAT1, transcription of IRF1 and SOCS1 as target genes and negative feedback formed by the latter protein. Most notably this model lacks negative regulation corresponding to cytoplasmic phosphatases, but on the other hand, includes a post-translational positive feedback loop, representing expression of STAT1 in response to activation of IFN-γ receptors (via unknown underlying mechanism). The scheme of biochemical reactions is shown in Figure 4.

Figure 4: Scheme of the JAK1/2-STAT1 pathway model from the work of Rateitschak et al.68 Waved (transcription and mRNA relocation) and dashed arrows represent delayed processes. Names of variables have been adjusted for a consistency of this review. See text for details.

The results indicate that the contribution of SOCS1 to the termination of STAT1 activation in PSC is limited. The authors suggest exhaustion of IFNγ and dephosphorylation of STAT1 by tyrosine phosphatases as critical steps of STAT1 inactivation in this experimental system. The proposed JAK1/2-STAT1 model68 successfully reproduced key laboratory findings and more importantly allowed to predict the results of new independent experiments not
previously used for model calibration. One of them compared the efficiency of different stimulation modes with respect to the expression/activation of different components of the STAT1 pathway. We considered this experimental design interesting for the following reason: Interferons are cytokines with a variety of clinical applications. Since interferon action also involves induction of negative feedback loops, the application mode (e.g., time between two applications; dose split versus a single high dose) is likely to influence biological efficiency. Thus, studies at the molecular level may provide deeper insights into the cellular basis of interferon responsiveness or resistance. The model developed specifically for PSC is of potential clinical interests because of the role played by STAT1 in the mediation of antifibrotic effects of IFN.

2.2. Network topology reconstruction

In all computational models described up to now one assumed a priori knowledge of the key biochemical reactions involved in the pathway. Consequently it was feasible to model the network dynamics using a set of differential equations. Taking one step back, a different modeling approach that can be found in the literature aims in inference of the network topology from experimental measurements. A methodology which was previously successfully applied in the field of gene regulatory network reconstruction (see, e.g., ref. 69). In the context of JAK-STAT signaling Kaderali et al. 70 reconstructed the network topology from gene knockout data obtained with the use of RNAi screening technology. The pathway is modeled as Bayesian network with probabilistic Boolean threshold functions. Moreover the efficient Monte Carlo procedure to sample from the posterior over model parameters was proposed.

Knockout RNAi dataset. The knockdown experiments was conducted for 10 genes involved in two variants of JAK-STAT pathway (JAK1/2-STAT1 and JAK1/TYK2-STAT1/2) in humane hepatoma cell line under three different conditions: no stimulation, IFN-α stimulation and IFN-γ stimulation. The main obstacle in the network reconstruction task is the fact that the number of possible network topologies increases exponentially with the number of nodes in the network. To reduce the complexity 10 genes of the pathway were manually grouped into six complexes: IFN-α receptor complex, IFN-γ receptor complex, JAK1 and TYK2 kinase complex, JAK1 and JAK2 complex, STAT1, STAT2 and IRF9 complex and STAT1 homodimer. The proposed method correctly reconstructed the core topology of JAK-STAT pathway. All edges were identified except of STAT1 phosphorylation by JAK1/TYK2.

Constraint-based approach. To reconstruct the JAK-STAT signaling network in a human B-cell, Papin and Palsson 71 applied the method called extreme pathways analysis. The methodology, proposed originally in the context of metabolic pathways, enables to study various properties of signaling systems such as: input/output relationships, crosstalk, correlated reaction sets and network redundancy. 72 Extreme signaling pathways represent the edges of the steady-state flux cone derived from convex analysis of systems of reaction equations (given by stoichiometric matrix) with (in)equality constraints such as mass balance and reaction irreversibility. The method was used to reconstruct whole JAK-STAT signaling network based on the stoichiometric matrix consisting of 15 different receptors and 15 corresponding ligands identified in human B cells. There were 297 reactions in total that gave rise to 147 extreme pathways characterizing fundamental functional states of the network. The structure of crosstalk in the JAK-STAT network, formally defined as a nonnegative linear combination of extreme pathways, was also analyzed. Such constraint-based approach may provide the tool for the description of biologically and medically significant properties of the pathway.
3. Discussion

Combination of theoretical and experimental analysis can provide valuable insight into mechanisms of signaling pathways. High quality data is the limiting factor in computational modeling of JAK1/2-STAT1 pathway. We briefly reviewed here analysis performed for existing kinetic models: a set of models derived from the model by Yamada et al.\textsuperscript{47} and a more recent model of JAK1/2-STAT1 pathway in pancreatic stellate cells build from the scratch on the independent dataset.\textsuperscript{68} Rigorous mathematical approach allows one to study the network dynamics, while techniques of sensitivity analysis and model reduction contribute to identification of key pathway components (potential drug targets).

We would like to emphasize, that all reviewed kinetic models of JAK1/2-STAT1 pathway neglect the aspect of intrinsic noise caused by stochastic fluctuations. Our preliminary comparison of mean and standard deviation of the stochastic process with the numerical solution of RRE, both underlying model by Yamada et al.\textsuperscript{47} suggest that deterministic model is a good approximation.\textsuperscript{73} However, there is an evident lack of the exhaustive, systematic analysis of influence of both intrinsic and, simultaneous, intrinsic and extrinsic noise on the behavior of JAK1/2-STAT1 signaling.

In principle, kinetic models building is a tricky and difficult task for many reasons. Except of an incomplete biological knowledge, there is no clear notion of model optimality. With respect to more mundane aspects of model building, standardization of published models via common formats, such as SBML,\textsuperscript{53} databases, such as BioModels,\textsuperscript{51,52} and vocabularies, such as Systems Biology Ontology\textsuperscript{74} is essential, yet still not widely practiced. In context of computational models of JAK1/2-STAT1 signaling pathway this is perfectly exemplified by a variants of the model by Yamada et al.\textsuperscript{47} In fact, Zi et al.\textsuperscript{58} and Soebiyanto et al.,\textsuperscript{66} most likely being unable to exactly reproduce original model, used as slightly different, smaller variant of the model (with just over 50 rate parameters). Similarly, as already mentioned, Shudo et al.\textsuperscript{59} applied their own adjustments to the model. Although in a model of such size none of these changes made significant qualitative differences, small quantitative differences are noticeable and confusing.

On the final note, we sketched recent attempts to an even more challenging task of inference of the network topology from experimental measurements. Such approach is invaluable in cases when the network of biochemical reactions is unknown. In the context of JAK1/2-STAT1 modeling the extreme pathway analysis approach\textsuperscript{71} seems more adequate to understand biochemical reactions network design principles than the Bayesian network inference\textsuperscript{70}.

Concluding, to appreciate the usefulness of a formal model, we need to be conscious that the main role of computational models of signaling pathways is compacting a large number of detailed knowledge and determining within the overall glance at the pathway and its dynamics, which of its elements, when and why are important. It happens that discoveries provided by modeling approach were already known, but their importance lies in the fact that these hypotheses had been confirmed independently by the iterative process of model development and calibration (this is e.g. the case of nucleocytoplasmic shuttling importance from the model by Swameye et al.\textsuperscript{3}). Likewise, computational model helped to demonstrate the importance of dimerization of STAT1 to avoid undesirable activation caused by the noise in the IFN activity\textsuperscript{59} The last example concerns the reported significance of the cytokine receptors dimerization before the IFN appearance for the stability of the pathway activation process for slightly different characteristics of intracellular environment\textsuperscript{44}. 
4. Acknowledgments

The work in this paper was partially supported by the Polish National Center for Science grant no. 2011/01/B/NZ2/00864 and by the Swiss National Science Foundation grant no. 141264.

Study was supported by research fellowship within “Information technologies: research and their interdisciplinary applications” agreement number POKL.04.01.01-00-051/10-00.

References


17

**Figure 1** Illustration of paradigm of interaction between modeling and experiments. Note that the conclusions obtained from the model affect the experimental setup. Similarly, measurements that provide data to bioinformatics analysis allow for model enhancements. In the desired scenario several iterations of such cycle are carried out.

**Figure 2:** Key steps of JAK-STAT signaling in response to IFNs. IFNs induce activation of JAKs, receptors and subsequently, of cytoplasmic STATs; all by phosphorylation. Active STATs essentially form dimers, which then are translocated to nucleus where they act as transcription factors in the specific gene promoter or enhancer regions. IFN signaling negative regulators include multiple cytoplasmic PTPs, nuclear PTP and PIAS proteins, as well as specific to IFN signaling SOCS1 which in principle inhibits STAT1 phosphorylation process. See text for more details.

**Figure 3:** JAK1/2-STAT1 pathway modules denoted by colors on the scheme of a model by Yamada et al.47 (a), and numerical simulations of modules output species representing concentration of, respectively, phosphorylated receptor dimer (b), phosphorylated nuclear STAT1 dimer (c), and SOCS1 (d) with respect to time. Additional decorations of the simulations graphs represent signaling properties, such as: peak activity time $\tau$, duration $\vartheta$ and amplitude $\alpha$ at the basal, steady state level $\pi$, corresponding to a constant IFN input of 10 nM. In such in-silico experimental setup, unbound active receptor reaches peak in ca. 25 min (b) and from that moment on it is gradually used up in activation of STAT1 proteins. In turn, phosphorylated STAT1 dimers accumulate in nuclei, with maximum concentration reached in ca. 1 hour (c). Their slow concentration descent is followed up by delayed expression of SOCS1 proteins (d), which gradually overtake active receptors (b). The remaining small excess of the latter induces a second, much weaker phase of signaling, starting in ca. 5-th hour (b-d). After that the signal is completely attenuated. Duration of activity of signaling molecules $\nu$ elongates downstream of the pathway, and extent of this effect depends on the strength of signaling $\alpha$ (cf. the first and the second phase of signaling; b-d).

**Figure 4:** Scheme of the JAK1/2-STAT1 pathway model from the work of Rateitschak et al.68 Waved (transcription and mRNA relocation) and dashed arrows represent delayed processes. Names of variables have been adjusted for a consistency of this review. See text for details.