Model-based analysis of Interferon-β induced signaling pathway

Jaroslaw Smieja, Mohammad Jamaluddin, Allan R. Brasier and Marek Kimmel

Supplement

Table of contents:

1. IFN–β signaling pathway

The key components of the signaling pathway that were taken into account in model building are shown in the Fig.1 in the paper. Justification for including or rejecting some of the known components is further in the text.

Upon binding target cells, IFN-β induces gene expression programs to produce an anti-viral state through several mechanisms (Der et al., 1998; Kalvakolanu, 2003). One mechanism involves enhancing peptide production from intracellular pathogens by inducing 26S proteasome catalytic activity by expression of the LMP2, and inducing cytosolic to ER transport by expression of the TAP1 and TAP2. Together, coordinate expression of the LMP2/TAP genetic element results in the extracellular display of pathogen derived peptides within the context of the MHC I (Lehner & Trowsdale, 1998; Yewdell, 2005). As a result, cytoxotic CD8-expressing T lymphocytes can then recognize and clear the infected cells. The second mechanism is to induce expression of various IFN stimulated genes, including MxA, oligoadenylate synthetase, PKR, speckled protein-100, and others that produce an anti-viral state by inhibition of viral translation and replication through largely unknown mechanisms (Sen, 2001).

Arguably the most important molecules mediating cell responses after IFN (both type I and II) stimulation are STATs. They comprise a family of several structurally and functionally related proteins, of which STAT1 and STAT2 are taken into account in the analyzed pathway.

In unstimulated cells, both STAT1 and STAT2 are constitutively expressed. They are reported to reside predominantly in the cytoplasm, though not necessarily in all cell types (Melen et al., 2001; Meyer et al., 2002). Our own experiments have shown that while most of the cellular STAT1 is cytoplasmic (Fig 8a,b), a significant portion of STAT2 can be found in the nucleus (Fig.12). Regardless of their localization, however, they are continuously shuttling between the nucleus and cytoplasm (Koester & Hauser, 1999; Begitt et al., 2000; Banninger & Reich, 2004). Due to their size, they require an active transport. It has been shown (Banninger & Reich, 2004; Kraus et al., 2003) that the nuclear shuttling of STAT2 is made possible through its association with the IRF9. Moreover, the majority of STAT proteins seem to be associated with high molecular mass complexes and the amount of free STAT monomers is very small (Melen et al., 2001; Braunstein et al., 2003).

Binding of the IFN-β to a cell receptor results in conformational changes in the cytoplasmic part of the receptor and subsequent activation of Janus kinases TYK2 and JAK1. This, in turn, leads to recruitment of STAT1 and STAT2 (and, possibly, other members of the STAT family) to the receptor and their tyrosine phosphorylation (see reviews (Levy & Darnell, 2002; Shuai & Liu, 2003;Vinkemeier, 2004)). There is no constitutive phosphorylation of STATs in unstimulated cells, as shown in multiple papers concerned with the JAK/STAT pathway. Phosphorylated STATs form hetero- and homodimers (Shuai & Liu, 2003; Taniguchi & Takaoka, 2001). In fact, it is postulated that the dimer formation is the process allowing release of phosphorylated STATs from the receptor (Goodburn et al., 2000). For the sake of simplicity, whenever the term

STAT1 homodimer or *STAT1|STAT2 heterodimer* is used in this paper, it means a dimer of phosphorylated STATs.

In cytoplasm, STAT1|STAT2 heterodimers form a complex with IRF9, called ISGF3 (Banninger & Reich, 2004; Taniguchi & Takaoka, 2001). Moreover, the heterodimers were also found in the nucleus in cells stimulated by IFN (e.g. (Ghislainet al., 2001, and our own experiments). In context of type I IFN actions, usually activity of ISGF3 is discussed and thoroughly analyzed. However, the STAT dimers also play an important role and, in fact, it is the homodimers that are postulated to be the primary source of induced expression of the genes involved in the analyzed regulatory circuit. In any way, both the dimers and ISGF3 complexes are rapidly (Fig. 9) transported into the nucleus where they serve as active transcription factors (TFs). This transport is facilitated by the importin-5α (Fagerlund et al., 2002; Meyer et al., 2002). It is justified to assume that, once in the nucleus, the complexes cannot be exported, since no associated molecule that would facilitate this process has been found so far. Once in the nucleus, they bind to high affinity sequences in target genes, recruiting p300/CBP coactivators, producing chromatin and factor induced acetylation, and inducing gene expression (Glass & Rosenfeld, 2000; Ray et al., 2005).

Though the ISGF3 complex activates transcription of many ISGs, it seems not to influence directly the pathway under investigation (or, at least, it can be assumed that the influence can be neglected in the mathematical model). Though it is possible that IFN-β treatment stimulates IFN-β production (Kroeger et al., 2002), our results have shown that

in the analyzed cell line IFN-β gene is not upregulated (results not shown). These results are consistent with another concept, in which the positive feedback mechanism, resulting in IFN-β production and secretion is fully activated only after viral infection (Taniguchi & Takaoka, 2001). Moreover, even for cells that do express IFN-β gene in response to such stimulation, the developed model will be valid if the initial concentration of IFN-β is sufficiently high. Since the goal of our research is to build a model applicable to a wide variety of cell types, we plan to enhance it in the future, modifying the term that corresponds to STAT phosphorylation so that it is a function of IFN concentration in the extracellular medium.

One of the genes transcribed early in response to IFN stimulation is IRF1, whose promoter region contains a GAS site (Harada et al., 1994), to which STAT1 homodimers bind (for more detailed discussion about the type of dimer see the section 2.3.1. in this *Supplement*). The newly synthesized IRF1 protein undergoes posttranslational modifications making it active (Taniguchi et al., 2001) and afterwards it is imported into the nucleus. Once there, it plays an important role in initiating transcription of late genes (as opposed to early genes, being activated directly by complexes of phosphorylated STATs). LMP2 and TAP1 are among the genes activated by complexes including IRF1 protein following IFN–β treatment (Min et al., 1998; Jamaluddin et al., 2001), as is STAT1. It has been found that the complex of unphosphorylated STAT1 and IRF1 binds to the shared promoter of LMP2 and TAP1 genes (Wright et al., 1995; Chatterjee-Kishore et al., 1998; Brucet et al., 2004), activating their transcription, while IRF1

complex with CBP is needed as the TF for STAT1 gene (Wong et al., 2002), thus creating a positive feedback loop.

There are several mechanisms negatively regulating cell response to IFN (see e.g. the reviews (Yasukawa et al., 2000; Chen et al., 2004; Wormald & Hilton, 2004)). One of them is dephosphorylation of STATs through activity of various phosphatases, found both in the cytoplasm and the nucleus (Hoeve et al., 2002; Mustelin et al., 2005). Dephosphorylation results in dissociation of dimers into monomers (Meyer & Vinkemeier, 2004), and of ISGF3 into STAT1 monomer and, possibly, STAT2|IRF9 dimer, eventually leading to nuclear export of the latter molecules and making them available to subsequent phosphorylation/dephosphorylation cycles (Levy & Darnell, 2002; Haspel & Darnell, 1999). The other mechanism of negative regulation involves a family of PIAS proteins, which are known to inhibit binding activity of STAT dimers and enhance their sumoylation (Rogers et al., 2003; Ungureanu et al., 2003). Finally, the pathway can be regulated by *de novo* produced SOCS proteins (Alexander & Hilton, 2004). However, reports about inducibility of SOCS-1 in response to type I IFN are ambiguous, with some works reporting SOCS-1 induction by type I IFN (e.g. Wang et al., 2000; Dalpke et al., 2001)) and other rejecting the idea or showing very late induction (Kamio et al., 2004). These differences most likely result from different pattern of SOCS-1 expression in different cell types (Levy & Darnell, 2002). Moreover, our own results found neither upregulation of SOCS-1 (Real-Time PCR, results not shown) nor the downregulation of phosphorylation in the analyzed cells (Fig. 2 in the main paper). The graphical representation of the pathway is presented in the Fig. 1 in the main text.

2. Modeling of basic processes – assumptions and simplifications

In this section, basic assumptions underlying mathematical description of processes taken into account are given. All processes taken into account in the model are presented in the Tables 1-2 in the main paper and the model parameters in Tables 1-5 in this *Supplement*. The following notation is used in the model description:

- Variables refer to cytoplasmic content if no subscripts are present, while nuclear content is represented by subscript *n*.
- mRNA transcripts are denoted by *t* subscript and always refer to the cytoplasmic content; the transport of mRNA to the cytoplasm is assumed to be very fast in relation to other processes and therefore neglected in the model.
- Phosphorylated form of proteins is indicated by *p* subscript.
- Subscripts *active* and *inactive* are used to distinguish respective states of the molecules

2.1. STAT1 and STAT2 proteins in unstimulated cells

The STAT proteins are known to form dimers in unstimulated cells (Banninger & Reich, 2004) and reportedly only a small fraction of them is found in the form of monomers (Braunstein et al., 2003). However, since it seemingly does not affect phosphorylation nor subsequent dimerization of the phosphorylated molecules, in our model those molecules are treated as monomers.

2.2. Phosphorylation and dephosphorylation

Phosphorylation of STAT proteins requires several steps, including ligand binding to a receptor, receptor dimerization, activation of JAKs and binding of STAT to the receptor. Moreover, according to some reports, STAT2 phosphorylation precedes STAT1 phosphorylation (Gotoh et al., 2003). In order to simplify the model and avoid assuming too many parameters whose values cannot be found in the literature and could not be measured in our experiments (e.g., initial number of free receptors, JAK concentration, kinetic rates for the processes), phosphorylation is described as a single-step process, with the rate initially proportional to protein concentration. Since there is no induction in SOCS-1 gene expression and Western Blots do not show a significant decrease in phosphorylated STATs (Fig. 2 in the main paper), no additional term is needed that would correspond to the negative control of phosphorylation. However, the dramatic increase in the level of STAT1 proteins does not translate into a similar increase in phosphorylated STAT1 level. This can be explained by the increased competition of the STAT1 proteins for the available receptors at their cytoplasmic end. Therefore we have introduced the saturation into the description of the phosphorylation process.

In the simplified model that is presented here, the phosphorylation rates do not depend on the extracellular IFN concentration. While it is true for the large value of IFN concentration used in our experiments, in the ultimate model the rates should be functions of this concentration and we have already planned the experiments to find these functions. Nevertheless, for a variety of experiments, in which high does of IFN is chosen to be the stimulus, there is no need for introducing that relation.

Dephosphorylation process has been simplified in the model in the similar manner. The biological process requires binding of specific phosphatases to achieve dephosphorylation. There are different kinds of phosphatases that react with the phosphorylated STAT proteins, both in the cytoplasm and in nucleus (Chen et al., 2004; Mustelin et al., 2005). Given that we are able to measure only the overall effect of phosphatases activity and not the specific reactions, the dephosphorylation is described by a simple term, proportional to the amount of phosphorylated molecules.

The model of dephosphorylation of STAT complexes calls for a separate description. Though its rate is also proportional to the dimers concentration, the process involves also degradation of the dimers. In our model we assume that as a result of dimer dephosphorylation, two unphosphorylated monomers are created (or, in the case of ISGF3 complex – STAT1 monomer and the STAT2|IRF9 dimer). The real process most likely entails several steps that lead to dissociation of the complex but it seems reasonable to assume that their dynamics is sufficiently fast, compared to other processes, and can be neglected.

2.3. Formation of complexes

It is assumed that, due to large number of molecules, complexes are created according to the law of mass action. However, two important points must be stressed and they are briefly discussed below.

2.3.1. STAT dimers

It has been postulated that dimerization of STATs is a required step before the release of phosphorylated STATs. While it can be true, it is not reflected in the model structure, since a relatively large constant rate of dimerization would correspond to the situation when dimers are formed in the vicinity of the receptor.

Additionally, it has been also found that STAT2 forms complexes with IRF9 in unstimulated cells, and its nuclear shuttling is the direct consequence of that (Banninger & Reich, 2004). Moreover, it is known that the ISGF3 complex is formed in the cytoplasm (Bluyssen et al., 1999; Taniguchi & Takaoka, 2001; Taniguchi et al., 2001). Therefore it is possible that STAT2 phosphorylation takes place without prior complex degradation (Lau & Horvath, 2002). In the model presented here, only STAT2 is taken into account. However, its dimerization with STAT1 actually corresponds to the formation of any complexes with STAT1. In view of the known mechanisms involving activity of ISGF3, it is very likely that it is this complex that is most closely approximated by the "heterodimer" variable in the model. Unless the model is expanded to include ISGF3 activated genes, this simplification does not affect its applicability .

2.3.2. IRF1|CBP complex

IRF1|CBP complex is known to be a TF for STAT1 gene (Wong et al., 2002). However, taking into account the role of CBP protein in transcriptional processes of other genes (White, 2001) it is very likely that the complex forms after binding of STAT1 to the promoter region of the gene. Even if the complexes were being formed before that event, lack of quantitative data on free CBP levels and of knowledge about other possible

mechanisms mediating creation of this complex renders its explicit modeling impossible. Instead, we chose to model it indirectly, as depicted in the subsection 3.9. devoted to modeling of transcription of late genes.

2.4. Nuclear shuttling of molecules.

We assume that the transport rate of a given substrate to or from the nucleus is proportional to its cytoplasmic or nuclear concentration, respectively. The rationale behind the structure of the equations governing nuclear and cytoplasmic transport is the same as the one presented by (Lipniacki et al., 2004). To account for different volumes of the nucleus and cytoplasm, a scaling factor $k_v = V/U$ (ratio of cytoplasmic and nuclear volumes) has been introduced. Therefore, if the only processes taken into account were nuclear import of cytoplasmic molecules whose concentration is denoted by $[x]$ and export of nuclear molecules whose concentration is denoted by $[x_n]$, their description would be given by:

$$\frac{d[x]}{dt} = -i[x] + e[x_n]$$

$$\frac{d[x_n]}{dt} = k_v i[x] - k_v e[x_n]$$

Where i and e are constant rates for nuclear import and export, respectively.

2.5. Gene transcription

There are two possible approaches to modeling relation between gene transcription and the concentration of active TF. In the simplest one it is assumed that the transcription rate is proportional to TF concentration. More elaborate model takes into account the upper

limit of transcription rate, determined by how fast a polymerase can move along DNA and what is the minimum distance between two polymerases along DNA chain in parallel transcription process. The kinetics assumed in this case is of Michaelis-Menten type.

If the transcription rate is small, compared to the theoretical maximum, the Michaelis Maenten kinetics can be approximated by the linear relation. In the case of genes transcribed in the analyzed pathway we have traced the maximum rate of transcription to check if it is indeed much smaller of the maximum theoretical value of 40 nucleotides/second per gene (Levin, 2000). . We have found that the transcription rate can be assumed to be proportional to the TF concentration. However, if the mathematical model was used for analysis of a disturbed system, in which TFs were accumulated in the nucleus, reaching relatively high concentrations (e.g. when nuclear export was blocked), transcription rate should be expressed in its nonlinear form, as $v_{max}*(TF)/(k+*(TF))$, where v_{max} denotes the theoretical maximum transcription rate and k is a parameter to be identified.

Another simplification lies in additive combination of both constitutive and stimulation–induced transcription. It might be more prudent to introduce a nonlinearity to the model that would switch the model between those two types of transcription. However, we decided not to include it in the model, since the induced rate is much greater than the constitutive rate.

For each gene, the mRNA concentration is calculated using the cytoplasmic volume. It is assumed that the nuclear export of newly synthesized mRNA is very fast and therefore its dynamics can be neglected. Moreover, splicing of pre-mRNA is not modeled as this process runs in parallel to transcription on the molecules that are formed during transcription.

2.6. mRNA translation

We assume that the translation rate is proportional to cytoplasmic transcript concentration. Translation rate does not depend on the transcript length, as multiple protein copies are being produced from one mRNA molecule simultaneously. The rate coefficient is the same for all transcripts, since there is no basis for assuming any preferences of the ribosomes.

2.7. Steady state in unstimulated cells

It is assumed that unstimulated cells are in a steady state, i.e. the concentration of all molecules both in the nucleus and cytoplasm is constant (which, however, does not mean that there is no nuclear shuttling). There are no phosphorylated STATs (Chatterjee-Kishore et al., 2000) at time 0, and, since IRF1 is present in very low amount in unstimulated cells, its steady state level is set to 0.

The direct implication of the steady state assumption is that not all model parameters are independent. Information about parameters that were calculated from the steady state

assumption is given in the section 4 and *Remarks* columns the Tables 1-5, where appropriate.

3. Model sensitivity to parameter changes

All the assumptions mentioned in the preceding section formed the basis for a mathematical model written in the form of ordinary differential equations. All processes included in the model are listed in the Tables 1-5. These tables contain also parameters that have been identified. The fit of simulation results to experimental data is shown in the Fig. 3 in the main paper.

However, the large number of parameters in the model naturally rises the question about robustness of the model with regard to parameter changes, in particular to changes in the initial conditions. To check this, we varied the parameter values and for each set of parameters, we run a separate simulation. First, each parameter in the model was increased, then decreased by 20%. Then, random changes in the parameter set have been introduced, with parameters either staying at the base level, or increased/decreased by 20%. In total 150 sets of parameters have been tested showing that the model is robust in terms of qualitative behaviour (Fig.3 in the main paper).

However, the robustness refers to qualitative behaviour, not the values of variables. As indicated in the *Methods* section in the main paper, each plot is normalized by the area under the curve. Therefore the results do not show differences in absolute values of variables. Since the main goal was to obtain qualitatively good fit, and the real concentrations are not known, such approach is reasonable. Nevertheless, looking at raw numbers provides an additional, valuable insight into the possible variability of cell

responses. Heterogeneity in cell population would result in variability in model parameters for different cells and analysis of various set of parameters might prove beneficial in investigation directed at differences in behavior among cells of different types.

For example, in the analyzed model different sets of initial conditions were introduced (with variations in values as described in the first paragraph of this section). The dynamics of the main variables present in the feedback loops are shown in Fig. 15. If only standardized values are compared, the behaviour of the key players in the pathway is very similar, not only qualitatively, but also quantitatively. However, the actual values obtained for each case show that it is true only for the qualitative behaviour.

4. Specific remarks on equations

The variable representing STAT2 concentration might actually relate to the dimers of unphosphorylated STAT2 and IRF9 and, correspondingly STAT2p to the the dimer STAT2p|IRF9 (see the *Discussion* in the main paper). Then, the variable denoted (STAT1|STAT2) would correspond to ISGF3 complex. Though this equivocal description clearly needs verification, unambiguous explanation is not necessary here unless ISGF3 activated genes are introduced into the model. Therefore, though the variable denotes heterodimer concentration, it might actually represent a compartment of all possible complexes that include phosphorylated STAT1. Alternatively, it can be referred to as ISGF3 concentration, since this complex seems to be a dominant one.

In the subsequent subsections we discuss the philosophy of retaining and removing of specific processes in the mathematical model.

4.1. Free STAT1 and STAT2 in cytoplasm and nucleus

The STAT2 is produced constitutively in the model, as neither our own nor others' results indicate that the STAT2 gene expression is induced in the pathway. Nonetheless, we have chosen to introduce the dynamics of STAT2 transcript, allowing for model expansion if STAT2 proved to be inducible by IFN stimulation. Therefore free STAT1 and STAT2 proteins come either from the respective mRNA tranlation or the dissociating dimers.

4.2. Phosphorylated free STATs in cytoplasm and nucleus

The implicit assumption is that even if postulated release of phosphorylated STATs is possible only after dimerization, the dimerization process should take part in the close vicinity of the receptor (implying relatively large value of association constants k_{s1s1} and k_{s1s2}). Furthermore, although it is quite plausible that two phosphorylated molecules forming a dimer dissociate, their influence on the whole system would be negligible, since the dimers are reported to be very stable (Levy & Darnell, 2002). Therefore, there is a negligible amount of phosphorylated monomers in the nucleus and formation of the dimers there can be neglected.

As far as the phosphorylated form of STAT1 is concerned, taking into account that dimerization and subsequent nuclear import are very fast, the number of those molecules in cytoplasm is relatively low. Hence, small differences between the plates significantly contribute to the experimentally observed oscillations (Fig. 8c).

Receptor-associated tyrosine kinase activity is rate-limiting in formation of phospho-STAT1. Therefore, the rate of phosphorylation is given by

$$\varepsilon = \frac{k_{s1\_phos}[STAT1]}{1 + k_{s1phos\_sat}[STAT1]}$$

However, for the first 2 hours this nonlinearity can be neglected and one can obtain the first-order reaction rate (and corresponding half-life time shown in the Table 3) of the process

$$k_{s1\_phos} = k^*_{s1\_phos} \cdot (1 + k_{s1phos\_sat} \cdot [STAT1](0)) \quad (1)$$

where $k^*_{s1\_phos}$ and k_{s1phos_sat} are parameters defined in the Table 3 and $[STAT1](0)$ is the initial concentration of STAT1 in cytoplasm.

4.3. Phosphorylated STAT1 complexes

Though dephosphorylation and subsequent dissociation of the dimers stem from activity of various phosphatases, they are not modeled explicitly here, except for the one that is postulated to be activated and responsible for the drop in STAT1 homodimer level. First, it is virtually impossible to distinguish between the results of processes controlled by different phosphatases since it would require real-time data on their binding to dimers. This drawback is even more exacerbated by the fact that it is not clear if all phosphatases

specific to STATs are known. Furthermore, it seems reasonable to assume that despite the multiple interactions that must occur during this process, it can be modeled by a simple relation that holds rate of constitutive dephosphorylation proportional to the concentration of a substrate. The only condition is that the phosphatase concentration is constant in the analyzed cell line.

Nuclear import of the complexes containing phosphorylated STAT1 is mediated by importin–α5. No molecule mediating nuclear export has been found so far and it seems that these complexes are trapped inside nucleus until they dissociate. Hence there is no nuclear export of the dimers in the model. Due to high stability of the dimers in the absence of active phosphatases (Haspel & Darnell, 1999) and relatively fast dephosphorylation rates, their irreversible degradation is neglected.

4.4. Negative regulation of STAT1 homodimers

The rapid decline of the IRF1 gene expression indicates a strong negative regulatory component in the JAK-STAT pathway (Fig. 3c in the main text). A careful comparison of nuclear STAT1 binding with that of IRF1 gene expression (Fig. 6 in this *Supplement*) clearly indicates that the decrease in the level of IRF1 mRNA is preceded by the fall of the homodimer level .This negative control might be achieved at any of the five stages: (i) an inhibition of tyrosine phosphorylation at the level of the IFN receptor/JAK kinases by SOCS or similar activity (Yasukawa et al., 2000), (ii) an inhibition of STAT1 nuclear import, (iii) an inhibition of STAT1 homodimeric DNA binding by the PIAS family of

proteins (Wormald & Hilton, 2004), (iv) a regulation of STAT1 homodimer level in nucleus, or (v) activity of an unknown repressor of the IRF1 gene.

We found that in the analyzed cell line an inhibition of STAT phosphorylation is not responsible for the negative regulation, as indicated by the phosphorylated STAT1 (Fig. 2 in the main paper) and STAT2 levels as well as SOCS-1 gene expression measured by Q-RT-PCR (results not shown). Additionally, the level of nuclear STAT1 and phospho-Tyr^{701} STAT1 persisted beyond that of STAT1 homodimer binding (Fig. 2 in the main text). Therefore, nuclear-cytoplasmic export of STAT1 could not account for the decline in STAT1 homodimer binding in EMSA. Furthermore, because cytoplasmic SIE DNA binding was undetectable (results not shown), these data suggested that DNA binding-competent STAT1 homodimers are efficiently imported into the nucleus.

In order to check the nature of the process regulating the homodimer levels, two different hypotheses can be proposed. The first one is degradation, with the possible involvement of PIAS, the second a phosphatase activated in the pathway, as regulators of STAT1 homodimer level.

It has been known that the PIAS family can actively regulate the JAK/STAT pathway. They have been shown to inhibit DNA binding of STAT dimers and to enhance their sumoylation (Rogers et al., 2003; Ungureanuet al., 2003), while not being inducible by the IFN (O'Shea & Watford, 2004). PIAS1 and PIASy could be potential candidates responsible for the particular effects observed in our experiments due to their affinity to

STAT1 (Coccia et al., 2002; Liu et al., 2004; Zhang et al., 2004). Judging from the the homodimer level time course (Fig. 3b in the main paper), the regulation mechanism could involve either the PIAS inhibitory effects on DNA binding or induction of active degradation. The latter could be a result of either ubiquitination or sumoylation of the homodimers. Though the ultimate effect of sumoylation is not yet known, taking into account different effects of this process (Wilson, 2004), it is possible that it would promote homodimer degradation.

Though the arguments for PIAS involvement in regulation of the analyzed pathway look plausible, several factors speak against this hypothesis. First of all, there are no reports about different forms of any given PIAS, rendering it active or inactive, and activation, as shown in the following subsection is crucial to obtain the time courses matching the experimental results. Furthermore, it has been shown in PIAS1 knockout studies that the PIAS1 does not influence IRF1 gene expression (Liu et al., 2004). Similarly, PIASy was reported to have only limited influence on IFN signaling pathway in mice (Roth et al., 2004). Nonetheless, it still remains to be seen if those results relate directly to human cells. But the most important argument against PIAS is that if sumoylation led to irreversible degradation, there would be a visible drop in total STAT1 levels, which is not observed in experimental data (Fig. 2 in the main paper). This is illustrated in Fig. 5 in this *Supplement*. The minimum number of total STAT1 molecules in the model built on degradation assumption is 15-fold lower than for the model assuming increased dephosphorylation and 100 times lower than the initial number of molecules. Such drop

in the total STAT1 level has never been reported, nor it is seen in our results. Therefore the hypothesis about degradation was rejected.

4.5. Free inactive and active PHY phosphatase

PHY protein is the only hypothetical molecule that had to be introduced to explain system behavior. According to the assumed model, it is localized in the nucleus.
Its activation is assumed to depend on the concentration of the heterodimers (or, according to the analysis at the beginning of this section, complexes containing phosphorylated STAT1, other than homodimers) - see Fig. 9 in this *Supplement*. There is no inactivation in our model. There is twofold rationale behind that simplification. First, judging from the observed results, the PHY, once activated, should remain in this state for a relatively long time to maintain the homodimer level down for several hours despite continuous phosphorylation of STAT1. For exactly that reason, it seems unlikely that the transformation into an inactive state is achieved without an interaction with some other molecule. Integrating inactivation process into the model would therefore require introducing even more hypothetical proteins and processes into the dynamical model. Considering this, the simplified model of PHY dynamics without inactivation process seems to be justified. However, this inhibits the model ability to imitate increase of homodimer level observed after 15 hours, as duly described in the *Discussion* section (5.2.1) in the paper.

It should be noted that in order to avoid increasing the number of hypothetical molecules and unknown parameters that could not be identifiable without knowing the nature of these molecules, it is assumed that the PHY is in steady state as far as its production and degradation is concerned. This does not influence the model conformity to the real processes unless both the production and degradation of this protein is very fast. This motion can be rejected since from the presented reasoning it follows that the PHY should be relatively stable.

4.6. The PHY\|STAT1\|STAT1 complex

It should be stressed that after the dissociation of the complex the STAT1 homodimer no longer exists, as depicted in the Fig.9.

4.7. Free IRF1 protein in the cytoplasm and nucleus.

Though activation of the IRF1 protein undoubtedly exhibits its own dynamics, it is neglected in the model following the assumption that the activation happens almost instantly. Otherwise, even if the rate of activation was assumed to be proportional to the concentration of newly synthesized IRF1 proteins, the model would have to be more complicated, including two different inactive IRF1 isoforms – one that can and another one that cannot be activated (see section 5.2.1. in the paper for more details).

The IRF1 protein is a relatively small molecule (31kDa), therefore it can shuttle to and from nucleus through the nuclear pores. As a result, both nuclear export and import of the active IRF1have to be included in the model.

If the rationale behind nuclear import is activation of IRF1, then clearly this mechanism is turned off at some time in the course of IFN stimulation. One of plausible explanations of this effect is a mechanism blocking activation of IRF1 in the cytoplasm (see section 8 of this *Supplement*). This explanation is even more appealing if we consider that it requires only two forms of IRF1 – active and inactive – to build a structural view of the pathway. Nonetheless, in order to take it into account, one should assume the dynamics of an unknown factor blocking the IRF1 activation, since it is clear that it cannot be based on cytoplasmic/nuclear levels of this protein alone. Therefore, we decided to pursue another explanation. The process that is incorporated in the model is based on the assumption that the IRF1 is irreversibly inactivated in the nucleus, followed by its nuclear export (Fig. 1 in the main text).

4.8. IRF1|STAT1 complexes

Formation of these dimers in the cytoplasm is neglected, since it is assumed that the nuclear import of the activated IRF1 proteins is very fast. Moreover, in our experiments (results not shown) DNA binding of STAT1|IRF1 complex was not detected in the cytoplasmic extract. Besides, even if STAT1|IRF1 dimers were formed in the cytoplasm, their transport would have to be facilitated by an importin, due to the size of STAT1. Such cytoplasmic dimers and their nuclear import have not been reported so far, hence we assumed that their influence on the processes, even if existing, would be negligible. Moreover, numerical tests performed for a wide range of parameters in a model including the cytoplasmic complexes (results not shown) did not yield good results. Therefore, we

have rejected hypotheses about cytoplasmic sequestration of IRF1 by unphosphorylated STAT1 which is produced in large amounts in the pathway.

4.9. IRF1, STAT1, TAP1 and LMP2 mRNA

Induced IRF1 mRNA production rate is assumed to be proportional to STAT1 homodimer concentration in the nucleus. Similarly, late TAP1 and LMP2 genes transcription is directly modeled as proportional to concentration of IRF1|STAT1 complex, their TF. These genes are short (about 1kbp and 2kbp, respectively), hence the dynamics of any additional processes (if they exist) can be neglected here. However, with the late gene STAT1, the procedure must be different. First, our model does not include its TF, IRF1|CBP concentration (see section 2.3.2). To account for the dynamics of concentration of these molecules, as well as other, unknown processes likely to be required by STAT1 transcription (e.g. chromatin remodeling – STAT1 is a relatively long gene), additional variables have been introduced. They do not represent any particular concentration nor processes. They approximate processes that occur between binding of a known TF and the full-swing transcription, and represent the dynamics of serially connected first order time-lag elements. For STAT1 gene four such elements were needed. This clearly can be proposed in view of experimental results showing significant delays between the peaks of TFs levels and expression of their respective late target genes. Indeed, CBP that is needed to form complex with IRF1 for induced transcription of STAT1 gene, is a known factor that affects the chromatin structure (White, 2001). In fact, these downstream events dependent on chromatin remodeling and promoter assembly can sometimes take significant amounts of time as observed in our earlier

mechanistic studies of STAT3 dependent transcription. Using a rapidly inducible luciferase reporter gene, we found that STAT3 dependent transcription can be delayed by as much as 12-24 h after STAT3 enters the nucleus and binds its target genes (Ray et al., 2002). We suggest that a similar rate limiting step occurs for the IRF1 dependent genes.

From the considerations given above it stems that binding of additional molecules to DNA or to a regulatory complex being formed on the promoter region can be described as a Poisson process so that the binding times are exponentially distributed random variables. In terms of deterministic modeling, binding of a single molecule can be represented as a first order lag element. Its time constant corresponds to the parameter λ_i of exponential distribution ($T_i = 1/\lambda_i$). Moreover, this element can be also used to model other processes, including subsequent steps of chromatin remodeling.

This approach allows including unknown processes in the model. Moreover, only a limited number of elements need to be included, regardless of the number of processes that must take place. The reason is that if the time constants significantly differ, and only the final output of the system is of interest, the faster processes can be neglected, reducing the model order. If they are similar, introducing additional elements does not affect the output significantly, at least for the type of input possible in the analyzed system (no high frequency oscillations). In practice, 3-4 elements are sufficient to reproduce the system responses.

Let ($STAT1_t$) denote the concentration of STAT1 mRNA, v_{s1t}, $k_{s1tprod}$ and k_{s1tdeg} induced production constant, constitutive production and degradation rates, respectively. Then, if we denote by $(TF_{s1})_n$ the output of appropriate dynamical element used as an input for STAT1 gene transcription, the equation describing transcript dynamics takes the following form:

$$\frac{d(STAT1_t)}{dt} = k_{s1tprod} + v_{s1t} \cdot (TF_{s1})_n - k_{s1t\_\deg} \cdot (STAT1_t)$$

To calculate $(TF_{s1})_n$ the following equations are used

$$\frac{dy_i}{dt} = -\frac{1}{T} \cdot y_i + \frac{1}{T} \cdot y_{i-1}$$

where $y_0 = (IRF1)_n$ (i.e. nuclear IRF1 protein concentration) i = 1,2,3,4, and $(TF_{s1})_n = y_4$.

5. Initial conditions and relations between parameters

The equations describing the processes can be used to find relations between parameters, if we assume steady state in unstimulated cells (parameters include also initial conditions). These relations are subsequently used to reduce the degree of freedom (number of parameters that can be independently changed) of the model.

All initial conditions are assumed values. For better clarity, they are given in the numbers of molecules, instead of the concentrations that are used in simulations. Two kind of assumptions have been made: first, the total number of proteins of a given kind (including those embedded in complexes); second, the distribution between cytoplasm and nucleus (given as a percentage of total number of proteins). The concentrations are calculated using assumed cytoplasmic and nuclear volumes.

For all parameters, their value has been rounded to the third most significant digit.

Let us assume the following notation:

$(STAT1)_{c\_ss}$ – steady state concentration of free STAT1 in cytoplasm

$(STAT1)_{n\_ss}$ – steady state concentration of free STAT1 in nucleus

$(STAT2)_{c\_ss}$ – steady state concentration of free STAT2 in cytoplasm

$(STAT2)_{n\_ss}$ – steady state concentration of free STAT2 in nucleus

$(STAT1_t)_{ss}$ – steady state concentration of STAT1 mRNA

$(STAT2_t)_{ss}$ – steady state concentration of STAT2 mRNA

$(LMP2_t)_{ss}$ – steady state concentration of LMP2 mRNA

Then, steady state assumption for unstimulated cells leads to the following relations (parameters are described in Tables 1-5 in this *Supplement*):

$$(STAT1_t)_{ss} = k_{s1deg}\,((STAT1)_{c\_ss} + (STAT1)_{n\_ss}/k_v)\,/k_{transl}; \quad \text{(S1)}$$

$$(STAT2_t)_{ss} = k_{s2deg}\,((STAT2)_{c\_ss} + (STAT2)_{n\_ss}/k_v)/k_{transl}; \quad \text{(S2)}$$

$$k_{s1tprod} = \mathrm{k}_{s1t\_deg}\,(STAT1_t)_{ss} \quad \text{(S3)}$$

$$k_{s2tprod} = \mathrm{k}_{s2t\_deg}\,(STAT2_t)_{ss} \quad \text{(S4)}$$

$$k_{l2tprod} = \mathrm{k}_{l2t\_deg}\,(LMP2_t)_{ss} \quad \text{(S5)}$$

$$k_{t1tprod} = \mathrm{k}_{t1t\_deg}\,(TAP1_t)_{ss} \quad \text{(S6)}$$

$$i_{s1} = (STAT1)_{n\_ss}\,(e_{s1} + k_{s1deg}/k_v)/(STAT1)_{c\_ss}; \quad \text{(S7)}$$

$$i_{s2} = (STAT2)_{n\_ss}\,(e_{s2} + k_{s2deg}/k_v)/(STAT2)_{c\_ss}; \quad \text{(S8)}$$

6. Comparing simulation results to experimental data

The cytoplasmic: nuclear volume is based on microscopic analysis of HeLa cells (Fig. 10). Here, the nuclei appear as clear halos surrounded by blue rims of cytoplasm. On average, there is a 1:2 ratio of cytoplasmic to nuclear volume.

It is extremely difficult to define a good standard that could be used to evaluate how well simulation corresponds to experimental data. One of possible approaches is to calculate a mathematical performance index for the test, e.g. the sum of squared errors taken at the measurement moments or the correlation coefficient. However, we chose to evaluate the quality of the fit using our subjective judgment – this way it's easier to 1) discriminate against measurements of poor quality without discarding them altogether and 2) build the model that reflects qualitatively the dynamics (in terms of oscillations, dumping etc.). The sum of squared errors can be used to find initial estimates for the coefficients, since there exist algorithms that solve the resulting parametric optimization problem.

7. Discrepancies between simulation and experimental data

On of the main differences between simulation and experimental data can be seen late in the course in homodimer levels. However, this discrepancy takes place only after the IRF1 gene has been activated and therefore it does not significantly influence any other part of pathway. This discrepancy is caused by a simplified explanation of induced degradation of STAT1 homodimer. In fact, we have chosen not to try to emulate the second peak, since it would result in making the model unnecessarily more complicated.

Nevertheless, even the simplified hypothesis that underlies homodimer dynamics provides a reasonably accurate prediction of IRF1 gene expression (Fig. 1c in the paper).

Despite the apparent discrepancy between experimental data and simulation for total STAT1 in cytoplasm (Fig.7a in the Supplement), the pathway seems to be correctly modeled: Taking into account the subsequent events that take place, cytoplasmic level of STAT must initially fall, since phosphorylated STATs are imported into nucleus after dimerization. In fact, careful examination of the blots shows that it is exactly the case. The discrepancy is to some extent due to an imperfect quantification procedure, but the largest obstacle in obtaining a good fit is the difference in initial conditions between cell plates used in experiment. Moreover, due to a very long half-life of STAT1 protein (16 h) the difference of even one or two STAT1 mRNA molecules can significantly contribute to the difference in initial STAT levels.

It should be stressed that the highly variable IRF1 mRNA levels apparent in the experimental result (Fig. 3c) in our opinion do not represent genuine oscillations in the system. Though one could argue that they correspond to the first cycle of STAT dimerization - nuclear import – dephosphorylation – nuclear export, this notion is not supported by the blot data. Besides, these oscillations would be too fast for a real biological system. The most likely explanation of the observed phenomenon is the variability in dynamics between replicate cell plates.

Comparison of experimental data and simulation for STAT2 is presented in Fig. 12. It must be stressed that all processes other than phosphorylation and complex

formation/dissociation have been neglected for STAT2. Those that were considered had to be included in the model to account for the STAT1dynamics resulting from complexes formation/dissociation. However, they are not sufficient to obtain a model reflecting the true STAT2 dynamics. Nevertheless, behavior of the cytoplasmic STAT2 is well reproduced by the model. Discrepancy for nuclear STAT2 can be explained by the quality of the blot for time points 10 and 20 minutes. Moreover, if the change in nuclear content of STAT2 is due to nuclear import, then the cytoplasmic and nuclear contents should correspond to each other as in simulation plots.

8. An alternative model explaining cytoplasmic IRF1 dynamics

The reason why the (inactive) IRF1 is accumulated in the cytoplasm, evades our explanation. Still, since it does not any role in the pathway, we decided not to pursue the explanation, since it would require introducing new, unknown elements into the pathway model. This approach was acceptable, as the nuclear IRF1, an active TF, is well represented in our model.

Nevertheless, we have attempted to find out the reason for the IRF1 sequestration in the cytoplasm. Since IRF1 forms complexes with unphosphorylated STAT1 in the nucleus, which subsequently act as active TFs, SIE DNA binding of IRF1 was measured in the cytoplasm. EMSA did not show binding activity of the cytoplasmic IRF1 even though IRF1 was present in cytoplasm as shown by Western Blot. This means that either STAT1 does not dimerize with IRF1 in cytoplasm, or that if it does, IRF1 is already in inactive form. Simulation performed with models based on the assumption of IRF1 sequestered in

the cytoplasm by STAT1 also did not yield good fits to experimental data. Therefore, an alternative model was built (see Fig. 13). It was assumed that an unknown repressor of activation was produced in the pathway. This required introducing two additional variables, one representing hypothetical mRNA and the other protein. As in the original model, IRF1 can be in one of two forms – inactive or active. *De novo* produced IRF1 protein is inactive. It is activated in the first-order process – no explicit activator is needed. However, there is a repressor of activation postulated – the larger concentration of the repressor, the smaller rate of activation. As in the original model, 1) only active IRF1 enter nucleus; 2) inactivation results in conformational change and subsequent nuclear export; 3) The rate of inactivation is proportional to the concentration of active IRF1.

It is assumed that the TF for the gene coding this repressor exhibits dynamics similar to IRF1 (it might be an autoregulatory negative feedback). The activation process for IRF1 is described by the following equation

$$\frac{d[IRF1_{active}]}{dt} = \frac{k_{i1\_activ}[IRF1_{inactive}]}{1 + k_{inirf1}[X]} \tag{S9}$$

where [*X*] denotes the cytoplasmic concentration of an unknown activation inhibitor, k_{i1_activ} and k_{inirf1} are constant parameters.

Under these assumptions the simulation results provide reasonable fit to the Western Blot data, as shown in the Fig. 14 (all other fits do not change).

References

Alexander WS, Hilton DJ: The role of suppressors of cytokine signaling (SOCS) proteins in regulation of the immune response. *Annu Rev Immunol* 2004, 22:503–29.

Banninger G, Reich NC: STAT2 Nuclear Trafficking. *J Biol Chem* 2004, 279(38):39199–39206.

Begitt A, Meyer T, van Rossum M, Vinkemeier U: Nucleocytoplasmic translocation of Stat1 is regulated by a leucine-rich export signal in the coiled-coil domain. *PNAS* 2000, 97(19):10418–10423.

Bluyssen HAR, Durbin JE, Levy DE: ISGF3g p48, a Specificity Switch for Interferon Activated Transcription Factors. *Cyt & Growth Factor Rev* 1996, 7(1):11-17.

Braunstein J, Brutsaert S, Olson R, Schindler C: STATs Dimerize in the Absence of Phosphorylation. *J Biol Chem* 2003, 278(36):34133–34140.

Brucet M, Marques L, Sebastian C, Lloberas J, Celada A: Regulation of murine Tap1 and Lmp2 genes in macrophages by interferon gamma is mediated by STAT1 and IRF1. *Genes Immun* 2004, 5:26–35.

Chatterjee-Kishore M, Kishore R, Hicklin DJ, Marincola FM, Ferrone S: Different Requirements for Signal Transducer and Activator of Transcription 1a and Interferon Regulatory Factor 1 in the Regulation of Low Molecular Mass Polypeptide 2 and Transporter Associated with Antigen Processing 1 Gene Expression. *J Biol Chem* 1998, 273(26):16177–16183.

Chatterjee-Kishore M, Wright KL, Ting JP-Y, Stark GR: How Stat1 mediates constitutive gene expression: a complex of unphosphorylated Stat1 and IRF1 supports transcription of the LMP2 gene. *EMBO J* 2000, 19:4111–4122.

Chen W, Daines MO, Khurana Hershey GK: Turning off signal transducer and activator of transcription (STAT): The negative regulation of STAT signaling. *J Allergy Clin Immunol* 2004, 114:476-89.

Coccia EM, Stellacci E, Orsatti R, Benedetti E, Giacomini E, Marziali G, Valdez BC, Battistini A: Protein inhibitor of activated signal transducer and activator of

transcription (STAT)-1 (PIAS-1) regulates the IFN-γ response in macrophage cell lines. *Cell Signal* 2002, 14:537– 545.

Cramer LA, Klemsz MJ: Altered Kinetics of Tap-1 Gene Expression in Macrophages Following Stimulation with both IFN-γ and LPS1. *Cell Immunol* 1997, 178:53–61.

Dalpke A H, Opper S, Zimmermann S, Heeg K: Suppressors of Cytokine Signaling (SOCS)-1 and SOCS-3 Are Induced by CpG-DNA and Modulate Cytokine Responses in APCs. *J Immunol* 2001, 166:7082-7089.

Fagerlund R, Melen K, Kinnunen L, Julkunen I: Arginine/Lysine-rich Nuclear Localization Signals Mediate Interactions between Dimeric STATs and Importin α5. *J Biol Chem* 2002, 277(33):30072–30078.

Frevel MAE, Bakheet T, Silva AM, Hissong JG,. KSA Khabar, Williams BRG: p38 Mitogen-Activated Protein Kinase-Dependent and –Independent Signaling of mRNA Stability of AU-Rich Element-Containing Transcripts. *Mol Cell Biol* 2003, 23(2):425–436.

Ghislain JJ, Wong T, Nguyen M, Fish EN: The Interferon-Inducible Stat2:Stat1 Heterodimer Preferentially Binds In Vitro to a Consensus Element Found in the Promoters of a Subset of Interferon-Stimulated Genes. *J. Interferon Cytokine Res* 2001, 21:379–388.

Glass CK and Rosenfeld MG (2000) The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev*, **14**(2), 121-141.

Goerlich D, Seewald MJ, Ribbeck K: Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J* 2003, 22:1088-1100.

Goodbourn S, Didcock L, Randall RE: Interferons: cell signalling, immune modulation, antiviral responses and virus countermeasures. *J Gen Virol* 2000, 81:2341–2364.

Gotoh B, Takeuchi K, Komatsu T, Yokoo J: The STAT2 Activation Process Is a Crucial Target of Sendai Virus C Protein for the Blockade of Alpha Interferon Signaling. *J Virol* 2003, 77(6):3360–3370.

Harada H, Takahashi E-I, Itoh S, Harada K, Hori T-A, Tanigushi T: Structure and Regulation of the Human Interferon Regulatory Factor 1 (IRF1) and IRF-2 Genes:

Implications for a Gene Network in the Interferon System. *Mol Cell Biol* 1994, 14(2):1500-1509.

Haspel RL, Darnell Jr. JE: A nuclear protein tyrosine phosphatase is required for the inactivation of Stat1. *PNAS* 1999, 96:10188–10193.

Heinrich PC, Behrmann I, Haan S, Hermanns HM, Mueller-Newen G, Schaper F: Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem J* 2003, 374:1–20.

Hoeve J, Ibarra-Sanchez M de J, Fu Y, Zhu W, Tremblay M, David M, Shuai K: Identification of a Nuclear Stat1 Protein Tyrosine Phosphatase. *Mol Cell Biol* 2002, 22(16):5662–5668.

Jamaluddin M, Wang S, Garofalo RP, Elliott T, Casola A, Baron S, Brasier AR: IFN–β mediates coordinate expression of antigen-processing genes in RSV-infected pulmonary epithelial cells. *Am J Physiol Lung Cell Mol Physiol* 2001, 280:L248–L257.

Jong de H: Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *J Comput Biol* 2002, 9(1):67–103.

Kalvakolanu DV (2003) Alternate interferon signaling pathways. Pharmacol Ther, 100, 1–29.

Kamio M, Yoshida T, Ogata H, Douchi T, Nagata Y, Inoue M, Hasegawa M, Yonemitsu Y, Yoshimura A: SOCS1 inhibits HPV-E7-mediated transformation by inducing degradation of E7 protein. *Oncogene* 2004, 23:3107–3115.

Koester M, Hauser H: Dynamic redistribution of STAT1 protein in IFN signaling visualized by GFP fusion proteins. *Eur J Biochem* 1999, 260:137-144.

Kraus TA, Lau JF, Parisien J-P, Horvath CM: A Hybrid IRF9-STAT2 Protein Recapitulates Interferon-stimulated Gene Expression and Antiviral Response. *J Biol Chem* 2003, 278(15):13033–13038.

Kroeger A, Koester M, Schroeder K, Hauser H, Mueller PP: Activities of IRF1. *J Interferon Cytokine Res* 2002, 22:5–14.

Lankat-Buttgereit B, Tampe R: The Transporter Associated With Antigen Processing: Function and Implications in Human Diseases. *Physiol Rev* 2002, 82:187–204.

Lau JF, Horvath CM: Mechanisms of Type I Interferon Cell Signaling and STAT-Mediated Transcriptional Responses. *Mount Sinai J. Medicine* 2002, 69(3):156-168.

Lehner PJ and Trowsdale J (1998) Antigen presentation: Coming out gracefully. *Curr Biol*, **8**, R605–R608.

Levin B: *Genes, Vol. VII*. Oxford University Press, Oxford., 2000.

Levy DE, Darnell Jr J: STATs: transcriptional control and biological impact. *Nat Rev Mol Cell Biol* 2002, 3:651-662.

Lipniacki T, Paszek P, Brasier AR, Luxon B, Kimmel M: Mathematical model of NF–κB regulatory module. *J Theor Biol* 2004, 228:195-215.

Liu B, Mink S, Wong KA, Stein N, Getman C, Dempsey PW, Wu H, Shuai K: PIAS1 selectively inhibits interferon-inducible genes and is important in innate immunity. *Nat. Immunol* 2004, 5(9):891-898.

Melen K, Kinnunen L, Julkunen I: Arginine/Lysine-rich Structural Element Is Involved in Interferon-induced Nuclear Import of STATs. *J Biol Chem* 2001, 276(19):16447–16455.

Meyer T, Begitt A, Lodige I, van Rossum M, Vinkemeier U: Constitutive and IFN-γ-induced nuclear import of STAT1 proceed through independent pathways. *EMBO J* 2002, 21:344–354.

Meyer T, Gavenis K, Vinkemeier U: Cell Type-Specific and Tyrosine Phosphorylation-Independent Nuclear Presence of STAT1 and STAT3. *Exp Cell Res* 2002, 272:45–55.

Meyer T, Vinkemeier U: Nucleocytoplasmic shuttling of STAT transcription factors. *Eur J Biochem* 2004, 271:4606–4612.

Min W, Pober JS, Johnson DR: Interferon Induction of TAP1: The Phosphatase SHP-1 Regulates Crossover Between the IFN–α/β and the IFN–γ Signal-Transduction Pathways. *Circ Res* 1998, 83:815-823.

Mustelin T, Vang T, Bottini N: Protein tyrosine phosphatases and the immune response. *Nat Rev Immunol* 2005, 5:43-57.

Ray S, Boldogh I and Brasier AR (2005) STAT3 NH2-terminal acetylation is activated by the hepatic acute-phase response and required for IL-6 induction of angiotensinogen, *J Gastroenterol*, **129**(5), 1616-32.

Rogers RS, Horvath CM, Matunis MJ: SUMO Modification of STAT1 and Its Role in PIAS-mediated Inhibition of Gene Activation. *J Biol Chem* 2003, 278(32):30091–30097.

Roth W, Sustmann C, Kieslinger M, Gilmozzi A, Irmer D, Kremmer E, Turck C, Grosschedl R: PIASy-Deficient Mice Display Modest Defects in IFN and Wnt Signaling. *J Immunol* 2004, 173:6189–6199.

O'Shea JJ, Watford W: A peek at PIAS. *Nat Immunol* 2004, 5(9):875-876.

Shuai K, Liu B: Regulation of JAK–STAT signalling in the immune system. *Nat Rev Immunol* 2003, 3:900-911.

Taniguchi T, Ogasawara K, Takaoka A, Tanaka N: IRF family of transcription factors as regulators of host defense. *Annu Rev Immunol* 2001, 19: 623–55.

Taniguchi T, Takaoka A: A weak signal for strong responses: interferon-α/β revisited. *Nat Rev Mol Cell Biol* 2001, 2:378-386.

Ungureanu D, Vanhatupa S, Kotaja N, Yang J, Aittomaeki S, Jaenne OA, Palvimo JJ, Silvennoinen O: PIAS proteins promote SUMO-1 conjugation to STAT1. *Blood* 2003, 102:3311-3313.

Vinkemeier U: Getting the message across, STAT! Design principles of a molecular signaling circuit. *J Cell Biol* 2004, 167(2):197–201.

Voet D, Voet JG: Biochemistry, 3rd Edition. John Wiley & Sons, Inc, USA, 2004.

Wang Q, Miyakawa Y, Fox N, Kaushansky K: Interferon–α directly represses megakaryopoiesis by inhibiting thrombopoietin-induced signaling through induction of SOCS-1. *Blood* 2000, 96(6):2093-2099.

White RJ: *Gene Transcription*. Blackwell Publishers, USA, 2001.

Wilson van G (ed.): *Sumoylation : molecular biology and biochemistry*. Horizon Bioscience, Wymondham, Norfolk, UK, 2004.

Wong LH, Sim H, Chatterjee-Kishore M, Hatzinisiriou I, Devenish RJ, Stark G, Ralph SJ: Isolation and Characterization of a Human STAT1 Gene Regulatory Element. *J Biol Chem* 2002, 277(22):19408–19417.

Wormald S, Hilton DJ: Inhibitors of Cytokine Signal Transduction. *J Biol Chem* 2004, 279(2):821–824.

Wright KL, White LC, Kelly A, Beck S, Trowsdale J, Ting JP-Y: Coordinate Regulation of the Human TAP1 and LMP2 Genes from a Shared Bidirectional Promoter. *J Exp Med* 1995, 181:1459-1471.

Yamada S, Shiono S, Joo A, Yoshimura A: Control mechanism of JAK/STAT signal transduction pathway. *FEBS Lett* 2003, 534:190-196.

Yasukawa H, Sasaki A, Yoshimura A: Negative Regulation Of Cytokine Signaling Pathways. *Annu Rev Immunol* 2000, 18:143–164.

Yewdell JW (2005) The seven dirty little secrets of major histocompatibility complex class I antigen processing. *Immunol Rev*, **207**, 8–18.

Zhang J, Xu L-G, Han K-J, Wei X, Shu H-B: PIASy represses TRIF-induced ISRE and NF-kB activation but not apoptosis. *FEBS Lett* 2004, 570:97–101.

Zi Z, Cho K-H, Sung M-H, Xia X, Zheng J, Sun Z: In silico identification of the key components and steps in IFN-□ induced JAK-STAT signaling pathway. *FEBS Lett* 2005, 579:1101–1108.

Tables

Table 1. Degradation, dissociation, production and inactivation rates

Reaction	Parameter	Value [$10^{-4}s^{-1}$]	$t_{1/2}$ [min]	Remarks

Degradation rates				
$[STAT1_t] \rightarrow \varnothing$	k_{s1t_deg}	3.85	30	Fitted
$[STAT2_t] \rightarrow \varnothing$	k_{s2t_deg}	3.85	30	Fitted
$[IRF1_t] \rightarrow \varnothing$	k_{i1t_deg}	0.64	180	Though it seems a large value for a regulatory element, especially taking into account the short half-life time of IRF1 protein, the value is in fact smaller than 465 min. reported in (Frevel et al., 2003); which would be too large to obtain a good fit for the decreasing slope of IRF1 mRNA
$[LMP2_t] \rightarrow \varnothing$	k_{l2t_deg}	0.64	180	In (Brucet et al., 2004; (Cramer & Klemsz, 2007) half-life time was reported to be decreased from 24 to 4 hrs due to IFN-γ stimulation
$[TAP1_t] \rightarrow \varnothing$	k_{t1t_deg}	0.96	120	
$[STAT1] \rightarrow \varnothing$	k_{s1deg}	0.128	900	In accordance to Heinrich et al., 2003)
$[STAT1_p] \rightarrow \varnothing$	k_{s1pdeg}	0.128	900	Assumed to be the same as for unphosphorylated form
$[STAT2] \rightarrow \varnothing$	k_{s2deg}	0.128	900	Assumed to be similar (equal to) the corresponding parameter for STAT1
$[STAT2_p] \rightarrow \varnothing$	k_{s2pdeg}	0.128	900	Assumed to be the same as for unphosphorylated form
$[IRF1_{act}] \rightarrow \varnothing$	k_{i1deg}	5.78·	20	In accordance to (Taniguchi & Takaoka, 2001), where the half-life time is reported to be shorter than 30 min
$[IRF1_{in}] \rightarrow \varnothing$	k_{i1_indeg}	5.78	20	for inactive form of IRF1; assumed to be the same as for an active form
$[STAT1\|IRF1] \rightarrow \varnothing$	$k_{s1i1deg}$	0.289	400	Assumed to be $0.01 \cdot k_{inv\_s1i1}$
Dissociation rates				
$[STAT1_p\|STAT1_p] \rightarrow$ [STAT1]+[STAT1]	k_{inv_s1s1}	4.62	25	Due to a very fast nuclear import, this parameter does not have much impact on the results as long as it is smaller or equal (and $t_{1/2}$ larger or equal) than corresponding parameter for dissociation in the nucleus. Dissociation for STAT dimers correspond to their dephosphorylation rates due to constitutively active phosphatases
$[STAT1_p\|STAT1_p]_n \rightarrow$ [STAT1]+[STAT1]	$k_{inv\_s1s1\_n}$	4.62	25	Estimates for dephosphorylation in the nucleus vary – 10-15 min (Levy &

$[STAT1]_n$+$[STAT1]_n$				Darnell, 2002), 15-30 min (Meyer & Vinkemeier, 2004)
$[STAT1_p\|STAT2_p]\rightarrow$ $[STAT1]$+$[STAT2]$	k_{inv_s1s2}	4.62	25	Fitted, under assumption that it should be of the same order as the respective constant for homodimers
$[STAT1_p\|STAT2_p]_n\rightarrow$ $[STAT1]_n$+$[STAT2]_n$	$k_{inv\_s1s2\_n}$	4.62	25	Fitted, under assumption that it should be of the same order as the respective constant for homodimers
$[STAT1\|IRF1_{act}]_n\rightarrow$ $[STAT1]_n$+$[IRF1]_n$	k_{inv_s1i1}	29	4	Fitted
$[STAT1_p\|STAT1_p\|PHY]_n$ $\rightarrow 2[STAT1]_n + [PHY]_n$	$k_{inv\_phys1s1}$	39	3	Fitted
Transcription rates				
$[STAT1_p\|STAT1_p]_n \rightarrow$ $[STAT1_p\|STAT1_p]_n$ $+[IRF1_t]$	v_{i1t}	0.05	N/A	Fitted
$[TF] \rightarrow [TF] + [STAT1_t]$	v_{s1t}	0.06	N/A	Fitted
$[IRF1]_n\rightarrow[IRF1]_n +$ $[LMP2_t]$ $[IRF1]_n\rightarrow[IRF1]_n +$ $[TAP1_t]$	v_{l2t} v_{t1t}	0.02	N/A	For TAP1 and LMP 2 genes; if no repressors are taken into account, they should be similar, since both share the same promoter and are relatively short (so there should be no significant differences in chromatin remodeling processes for them)
Translation rate (as in (Lipniacki et al., 2004))				
$[STAT1_t] \rightarrow [STAT1]$, $[IRF1_t] \rightarrow [IRF1_{active}]$	k_{transl}	500	N/A	As in (Lipniacki et al., 2004)
Inactivation rates				
$[IRF1_{active}]_n \rightarrow$ $[IRF1_{inactive}]_n$	$k_{inacti1}$	7.7	15	Fitted

For better clarity, rate constants are transformed into respective half-life times, where appropriate. Dissociation rates for dimers correspond to their dephosphorylation rates due to constitutively active phosphatases.

Table2 Transport parameters

Reaction	Parameter	Value $[10^{-2}s^{-1}]$	$t_{1/2}$ [min]	Remarks
$[STAT1]_n \rightarrow [STAT1]$	e_{s1}	0.14	8	Fitted
$[STAT2]_n \rightarrow [STAT2]$	e_{s2}	0.14	8	Fitted
$[STAT1] \rightarrow [STAT1]_n$	i_{s1}	0.00817	141	Calculated from steady state assumption (SI7)
$[STAT2] \rightarrow [STAT2]_n$	i_{s2}	0.049	141	Calculated from steady state assumption (SI8)
$[STAT1_p\|STAT1_p] \rightarrow [STAT1_p\|STAT1_p]_n$	i_{s1s1}	2.31	0.5	Fitted
$[STAT1_p\|STAT2_p] \rightarrow [STAT1_p\|STAT2_p]_n$	i_{s1s2}	2.31	0.5	Fitted
$[IRF1_{active}] \rightarrow [IRF1_{active}]_n$	i_{i1}	2.31	0.5	Fitted
$[IRF1_{active}]_n \rightarrow [IRF1_{active}]$	e_{i1}	0.58	2	Since IRF1 is an active TF, nuclear export should be much smaller than nuclear import; here we have assumed $e_{i1} = 10^{-4} \cdot i_{i1}$
$[IRF1_{inactive}]_n \rightarrow [IRF1_{inactive}]$	e_{i1_in}	0.00023	500	Fitted

Table 3. Phosphorylation and dephosphorylation rates for monomers

Reaction	Parameter	Value $[s^{-1}]$	$t_{1/2}$ [min]	Remarks
$[STAT1] \rightarrow [STAT1]_p$	$k^*_{s1\_phos}$	0.900	32	Phosphorylation is described by a nonlinear relation. However, for the first 2 hours this nonlinearity can be neglected and one can obtain the first-order reaction rate and corresponding half-life time of the process
	k_{s1_phos}		N/A	$k_{s1\_phos} = k^*_{s1\_phos} \cdot (1 + k_{s1phos\_sat} \cdot [STAT1](0))$
$[STAT2] \rightarrow [STAT2]_p$	$k^*_{s2\_phos}$	0.427	30	remark analogous to the above
	k_{s2_phos}		N/A	$k_{s2\_phos} = k^*_{s2\_phos} \cdot (1 + k_{s2phos\_sat} \cdot (STAT2)(0))$
$[STAT1]_p \rightarrow [STAT1]$	k_{s1_dephc}	$0.578 \cdot 10^{-3}$	20	Fitted, under assumption that it should be of the same order as the

				should be of the same order as the respective constant for homodimers
$[STAT2]_p \rightarrow [STAT2]$	k_{s2_dephc}	$0.578 \cdot 10^{-3}$	20	Fitted, under assumption that it should be of the same order as the respective constant for homodimers

For interpretation of $k^*_{s1\_phos}$ and $k^*_{s1\_phos}$ and explanation of the same order of half-life times despite different orders of parameter values see equation (1).

Table 4. Second-order reaction rates

Reaction	Parameter	Value $[\mu M^{-1}s^{-1}]$	Remarks
$STAT1_p + STAT1_p \rightarrow [STAT1_p\|STAT1_p]$ $STAT1_p + STAT2_p \rightarrow [STAT1_p\|STAT2_p]$	k_{s1s1} k_{s1s2}	4 10	Since two molecules can't react if they don't collide, the upper limit on rate constants for second-order reactions is imposed by the rate of diffusion, estimated at 109 M-1 sec–1 in water (Voet & Voet, 2004). The rates for homo- and heterodimer formation are quite large in that respect, reflecting the fact that due to the course of phosphorylation process there is quite large possibility of dimerizing right after phosphorylation. These rates are similar to those identified in (Yamada et al., 2003; Zi et al., 2005) for IFN-γ activated pathway
$STAT1_n + [IRF1_{active}]_n \rightarrow [IRF1\|STAT1]_n$	k_{s1i1}	0.01	Fitted
$[STAT1_p\|STAT2_p]_n + [PHY_{inactive}]_n \rightarrow [STAT1_p\|STAT2_p]_n + [PHY_{active}]_n$	$k_{activation}$	0.0007	Fitted
$[STAT1_p\|STAT1_p]_n + [PHY_{active}]_n \rightarrow [STAT1_p\|STAT1_p\|PHY_{active}]_n$	$k_{phys1s1}$	0.7	Fitted

Table 5. Other parameters

Parameter	Value	Remarks
Constitutive mRNA production rates – all in [$\mu M \cdot s^{-1}$], calculated from the steady state assumption (S3)-(S6)		
$k_{s1tprod}$	$0.274 \cdot 10^{-8}$	from (SI3)
$k_{s2tprod}$	$0.182 \cdot 10^{-8}$	from (SI4)
$k_{t1tprod}$	$0.142 \cdot 10^{-8}$	from (SI5)
$k_{l2tprod}$	$0.142 \cdot 10^{-8}$	from (SI6)
Time constant for inertial elements used to model STAT1 gene transcription		
T	42 [min]	
Other parameters		
k_v	0.5	The cytoplasmic/nuclear volume ratio is based on microscopic image analysis of HeLa cells (see Fig.10)
U	$1.2 \cdot 10^{-12}$ [l]	Nuclear volume assumed as in (Goerlich et al., 2003)
V	$0.6 \cdot 10^{-12}$ [l]	Cytoplasmic volume, calculated as $U \cdot k_v$
$k_{s1\_phos\_sat}$	10^4 [μM^{-1}]	Saturation parameter for phosphorylation of STAT1
$k_{s2\_phos\_sat}$	10^4 [μM^{-1}]	Saturation parameter for phosphorylation of STAT2

Table 6. Initial conditions

molecule	value (number of molecules)	Remarks
*STAT*1 (total)	10^5	Assumed; since STATs are one of the main signaling molecules, not only in the analyzed pathway, it is justified to assume large number of them in unstimulated cells
*STAT*2 (total)	$6.67 \cdot 10^4$	It seems reasonable to assume that the number of *STAT*2 molecules is smaller than *STAT*1, taking into account that *STAT*1 is involved in a larger number of processes
*IRF*1 (total)	0	stems from assumption on *IRF*1 mRNA
*STAT*1	$9 \cdot 10^4$	Most (90%) molecules assumed in the cytoplasm; consistent with literature and data (though e.g. (Meyer & Vinkemeier, 2004) reports that in HeLa cells more than 40% of STAT1 is located in the

		nucleus)
*STAT*2	$4 \cdot 10^4$	60% of all molecules assumed in the cytoplasm; though some have claimed that STAT2 is predominantly cytoplasmic in unstimulated cells (Banninger & Reich, 2004), our data indicates otherwise (results not shown)
$STAT1_n$	10^4	10% of all molecules assumed in the nucleus
$STAT2_n$	$2.67 \cdot 10^4$	40% of all molecules assumed in the nucleus
$STAT1_p$, $STAT2_p$, $STAT1_{pn}$, $STAT2_{pn}$, $STAT1_p \vert STAT1_p$, $STAT1_p \vert STAT2_p$, $(STAT1_p \vert STAT1_p)_n$, $(STAT1_p \vert STAT2_p)_n$,	all 0	there is no constitutive phosphorylation of STAT proteins (Chatterjee-Kishore et al., 2000)
*IRF*1, $IRF1_n$, $(IRF1 \vert STAT1)_n$	all 0	stems from assumption about total *IRF*1
*IRF*1 mRNA	0	there is very low basal level of *IRF*1, (Kroeger et al., 2002)
*LMP*2 mRNA	8	Assumed
*TAP*1 mRNA	8	Assumed
*STAT*1 mRNA (*)	2.57	Calculated from steady state assumption (S1). It must be stressed that this is the average number, hence it is not an integer. The value indicates large variance in actual number of proteins, given the proteins' long half-life time and differences that can arise from having even a few more mRNA molecules in a cell
*STAT*2 mRNA (*)	1.71	Calculated from steady state assumption (S2)

List of additional figures

(numbering is continued with respect to figures in the main text)

Figure 9. A schematic illustration of mechanism regulating active dephosphorylation of STAT1 homodimers. ISGF3 (modeled as STAT1|STAT2 heterodimer) induces activation of PHY phosphatase postulated to accelerate dephosphorylation of STAT1 homodimers

Figure 10. EMSA blots showing both cytoplasmic and nuclear ISGF3.

Figure 11. HeLa microscopic images showing the nucleus size.

Figure 12. Comparison of Western Blot data and simulation results for STAT2: a) cytoplasmic and b)nuclear

Figure 13. Hypothetic mechanism of IRF1 cytoplasmic sequestration

Figure 14. Comparison of Western Blot data and simulation results for cytoplasmic IRF1: a) original model presented in the paper, b) modified model based on the assumption of inhibition of IRF1 activation.

Figure 15. Comparison of the main variables in the feedback loop for various values of initial conditions: a) STAT1 homodimer, b) IRF1 mRNA, c) nuclear IRF1 protein, d) STAT1 mRNA. Left column shows standardized results, right column shows unscaled results. In all plots thick line represents results obtained for the basic set of parameters

Figure 4

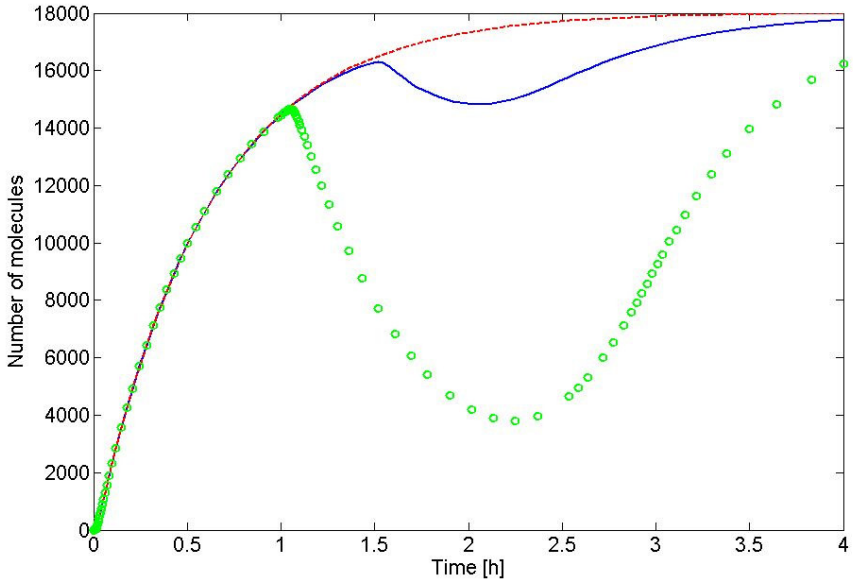


Figure 5

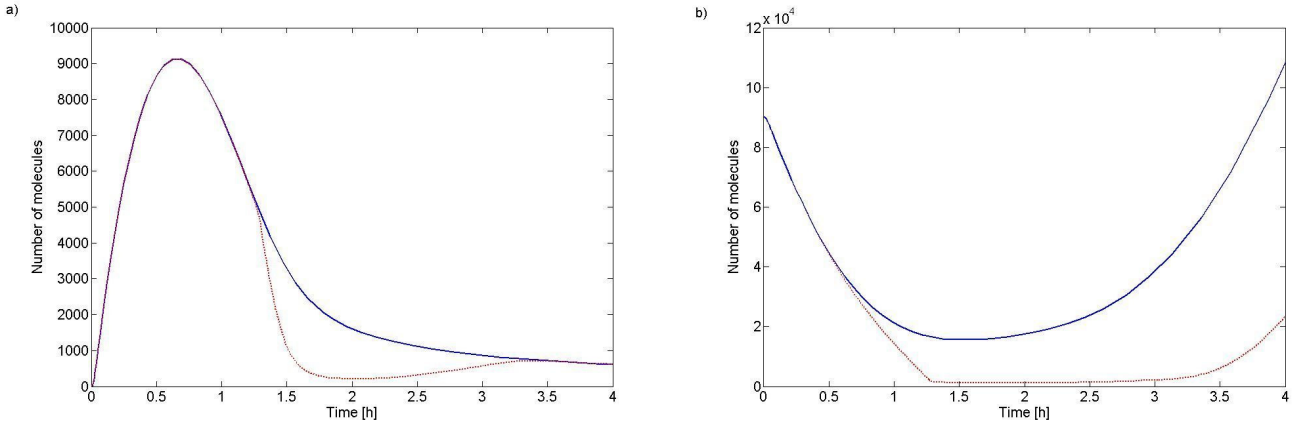


Figure 6

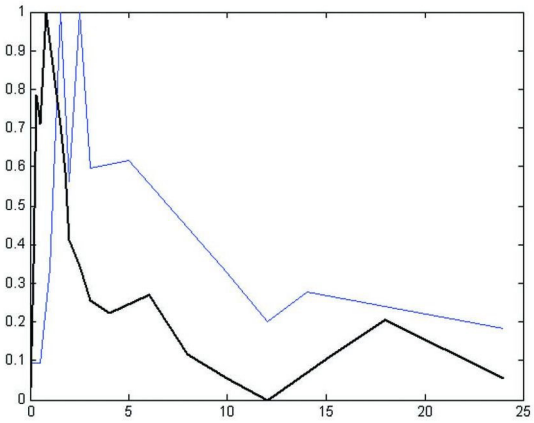


Figure 7

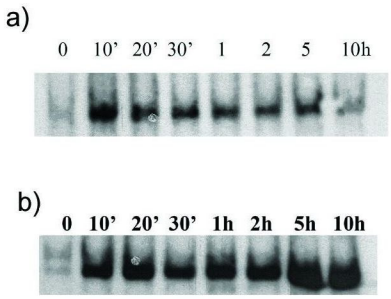


Figure 8

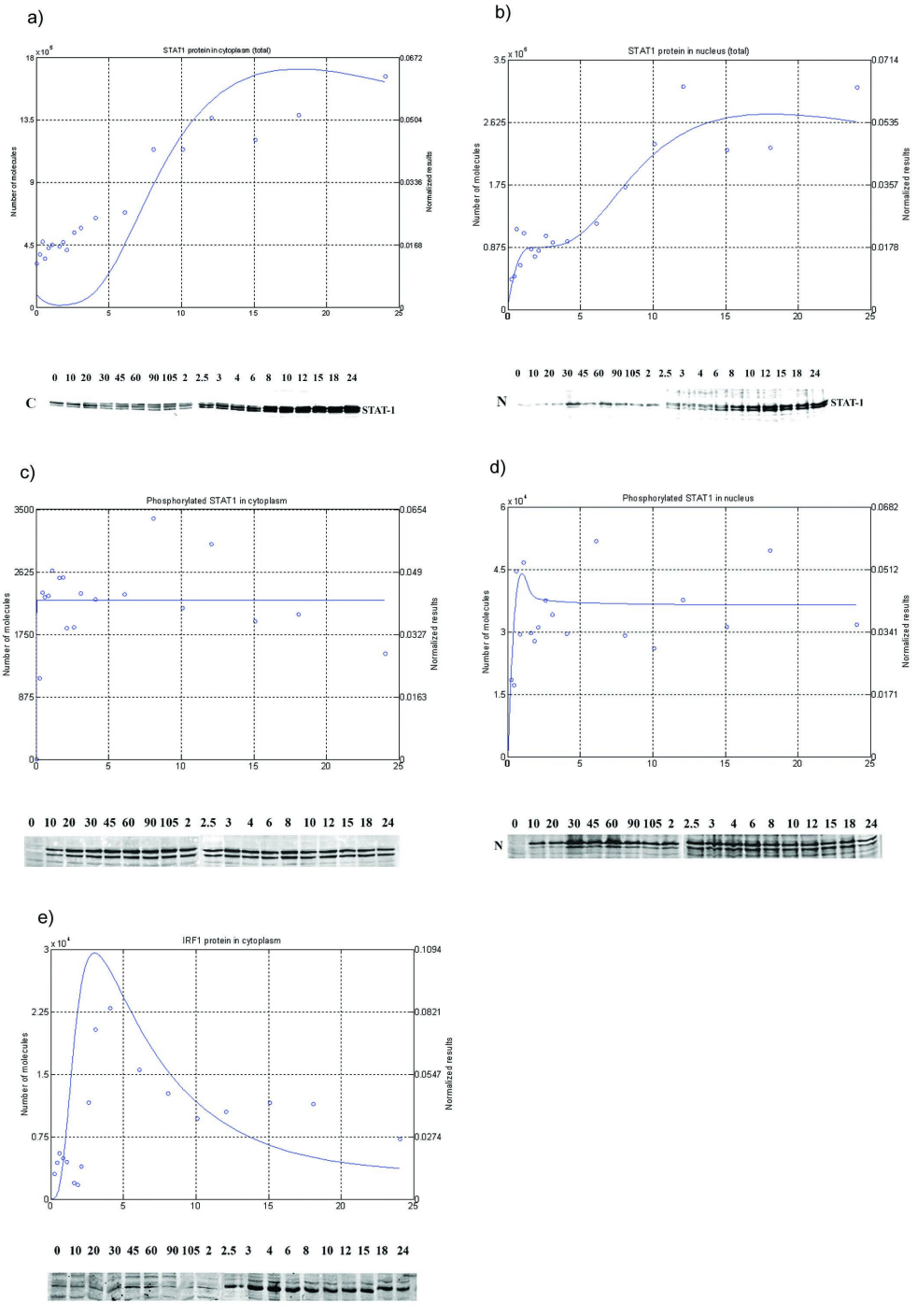

a)
STAT1 protein in cytoplasm (total)
Number of molecules
Normalized results
0 10 20 30 45 60 90 105 2 2.5 3 4 6 8 10 12 15 18 24
C
STAT-1
b)
STAT1 protein in nucleus (total)
Number of molecules
Normalized results
0 10 20 30 45 60 90 105 2 2.5 3 4 6 8 10 12 15 18 24
N
STAT-1
c)
Phosphorylated STAT1 in cytoplasm
Number of molecules
Normalized results
0 10 20 30 45 60 90 105 2 2.5 3 4 6 8 10 12 15 18 24
d)
Phosphorylated STAT1 in nucleus
Number of molecules
Normalized results
0 10 20 30 45 60 90 105 2 2.5 3 4 6 8 10 12 15 18 24
N
e)
IRF1 protein in cytoplasm
Number of molecules
Normalized results
0 10 20 30 45 60 90 105 2 2.5 3 4 6 8 10 12 15 18 24

Figure 9

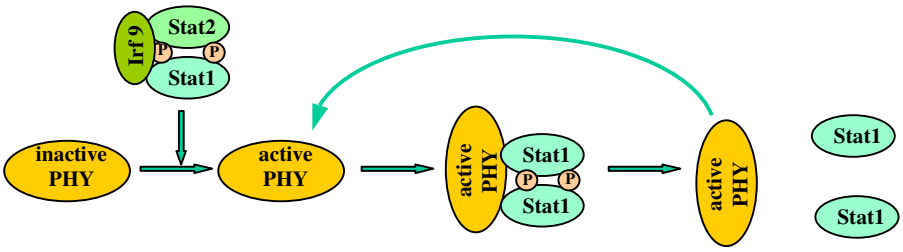


Figure 10

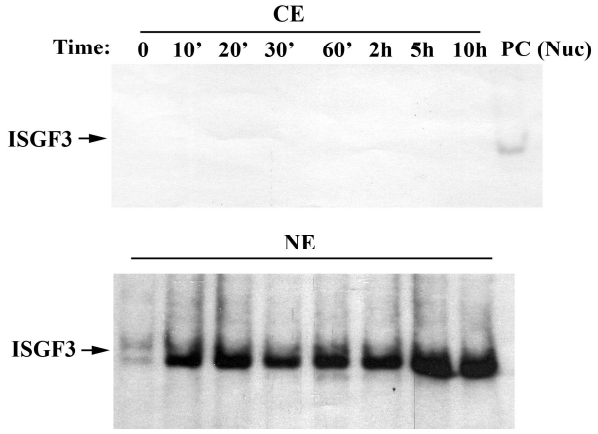


Figure 11

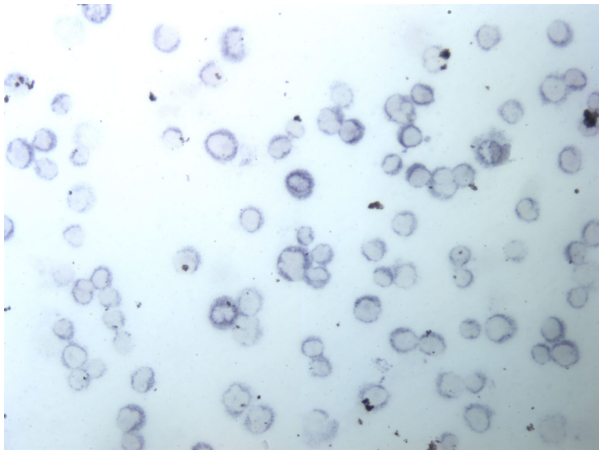


Figure 12

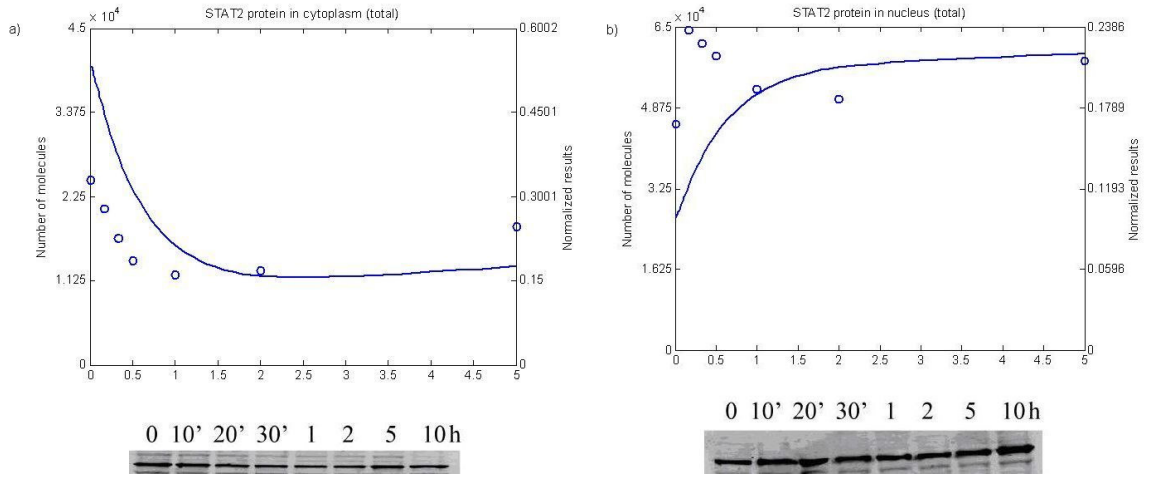


Figure 13

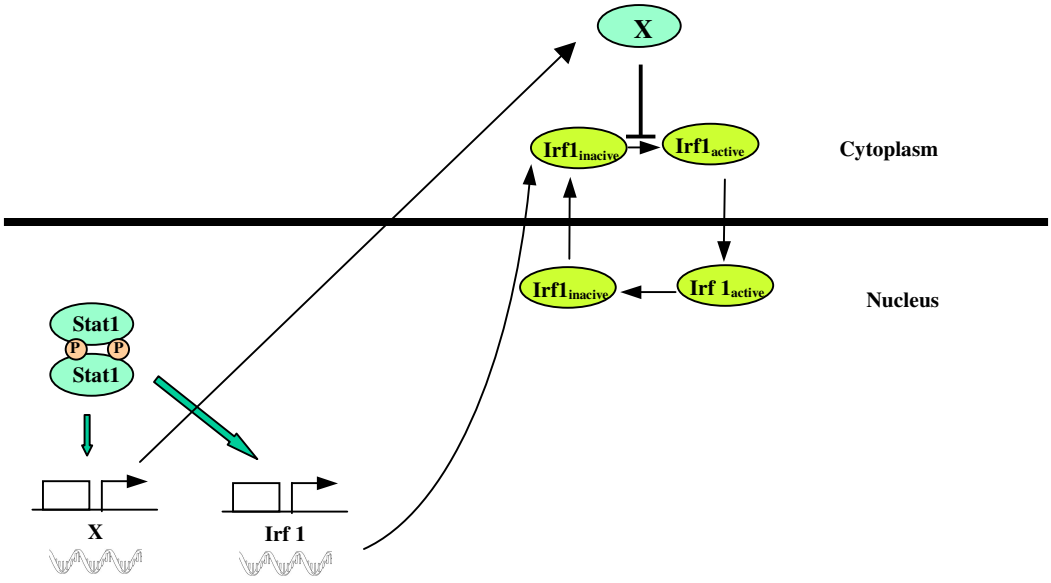
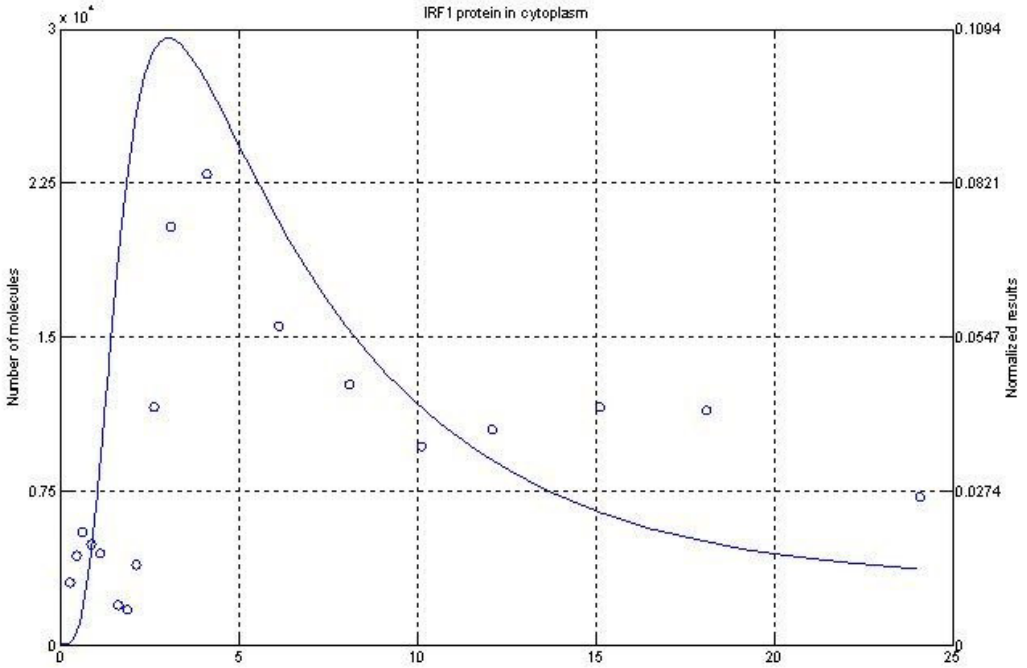


Figure 14

a)


IRF1 protein in cytoplasm
x 10^4
Number of molecules
Normalized results
3
2.25
1.5
0.75
0
0.1094
0.0821
0.0547
0.0274
0
0
5
10
15
20
25

b)

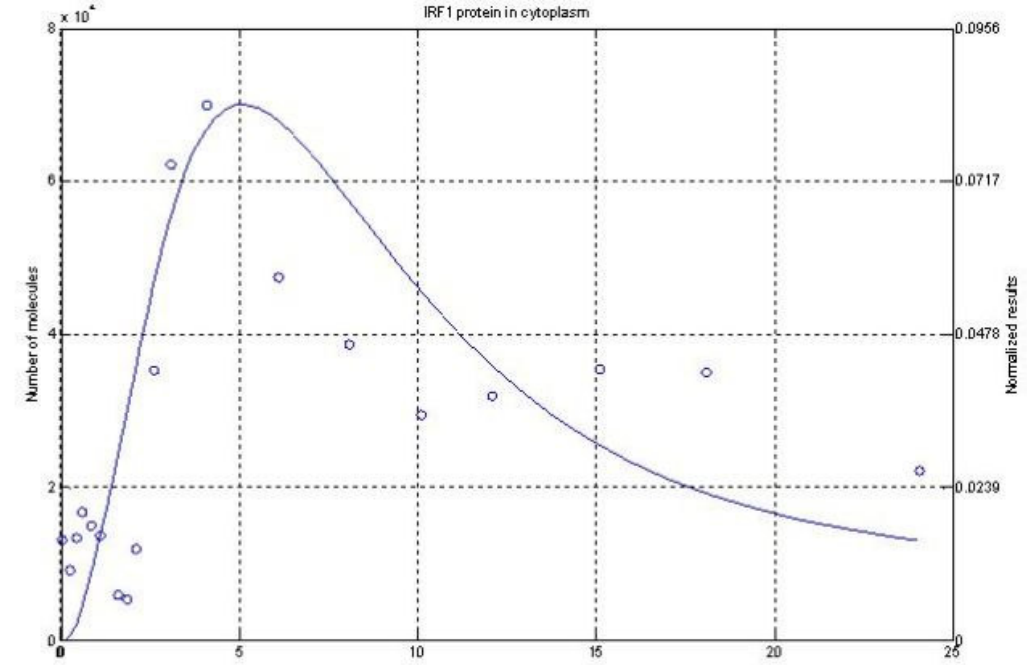
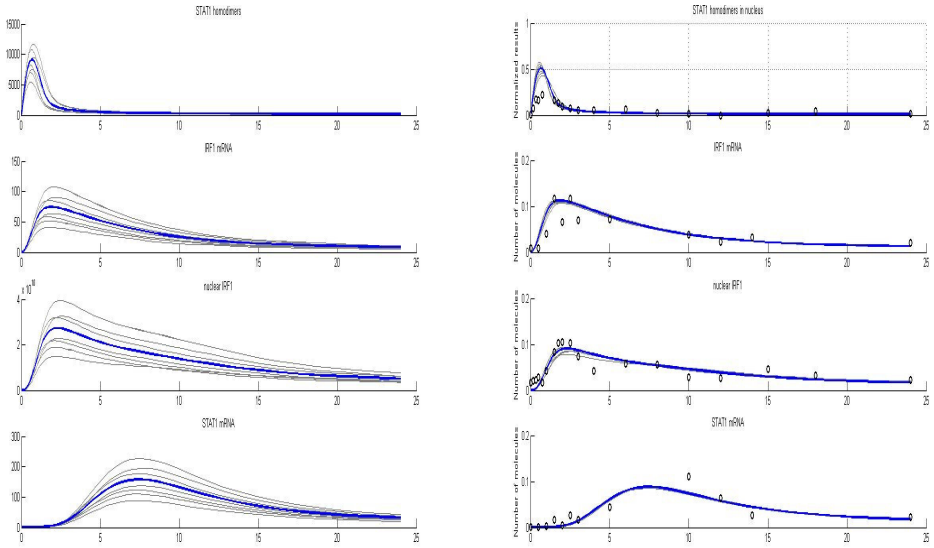

IRF1 protein in cytoplasm
x 10^4
Number of molecules
Normalized results
8
6
4
2
0
0.0956
0.0717
0.0478
0.0239
0
0
5
10
15
20
25

Figure 15


STAT1 homodimers
STAT1 homodimers in nucleus
Normalized results
IRF1 mRNA
Number of molecules
nuclear IRF1
STAT1 mRNA