LUDWIG-MAXIMILIANS-UNIVERSITÄT
TECHNISCHE UNIVERSITÄT MÜNCHEN

Institute of Bioinformatics and Systems Biology (IBIS)

Research group Computational Modeling in Biology (CMB)

Helmholtz Zentrum München

Masterarbeit

in Bioinformatik

Analyzing the behavior of a regulatory T Cell model using an Bayesian sampling approach

Dennis Rickert

Aufgabensteller:	Prof. Dr. Fabian Theis
Betreuer:	Prof. Dr. Fabian Theis
Abgabedatum:	30.04.2012

Erklärung zur Selbständigkeit

Ich versichere, das ich diese Masterarbeit selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

30.04.2012 _____

Dennis Rickert

abstract

The transcription factor NFkB is of high medical relevance, as misregulation of NFkB is associated with a wide spectrum of diseases, ranging from auto-immune diseases, impaired ability to combat infections and cancer to schizophrenia and other neurological disorders. The regulation of NFkB is structured into several highly complex, interacting pathways that have been analyzed with system biological modeling approaches for more than ten years. In this thesis, we adapt an existing model of the canonical NFkB pathway to analyze data generated by novel experimental setups. The established models are based on different biological systems (i.e. fibroblast cells instead of Jurkat cells), which requires us to adapt a significant number of parameters. In addition, our experimental data is significantly less detailed than established models, as we only have access to pooled whole cell extract protein concentrations instead of spatially resolved multi compartment data. Despite these problems, we illustrate that stringent application of parameter inference heuristics allows us to adapt the existing models to explain the observed experimental data. In addition, we are able to show that based on our experimental data, several reactions happen according to complex kinetics which have been overly simplified in the established models. Finally, we show that adapting the model to the resolution of our experimental data allows us to reduce the number of free parameters to the point where Bayesian analysis of the parameter space becomes viable. To our knowledge, all previous models of NFkB suffered from a model complexity that prevent such approaches.

Der Transkriptionsfaktor NFkB hat große medizinische Relevanz, da Misregulation von NFkB im Zusammenhang mit einem großen Spektrum von Krankheiten steht. Diese reichen von auto-immune Krankheiten, eingeschränkter Fähigkeit Infektionen zu bekämpfen über Krebs bis zu Schizophrenie und anderen neurologischen Krankheiten. Die Regulation von NFkB erfolgt über mehrere komplexe, untereinander interagieren Signalwege, welche seit über zehn Jahren mit systembiologischen Modelansätzen analysiert werden. In dieser Arbeit adaptieren wir ein existierendes Model des kanonischen NFkB Signalwegs um die experimentellen Messungen von neuartigen experimentelle Ansätzen aus zu werten. Die etablierten Modelle der NFkB Regulation basieren auf anderen Zelltypen, nämliche Gewebszellen anstelle von Jurkatzellen, wodurch eine Reihe biologischer Parameter neu abgeschätzt werden müssen. Zusätzlich besitzen unsere experimentellen Daten eine geringere Auflösung als die der etablierten Modelle, da wir nicht in der Lage sind zwischen Proteinen in Zytoplasma und Zytosol zu unterscheiden. Wir zeigen das es durch Anwendung von Parameter Inferenz Heuristiken möglich ist, die existierenden Modelle an unsere Datenauflösun zu adaptieren. Zusätzlich sind wir in der Lage zu zeigen das, basierend auf unseren neuen experimentellen Daten, einige Reaktionen in den etablierten Modellen zu sehr vereinfacht werden und besser durch andere Ansätze modeliert werden. Zuletzt zeigen wir das wir, indem wir die Komplexität des Models an die Auflösung unserer experimentellen Daten anpassen, in der Lage sind Bayesianische Parameter Analyse durch zu führen. Dies war unserer Wissens nach für die etablierten Modelle aufgrund ihrer größeren Komplexität nicht möglich.

Contents

Figure overview:

1. Introduction

In this thesis we analyze the activation of the transcription factor NFkB in T cells based ***on ODE models of the canonical NFkB pathway***. Models of this pathway have been studied for more than ten years, yet no gold standard model exists and new experimental evidence frequently leads to the adaption and refinement of existing models.

The family of NFkB type transcription factors is involved in the regulation of a wide range of different tissue- and celltypes. ***Misregulation or disruption of NFkB*** signaling is associated with numerous serious diseases, such as, but not limited to auto immune diseases, HIV, smallpox and other bacterial and viral infections, different types of cancers and schizophrenia.

The common property of all members of the NFkB family is the ability to ***bind to a common DNA motif***. Based on the presence of the binding motif in the regulatory region of a gene, hundreds of potential targets for the NFkB family can be identified. These targets belong to a wide spread of different categories, including various types of proteins relevant for immune responses (cytokines, chemokines, immunoreceptors, proteins involved in antigen presentation), stress- and apoptosis signaling, regulation of cell surface and membrane properties and additional transcription factors.

The wide range of different NFkB targets is reflected in the number of biological processes NFkB activation is a part of. NFkB frequently acts as an ***anti-apoptotic signal*** in normal tissue, mitigating pro apoptosis signals such as those resulting from DNA damage. In this role NFkB is often found to be constitutively active in tumors. On the other hand NFkB is involved in ***immune responses*** and the activation of NFkB is an essential step in both the adaptive and the innate immune response. Beyond this, NFkB is also known to have tissue specific roles, for example in neurons.

Given the diversity of processes NFkB is involved in, it is not surprising that it also activated and regulated in different, partially tissue specific ways. The ***activation of NFkB*** is frequently divided into the so called ***classical*** and the ***alternative pathway***, based on the molecular mechanisms result in NFkB activation. Of these two pathways, the canonical pathway is generally considered to be better understood.

The earliest computational model of the canonical pathway that explored the regulation and feedback loops involved in NFkB activation was published in 2002 [1:Hoffmann, A. et. al.]. As reviewed by [2:Cheong, R. et. al.] ***over twenty different variants of the model*** have been derived in the following six years, each with adaptions ranging from small modifications (slightly different parameters, exchange of reaction dynamics) to redesign of large parts of the model topology (addition of multiple different IkB proteins).

A common property of most of these models is that they utilize ***manually adapted parameter sets*** to explain the experimental observation. While this allows the researchers the pick biologically conceivable parameters, it is an approach that can limit the value of predictions derived from the model. Models with manually optimized parameters show ***researcher bias*** that is transferred into all predictions based on this model. In contrast, ***Bayesian approaches*** to parameter optimization supply an ensemble of valid parameter sets. Predictions based on such an ensemble will often show less researcher bias, but require a model that is accessible to numerical optimization.

One aspect that limits the analysis of existing models using numerical optimization is the ***high degree of detail*** found in these models. For example, all published models are ***multi compartment models***, that include both cytoplasm and nucleoplasm. This introduces a significant number of transport reactions in the model. Estimating the transport rates based on experimental observations is a challenging task, as experiments that differentiate between different compartments are very involved.

Another example are ***enzymatic reactions***. In most models of the NFkB pathway, these are frequently modeled as a set of binding, processing and unbinding reactions. While such an approach ***illustrates*** in detail what happens on a molecular level, other ways to model enzymatic reactions, such as the ***Michaelis Menten approximation***, introduce fewer free parameters into the model and are often more accessible to numerical optimization.

In this thesis we will study the following questions:

- Can we adapt the existing computational models, which are mostly based on observation of fibroblasts to explain the stimulation patterns observed in Jurkat cells / T Cells?
- Which degree of model complexity is justified for experimental data based on whole cell pooling (i.e. data that doesn't differentiate between cytoplasm and nucleoplasm)?
- For our chosen degree of model complexity, is it computationally feasible to run a Bayesian parameter analysis or are we limited to a maximum likelihood / researcher biased model analysis?
- Which technical approaches are required to optimize parameter sampling?
- How can parameter ensembles based on Bayesian sampling help our understanding of the regulation of NFkB?

2. Background Information

The NFkB/Rel family of transcription factors is of high clinical relevance and has been studied for more than twenty years, both from a more theoretical modeling perspective and in clinical studies. The ***diseases most commonly associated*** with NFkB/Rel misregulation are auto immune diseases, chronic inflammation, immunodeficiency and various types of cancer [3:Courtois, G. et. al.]. In its role as a major regulator of the immune response, ***NFkB is a target of various viral and bacterial infections*** that try to suppress the immune response by inferring with the regulation of NFkB [4:Hiscott, J. et. al.]. In addition, NFkB has also other, more cell type specific roles, for example regulating synaptic plasticity in neuronal cells [5:O'Neill, L. A. et. al.].

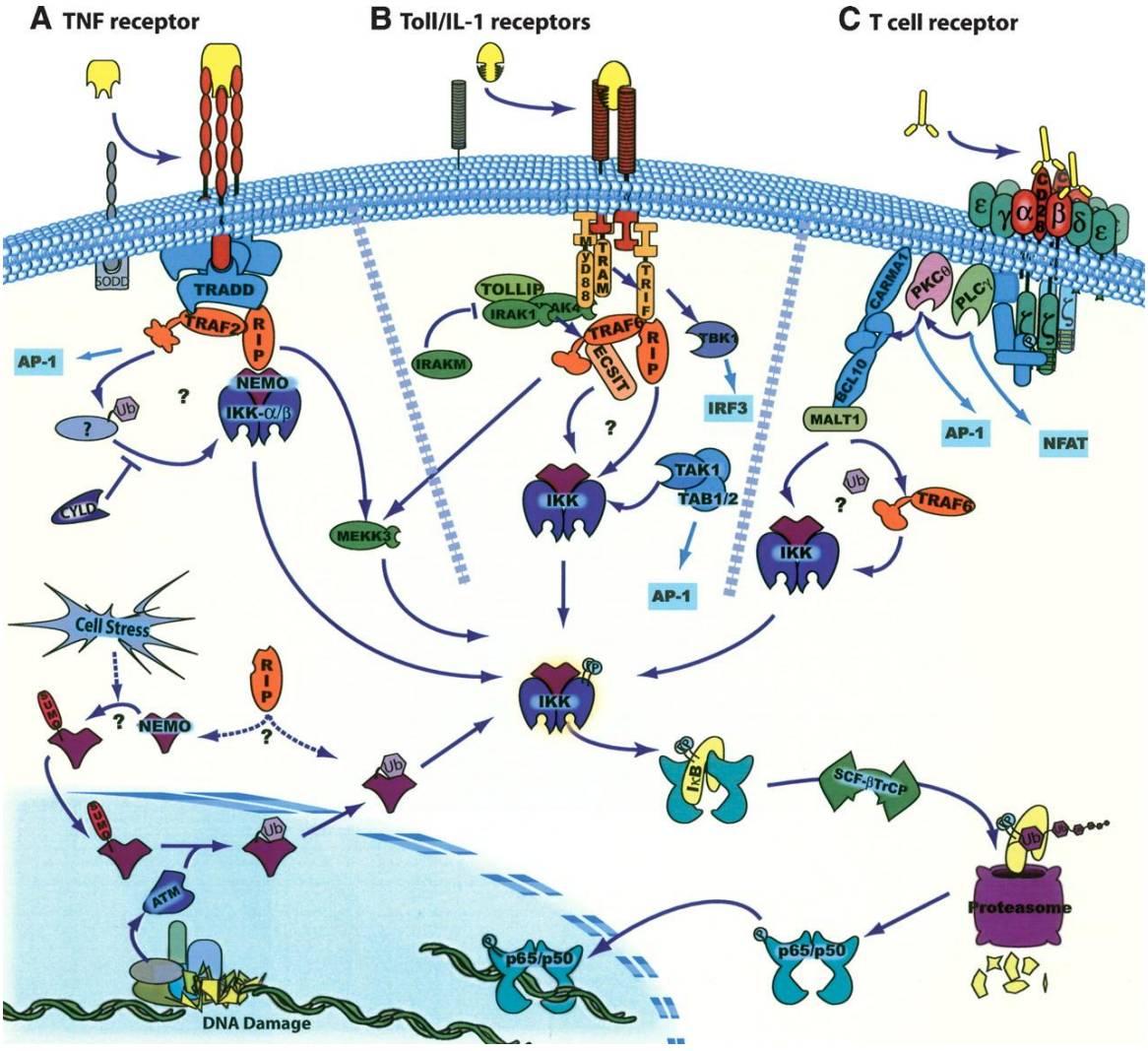


Figure 1: Initiation of NFkB signaling can be induced by different receptors and cellular events.

Figure from [6:Hayden, M. S. et. al.]. Different signaling events result in the activation of NFkB. Activation of TNF-, Toll/IL- and T Cell receptors all result in the formation of large membrane bound protein complexes that start signaling cascades which eventually result in the activation of NFkB. However, intracellular factors like DNA damage or cellular stress are also able to activate NFkB.

NFkB activation happens along multiple ***different signaling chains***, initiated by a variety of cellular events. The most frequent are stimulation of TNF-, T Cell- and Toll/IL-1 receptors, cellular stress and DNA damage[6:Hayden, M. S. et. al.]. Each of these pathways involves a significant number of individual molecular players as summarized in Figure 1. Most of these pathways are connected via crosstalk events, further complicating their analysis.

If all pathways are considered, more than 780 molecules are known to inhibit the activity of NFkB either directly or indirectly [7:Gilmore, T. D. et. al.]. This illustrates the need for a structured approach to analyze NFkB signaling. The most common differentiation in NFkB pathways is the distinction between the ***classical*** and ***alternate pathway***.

2.1. Pathways of NFkB activation

The distinction between the classical (or canonical) and alternate (or non-canonical) pathway of NFkB activation is mainly based on the ***final steps of each pathway that lead to NFkB activation***. Both in the canonical and the non-canonical pathway a multimeric protein complex including multiple IKK (IkB Kinase) subunits is assembled and activated. The combination of catalytic and regulatory subunits of the complex determines both its substrate specificity and its classification as either classical or alternate pathway. Both pathways of NFkB activation have been extensively reviewed, e.g. in [8:Bonizzi, G. et. al.].

Classical pathway of NFkB activation

The key IKK complex in the classical pathway consists of the catalytic subunits IKKα, IKKβ and the regulatory subunit IKKγ or NEMO (NFkB essential modulator). Together, they form a complex that ***phosphorylates IkB molecules,*** most notoriously IkBα.

IkB proteins have an ankyrin repeat called motif, that allows them to bind to NFkB, especially p65/p50 (RelA/RelB). Binding of IkB to NFkB ***inhibits the ability of NFkB to bind DNA,*** masks the nuclear localization signal (preventing NFkB from being imported into the nucleus) and results in the export of the bound IkBaNFkB complex from the nucleus into the cytoplasm.

Phosphorylation of IkB causes its rapid degradation. The exact mechanisms of this degradation are currently still under investigation. A sufficiently strong IKK activity can result in the phosphorylation and subsequent degradation of almost the complete IkBα pool of a cell in 15 minutes, resulting in the unbinding and ***activation of the previously sequestered NFkB***.

Upon degradation of IkB, NFkB is imported into the nucleus and acts as a transcription factor. An interesting aspect is that in doing so, NFkB constitutes a ***negative feedback*** to its own activity, by inducing the ***synthesis of additional IkBa***.

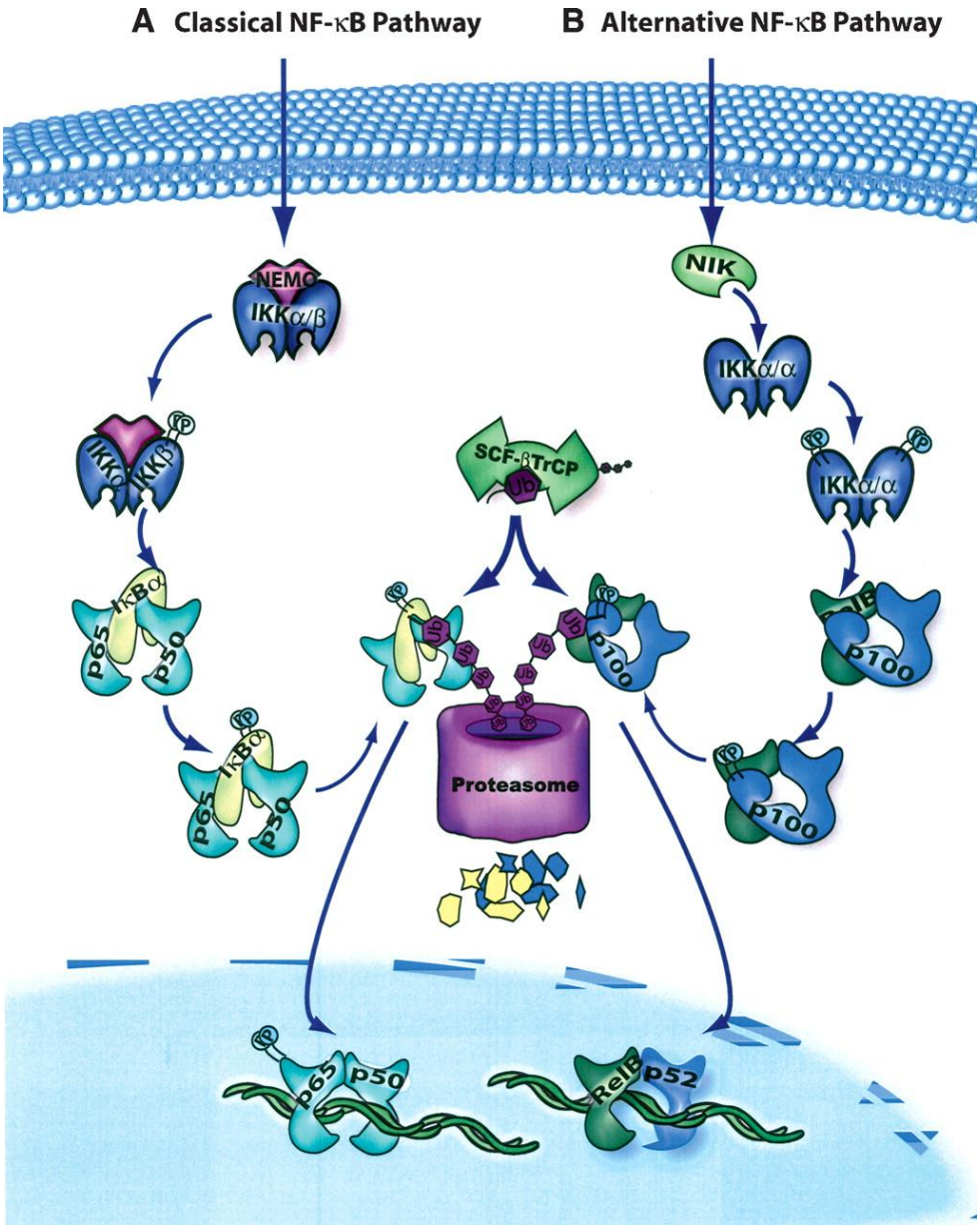


Figure 2: Final steps of NFkB activation via different pathways converge with the activation of different IKK complexes

Figure from [6:Hayden, M. S. et. al.]. The large number of different upstream activation signals of NFkB converge on the level of active IKK complexes. In the classical pathway, a complex consisting of IKKα, IKKβ and NEMO is generated. This complex activates NFkB by degrading IkBa. In the alternate pathway, the different activation signals converge by the activation of an IKK complex consisting of two IKKα units.

The IKK complex, especially the NEMO subunit, serves as a basis to ***unify and integrate diverse upstream*** signals. Not all of the regulators, activators and inhibitors of this complex and their connection to the upstream signaling events that result in their activation are known. However, an important step is the binding of ***ubiquitin chains*** to NEMO. As discussed in [9:Oeckinghaus, A. et. al.], this is facilitated by the Carma1–Bcl10–Malt1 (or ***CBM***) complex. This complex is in turn activated by the protein kinase ***PKC0***.

In addition, it is known that the CBM complex is ***regulated by the protein A20***. This has been studied in detail by e.g. Düwel, Krappmann et.al. e.g. in [10:Duwel, M. et. al.]. A20 can deubiquitinate the Malt1 subunit of the CBM complex, which prevents IKK and CBM interaction and serves as an inhibitor of NFkB. A20 is an additional protein that ***integrates multiple cellular signals***. Upon T Cell stimulation, A20 is degraded via the proteasomal pathway. Once NFkB is activated via the canonical pathway, it induces the transcription of additional A20, resulting in a net recovery of the available A20 protein. In this way, ***A20 constitutes a second negative feedback loop*** for NFkB activity.

The classical pathway of NFkB activation is frequently stimulated as part of the ***innate immune response***. The innate immune system is the non-adaptive or non-specific part of the immune system, which involves responses such as ***activation of leukocytes*** and ***inflammatory responses***. The most frequent inducers of the canonical pathway of NFkB regulation are TNF-, T Cell- and Toll-like receptors.

Alternate pathway of NFkB activation

In contrast to the canonical pathway, the alternate pathway is less well explored. It is known to influence both development of ***lymphoid tissue*** and regulate ***adaptive immune responses***. The induction of this pathway results in the formation of an IKK complex consisting of two IKKα subunits without either IKKβ or NEMO. It is activated by the NFkB-inducing kinase (NIK).

The active IKK complex of the alternate pathway does not target IkBα. Instead, it induces the ***proteolytic processing*** of the inactive NFkB family member p100 into the active form p52. p100 has an ankyrin repeat domain in addition to its DNA binding domain. Prior to processing, this domain binds to p100 itself, inhibiting its transcription factor activity in a way similar to the way IkBα inhibits other NFkB molecules.

2.2. Key proteins of the canonical NFkB pathway

We already briefly mentioned the key proteins of the canonical NFkB pathway in the previous section. However, for the design of a computational model that recapitulates the canonical pathway, it is useful to summarize the most relevant properties of each member of the pathway. The various NFkB, Rel, IkB and IKK proteins are summarized in Figure 3.

NFkB family of proteins

The members of the NFkB family of proteins all share the so-called ***Rel-homology-domain RHD***. This domain enables NFkB proteins to bind to the NFkB specific DNA promoter motif and thus regulate the transcription of genes. However, NFkB only binds to DNA after ***forming NFkB dimers***. A wide variety of different dimers, consisting of either identical (homodimers) or different (heterodimers) member of the NFkB family are known. The different possible dimers have varying effects on the target gene transcription.

In addition, the Rel members of the family, RelA, RelB, c-Rel, Dorsal and Dif also have a ***transcription activating domain*** (TAD). This domain enables the initiation of the gene transcription. NFkB dimers that contain at least one member of the Rel group usually promote gene transcription, whereas dimers that lack a TAD are inhibitory.

Rel proteins are inhibited by ***binding to IkB monomers***. In contrast, the proteins p105/p50 and p100/p52 are synthesized in a ***precursor form***. They carry an ankyrin repeat domain, that blocks the DNA binding domain of the precursor form in a way that is similar to the inhibition of Rel proteins by IkB monomers.

IKK family of proteins

The IKK family of proteins consists of the different subunits IKKα, IKKβ and NEMO. Different sets of subunits can form active IKK complexes, that act ***by phosphorylating specific target sites*** of other proteins. For the canonical pathway, this complex consists of each one subunit of IKKα, IKKβ and NEMO. While IKKα and IKKβ are the catalytically active subunits, NEMO is assumed to regulate the activity of the IKK complex.

The IKK complex of the canonical pathway specifically targets IkB proteins for phosphorylation. The NEMO subunit interacts with several other proteins which ***regulate the activity of the IKK complex***. One example for such an interaction is the ubiquitin mediated interaction with the Carma1–Bcl10–Malt1 (CBM) complex. This interaction is required for the activity of IKK. Some inhibitors of NFkB, for example the ***protein A20*** disrupt these ubiquitin chains and so inhibit NFkB.

IkB family of proteins

Different IkB proteins all act as inhibitors of the various Rel proteins. They ***bind to the Rel-type NFkB molecules*** via the ankyrin repeat domain and prevent bound NFkB from acting as a transcription factor. In addition, IkB proteins can acts as a ***"shuttle"*** that moves bound NFkB from the nucleus to the cytoplasm. For IkBa, the details of this process have been studied e.g. by [11:Sachdev, S. et. al.]. Unbound IkBa has a nuclear localization signal (NLS) in the second ankyrin region, that causes its import into the nucleus. Once it binds to NFkB proteins in the nucleus, not only does it block the NLS of NFkB, but the NLS IkBa also becomes blocked. The details of these processes might be different for the different IkB isoforms, justifying the assumption of different import- and export rates for the different proteins.

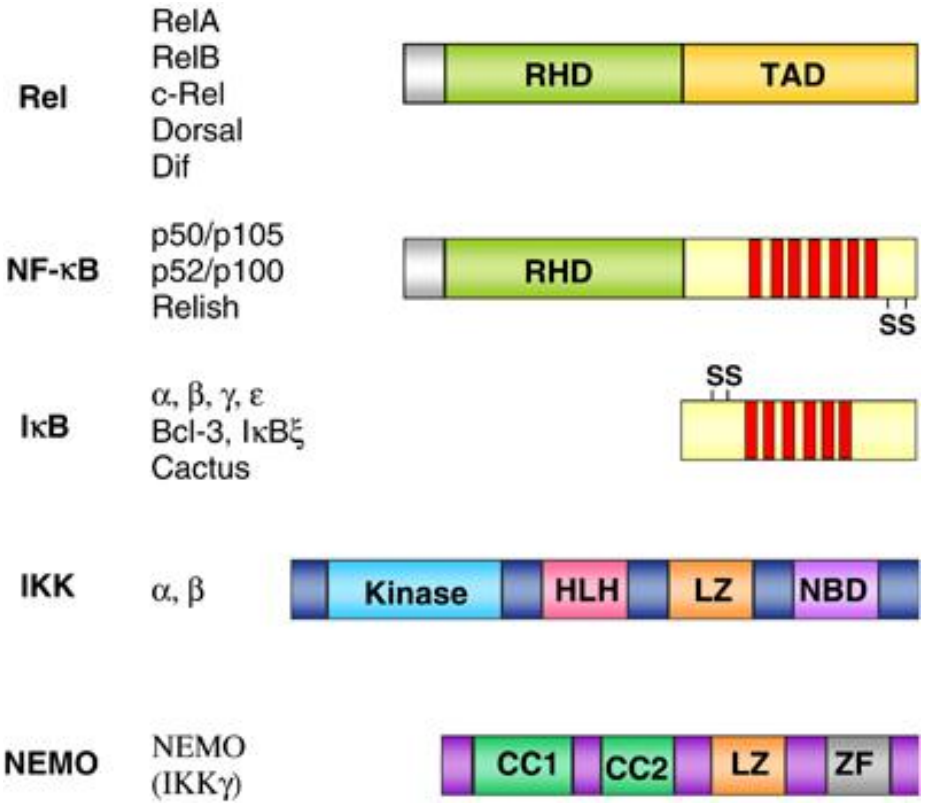


Figure 3: Overview of various key proteins of NFkB regulation and their key features

Figure from [12:Gilmore, T. D.]. Key feature of NFkB/Rel proteins is the RHD, the Rel homology domain which facilitates binding to the NFkB specific DNA binding motif. Beyond this, the Rel proteins also have a transcription activation domain (TAD), which allows them to start gene transcription once they're bound to the DNA. Both NFkB proteins and IkB proteins have an ankyrin repeat domain (sequence of red stripes), that binds to- and block the RHD. IkB binds to Rel and so inhibits its transcriptional activity. In contrast, the ankyrin repeats of NFkB have an auto inhibitive function. The RHD of a NFkB protein is blocked by its own ankyrin repeat domain. Only when NFkB is proteolytically processed is this domain removed and NFkB becomes activated.

Unbound IkBa monomers are rather ***unstable proteins***, even when not actively degraded. The half-life time of unbound IkBa has been estimated to be around 40 minutes, however some cell type specific variation of this rate have been reported [13:Krappmann, D. et. al.]. Binding of IkBa to NFkB stabilizes the IkBa protein, possible by refolding of some domains of IkBa. However, the turnover of the bound IkBa NFkB complex has been issue of multiple experiments with partially contradictory results. We discuss this in more detail in section 6.1.5.

The ***phosphorylation of IkB*** adds another layer of complexity to estimating the stability of IkB. Initially it was assumed that IKK has higher affinity for bound IkBa than unbound IkBa. However recent studies suggest that the higher ratio of phosphorylated bound IkBa reported in older studies might result from the higher concentration of bound IkBa found in vivo in cells, and that ***IkBa is targeted by IKK*** with the same specificity whether it is bound or unbound [14:Mathes, E. et. al.]. In addition, Mathes et.al. report that the phosphorylation of unbound IkBa apparently does not result in its immediate degradation; only when IkBa is both phosphorylated and bound to NFkB it is degraded rapidly. This observation is very recent; all established models of the NFkB pathway that we are aware of make no distinction in the degradation between bound- and unbound phosphorylated IkBa.

A20/TNFAIP3

A20 is also known as TNFAIP3 (tumor necrosis factor alpha induced protein 3). It is one of the feedback proteins that both ***regulates- and is regulated by NFkB***. In addition, A20 is also degraded by upstream event of TCR signaling. Knockout studies suggest that ***A20 is required to end NFkB activity associated with inflammatory processes***; mice with A20 total knockouts frequently suffer from chronic inflammation.

Both based on its domain structure and on experimental evidence, A20 is known to act as an ***ubiquitin editing enzyme***, able to add, remove and modify existing ubiquitin chains. A known target of A20 is the CBM complex that is important in the activation of IKK in the canonical pathway. A20 deubiquitinates a subunit of the CBM complex, preventing the interaction of IKK and the CBM complex [10:Duwel, M. et. al.]. However, beyond this A20 also targets multiple other members of the NFkB pathway, e.g. TRAF and RIP. [15:Heyninck, K. et. al.]. Due to this, while its general role as an inhibitor of NFkB activity can be considered established, the detailed mechanisms are still subject to research.

3. Material

3.1. Modeling of the canonical NFkB pathway

Computational modeling of the canonical NFkB pathway was started around 2002 with the model published by [1:Hoffmann, A. et. al.]. They choose to model NFkB activation using an ***ODE model*** based on mass action kinetics. In the last ten years, this model has been adapted a significant number of times, however some features a shared between almost all descendant models.

Almost all models of the canonical NFkB pathway are ***two compartment models*** that make a distinction between proteins found in the nucleus and in the cytoplasm. As a result, core import/export processes are essential parts of these models.

The parameters of most models have been adapted manually, often based on ***biological reasoning*** such as an assumption of the largest conceivable biological reaction speed. However, as the Authors freely admit, not all parameters in the different iterations of the model are biologically reasonable. Parameters that could not be determined based on biological reasoning have been adapted manually based on a manually implemented maximum likelihood approach. Most models show significant degrees of parameter under determination. This under determination has been subject to multiple computational studies, however most of these studies have been focused on the analysis of the researcher biased parameter set. Global studies of parameter under determination, e.g. by utilizing Bayesian approaches are not known to us.

A large number of experiments utilized to determine the valid ***parameter ranges have been based on fibroblast cells***. Even model iterations that study different cell types frequently reutilize these parameters. This can lead to some problems when other cell types are studied. For example, it is known that the nucleus to cytoplasm ratio differs strongly between T Cells and fibroblasts. As this ratio is essential to the reaction rate of core import-export processes, the established models might be unsuitable to analyze T Cell based experiments.

3.1.1. Modeling by Hoffmann group

The group of Alexander Hoffmann is probably the predominant group in the design and study of computational models of the canonical NFkB pathway. Not only did they publish the original model, but they also contributed to a large number of adaptions to this initial model.

The general focus of most of their models is the regulation of NFkB activity that happens ***below the level of IKK activation***, such as the interaction of different IkB isoforms with NFkB. The input of these models consists of IKK activity profiles, i.e. quantified profiles of IKK activation. In contrast ***molecular complexes upstream of IKK*** are mostly ignored. This includes the negative feedback loop constituted by A20 transcription induced by NFkB. The most recent model

provided as web model on their group homepage ([16:Basak, S. et. al.], illustrated in Figure 4) recapitulates the interaction of ***three different IkB isoforms*** with NFkB and crosstalk events to the alternate pathway.

While the models by Hoffmann et.al. recapitulate the regulation of NFkB below the level of IKK with a remarkable degree of detail that is supported by significant amounts of experimental data, they are of limited use for the work presented in this thesis. The experimental effort required to measure and quantify multiple different IkB isoforms is significant and exceeds the resolution provided by our experimental data. As we are not able to match the resolution of the model, mapping of our experimental quantifications to the model states is not possible.

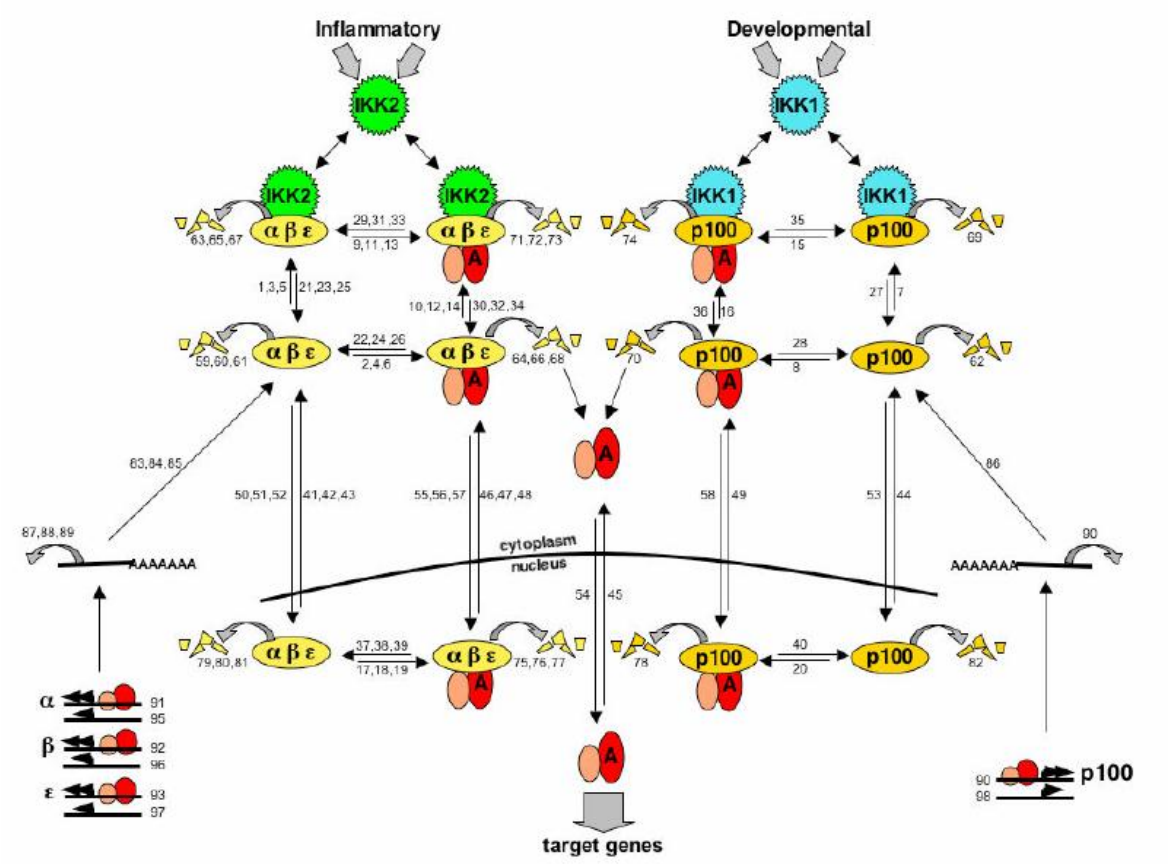


Figure 4: Most recent web model of NFkB regulation provided by Hoffmann et.al.

Three different IkB isoforms bind and inhibit NFkB and are degraded by IKK2, the IKK complex of the canonical pathway. In addition, crosstalk to IKK1, the complex of the alternate pathway is included. The model input are IKK2/1 activity profiles.

3.1.2. Modeling by Lipniacki group

In contrast to the models by Hoffmann, the model published in [17:Lipniacki, Tomasz et. al.] includes the ***A20 feedback loop***. The regulation of NFkB activity is started upstream of IKK, using binary on/off switches that correlate to the model stimulation by TNF. Another major difference between this model and the models of the Hoffmann group is that the ***different IkB isoforms are summarized*** a single IkB state. The model is visualized in Figure 5. As we specifically aim to analyze the ***result of A20 knockdown experiments***, this model is more suited to our needs than the Hoffmann et. al. models which do not include the A20 feedback loop.

However, we still have to make adaptions to the model to bring it to the same resolution as our experimental data. Currently, we are limited to ***pooled experimental measurements*** that quantify the amount of protein in the complete cell, but do not provide a distinction between proteins in the cytoplasm and in the nucleus.

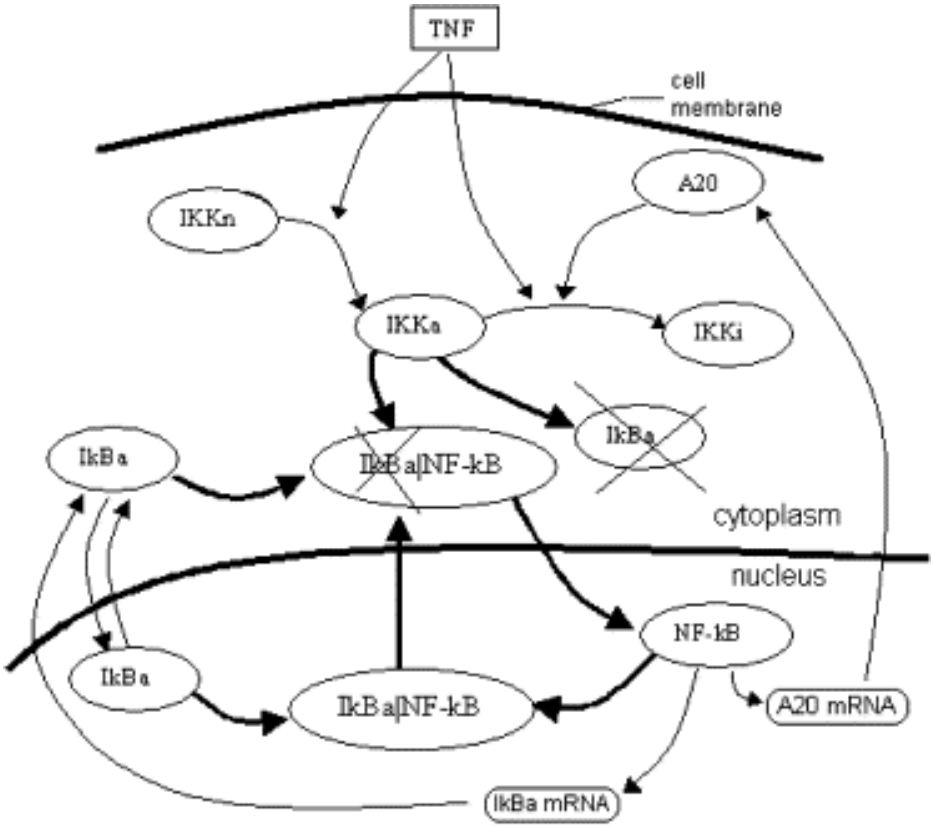


Figure 5: Original model by Lipniacki et.al.

The original model by Lipniacki contains only one IkB isoform. It is focused on including regulatory events and feedback loops upstream which target IKK. Visualization taken from original Lipniacki paper.

3.2. Experimental measurements

Experimental measurements and quantifications were provided by Michael Düwel, Daniel Krappmann and Richard Griesbach from the Helmholtz Zentrum München. All measurements of protein concentrations were performed using the pooled lysate of multiple cells. Cells were lysed completely, resulting in total cell protein concentrations instead of separate cytoplasm and nucleoplasm concentrations. A20 and IkBa protein concentrations were quantified using western blotting and suitable antibodies. In addition to A20 and IkBa, Actin concentration was determined and used as a housekeeping protein to normalize measurement intensities. NFkB concentration was quantified using EMSA (electrophoretic mobility shift assay). In addition to NFkB, the unspecific band on the EMSA gels was also quantified and used for data normalization. In detail, we analyzed the following experimental measurements. The raw experimental quantifications of all states found in the model by Lipniacki are visualized in Figure 6. Two different experimental setups were analyzed in two replicates each:

Two replicates of stimulated mock transfected / wildtype cells Jurkat cells.

The experiments siGFP I and siGFP II were performed by stimulating Jurkat cells transfected with "mock" siRNA, i.e. siRNA that does not affect any genes relevant to NFkB signaling. This was done to ensure that differences between cells with and without A20 knockdown were not due to the transfection process, which can cause significant stress for cells. For our analysis, these experiments are considered wildtype cells. Cells were observed for ten hours, with more frequent measurements in the first hours. In detail, measurements were taken at t = 0, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 6.0, 7.0 and 10.0 hours.

Two replicates of stimulated A20 knockdown cells.

Two stimulation experiments were performed with Jurkat cells that had been transfected with A20 siRNA (siA20 I/II), causing a knockdown of both the base- and induced A20 protein level. Each experiment was performed together with one of the mock transfection replicates, and protein concentrations were determined by western blotting on the same gel. This allows estimation of knockdown efficiency by comparing the measured intensities at time 0 hours. In contrast, the absolute intensities of different experimental setups are not comparable, as absolute values from different gels are shifted by a normalization factor. The time points at which protein concentrations were quantified were identical to the wildtype measurements.

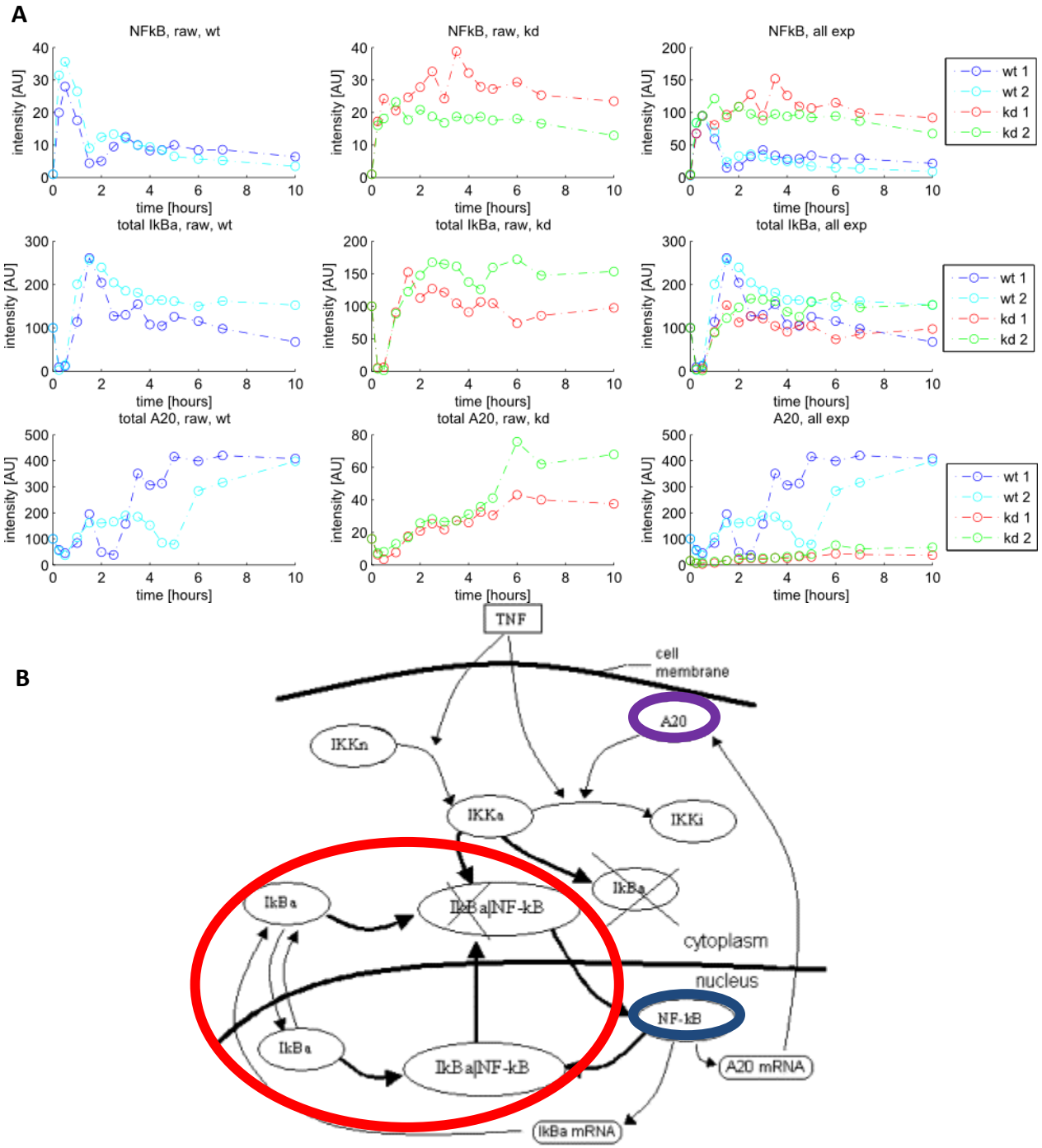


Figure 6: Overview of raw experimental data and pathway states measured

A. Overview of raw experimental data for wildtype and knockdown cell lines. From left to right: wild-type, knockdown and summary of all experiments. From top to bottom: Measured free NFkB, total IkBa, total A20.

B. Original visualization of the model by Lipniacki et.al. Colored circles correspond to experimentally measured concentrations. Red circle: total IkBa (bound and unbound, nucleoplasm and cytoplasm). Blue circle: total free NFkB. Purple circle: A20 protein. The IkBa measurement (red) summarizes a total of two distinct molecular species in two different compartments.

3.3. Normalized experimental data

The experimental measurements as visualized in Figure 6 are the experimental basis for our analysis of the NFkB pathway. This data has already been normalized to account for some systemic experimental variances, such as uneven illumination of the gel membranes. In addition, we normalized the relative cell number using the housekeeping protein Actin (for IkBa and A20 concentration) and the unspecific EMSA band (for NFkB).

Based on how the different experiments were distributed over multiple membranes, the illustrated values are only of limited comparability. Protein concentrations quantified on different gels are usually shifted by an unknown experimental factor. This is not a problem within each time course, but only relevant when comparing different experimental replicates.

For IkBa, we estimated this factor based on the assumption that the total amount of IkBa at time t=0 was identical in all experimental setups. This value was to 115[AU], where AU stands for arbitrary units (e.g. molecule numbers). The choice 115 is based on the data normalization performed by Lipniacki et.al. They assume that the total amount of NFkB in the system is a constant value (i.e. NFkB is neither produced nor degraded) and that almost all NFkB is bound to IkBa. In addition, there is an excess of free IkBa of up to 15% of the bound IkBa. We assume that the total amount of NFkB is approximately 100[AU], and accordingly the total IkBa is 115[AU].

This approach is not directly applicable to the free NFkB found in the steady state. While some free NFkB is found, the exact quantification of the initial amount is difficult, as it is at least one order of magnitude lower than the maximum activation observed. Instead, we normalized the value at t = 30 min. At 30 minutes, almost the complete IkBa has been degraded. Therefore, the complete 100[AU] NFkB should be observable in our measurement. This allows us again to determine the normalization constant. In the wildtype cell lines, the clear peak in NFkB activity at 30 minutes illustrates that this assumption is plausible. However, in the knockdown cell line siA20I, the time interval 1-7h contains multiple data points with an intensity higher than the one measured at t = 30 min. It is not clear whether these are the result of experimental noise or if NFkB indeed continues to rise beyond the 30 minute mark in knockdown cells. In this thesis, we assume that they are the result of experimental noise; any measured NFkB concentration that after normalization would be larger than 100[AU] is set to a maximum value of 100[AU]. However, we plan to explore this aspect in later experiments.

For A20 we used an approximation similar to the IkBa approximation, i.e. we assumed that the initial concentration in wild type cells at time t = 0 was a fixed value set to 100 [AU]. Estimation of initial concentration of knockdown cells was based on the evaluation of the knockdown efficiency performed by our experimental partners. As they estimated that the knockdown reduced A20 to ~1/6, we set the initial concentration of A20 in knockdown cell lines to 16[AU].

3.4. Error function and noise model

Fitting a mathematical model to a number of experimental observations almost always requires a quantitative measure of the fitting quality, often called an ***error function***. Error functions are usually set up in such a way that they are equal to zero if the model explains the data perfectly, and positive otherwise. Depending on both the type of model, the experimental measurements and the desired application, a large number of error functions exists.

One of most frequently employed error functions is the ***sum of squared residuals*** (SSR). It is based on the method of least squares, proposed 1794 by K. Gauss. It is based on the statistical assumption that the deviation between the real value of the observed system and the measured value is ***normally distributed*** with a mean centered around the real value and a fixed standard deviation ***that does not vary between experiments***. If this is the case, the SSR is calculated as follows:

$$SSR(x^{sim}) = \sum_{i=1}^{n} \frac{(x_i^{sim} - x_i^{real})}{{\sigma_i}^2}$$

Neither x_i^{real} (the real value of the system) nor ${\sigma_i}^2$(the normally distributed variance of the observation process) are intrinsically known to us; we have to estimate them from the experimental measurements. We use the mean of all observations for each time point as the estimate of the real system value.

In contrast, estimating the variance is more complicated. With only two experimental replicates, estimating the variance for each time point is not an option, as the resulting estimate would be very unreliable. In order to compensate for the small number of experimental replicates, we have to calculate the ***pooled variance*** of multiple time points.

In order to calculate the pooled variance of multiple observations, we have to specify whether we assume that the ***noise of our observations is absolute or relative***. If we assume an absolute noise model, the variance at each time point is identical. In contrast, in a relative noise model, the noise of each measurement is scaled with the size of the signal, e.g. a strong signal has a larger absolute variance than a smaller signal, but the same relative variance. For this analysis, we assume a relative variance model. This is mostly due to the A20 measurements; if we calculate the absolute standard deviation of all A20 time points, we find that the result is about 60[AU] (due to some possible outliers in the wildtype data series). This value is larger than most observations in the knockdown cell line, implying that under the assumption of absolute noise, the entire knockdown cell line is "drowned" in noise. The mean of our observed data as well as estimated relative errors are visualized in Figure 7.

3.5. Spline approximation of experimental measurements

During our analysis of the NFkB network model, we will utilize spline interpolations of our experimental observations as (sub-) model inputs to analyze several input/output pairs. These splines are utilized to approximate the value between our experimental observations by differentiable values, guaranteeing that we can use the splines as ODE inputs. The splines are visualized in Figure 7. They were generated using the smooth and spaps functions from the Matlab Curve fitting toolbox to generate the shape of the splines, and the alglib library for c++ to port the splines into an c environment. It should be noted that the parameters of the smoothing functions were chosen manually to generate a spline shape which our experimental partners considered to represent the qualitative processes accurately.

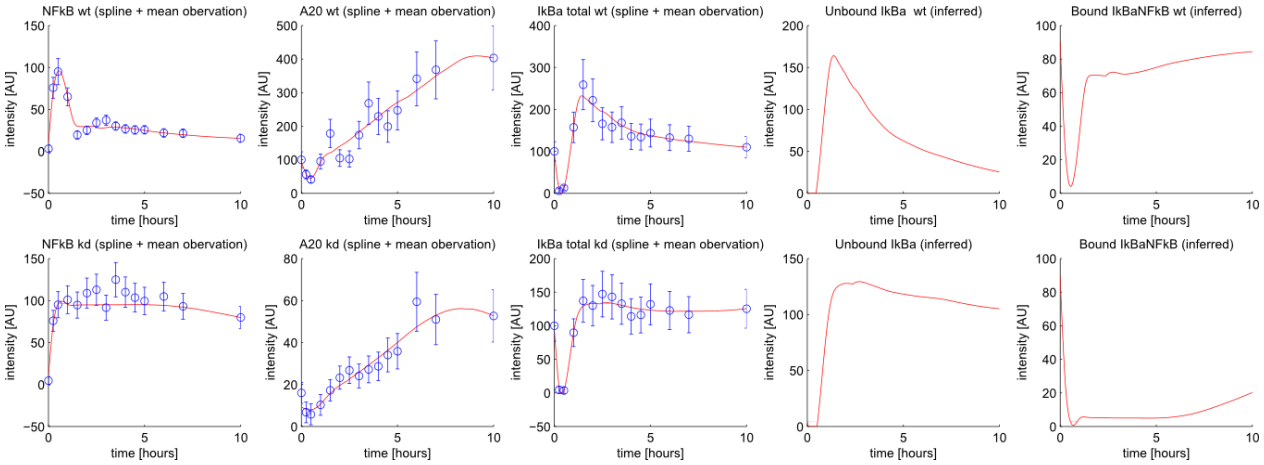


Figure 7: Mean of experimental measurements with error bars and interpolating splines.

4. Methods

4.1. Simulation of ODE models

ODE models x are mathematical systems that are described by their first derivatives with respect to the time dx/dt. Together with an initial state $x_{t=0}$, it is possible to simulate an ODE using numerical approaches. ODEs are used to describe a large variety of different systems, ranging from physics, economics, biological environments to biochemical pathways.

In biochemical systems, x has typically multiple dimensions, e.g. one for each concentration in the model. The initial state $x_{t=0}$ summarizes the concentrations in the system under "resting" conditions, i.e. prior to experimental stimulation. More details regarding the general approach of modeling biochemical systems as ODEs can be found in textbooks. We solved the ODEs of our system using the sundials c++ library, as it is one of the most efficient ODE solvers freely available.

4.2. Bayesian parameter sampling using Metropolis Hastings MCMC sampling

Our main focus in this thesis is not to provide a complete MCMC analysis of the pathway model, but rather to illustrate that the general model complexity of our adapted model is small enough to allow the application of MCMC algorithms. Therefore, we limit us to a short review of the most important properties of MCMC based parameter sampling and again refer to textbooks for more details regarding their utilization.

Metropolis Hastings MCMC starts off with a random initial parameter vector. This parameter vector is slightly varied at each so called ***proposal step***. A proposal for a new parameter can either be accepted or rejected. If the new proposal results in an improvement of the error function, it is automatically accepted. Otherwise, it is accepted with a certain probability that depend on the difference between the error value of the new proposal and the old parameter set. The proposal function has to be chosen according to certain mathematical properties; however some very simple proposals, such as a normally distributed step are often sufficient.

By iterating proposal steps, the Metropolis Hastings algorithm generates a sequence of parameters. Due to the construction of the algorithm, this chain has a higher density of "good" parameter sets. Ideally, most samples derived from an MCMC algorithm will represent good parameter fits with low associated error values.

4.3. Parameter Optimization using Simulated Annealing

Simulated Annealing is a parameter optimization technique similar to MCMC sampling. Again, we start with an initial parameter vector that we vary according to some proposal distribution. The acceptance formula for Simulated Annealing starts out similar to the formula for Metropolis Hastings MCMC, however it contains an additional term that influences the acceptance probability, the temperature. At a high temperature, the probability to accept a small increase in the error function is higher than at a low temperature.

The simulated annealing algorithm is started just like a Metropolis Hastings MCMC run, with proposals generated and accepted or rejected at each step. However, according to some fixed schedule, the temperature of the chain decreases. This eventually results in a situation where all increases in the error function are rejected. This will almost always result in the final accepted parameter set being at least a good local optimum (though not necessarily global)

4.4. Modeling of common biochemical reaction types in ODE's

Several basis types of biochemical reactions are often modeled similar in ODEs. Here, we give a short overview of these standard reactions found in our model and established models of NFkB signaling:

Basal decay of proteins and mRNA

$$R_{protein}decay|base = k_{protein}decay|base * protein$$
$$R_{mRNA}decay|base = k_{mRNA}decay|base * mRNA$$

Dimerization and dissociation of proteins

$$R_{protein}bind = k_{protein}bind * protein_1 * protein_2$$

Transcription and translation of mRNA

Transcription kinetics: If no transcription factor saturation occurs and n_{th} order transcription factor cooperativity is assumed (n_{th} order mass action kinetic):

$$R_{mRNA}transcription|induced = k_{mRNA}transc|ind * transcriptionfactor^n$$

Transcription kinetics: If transcription factor saturation occurs and n_{th} order transcription factor cooperativity is assumed (n_{th} order hill kinetic):

$$R_{mRNA}transcription|induced = k_{mRNA}transc|ind * \left(\frac{transcriptionfactor}{transcriptionfactor + k_{halfmax}}\right)^n$$

Enzymatically induced decay of Proteins

Reaction equation if time determining step is protein/enzyme binding (Mass Action kinetic):

$$R_{protein}decay|induced = k_{protein}decay|induced * enzyme * protein$$

Reaction equation if time determining step is protein modification (Michaelis Menten Kinetic):

$$R_{protein}decay|induced = k_{protein}decay|induced * enzyme * \frac{protein}{protein + K_{MM}}$$

5. Overview of all computational models implemented and utilized

In this chapter, we will give a short overview of all models and submodel we implemented during our design of the NFkB pathway model. More details to the models are given in chapter 6.

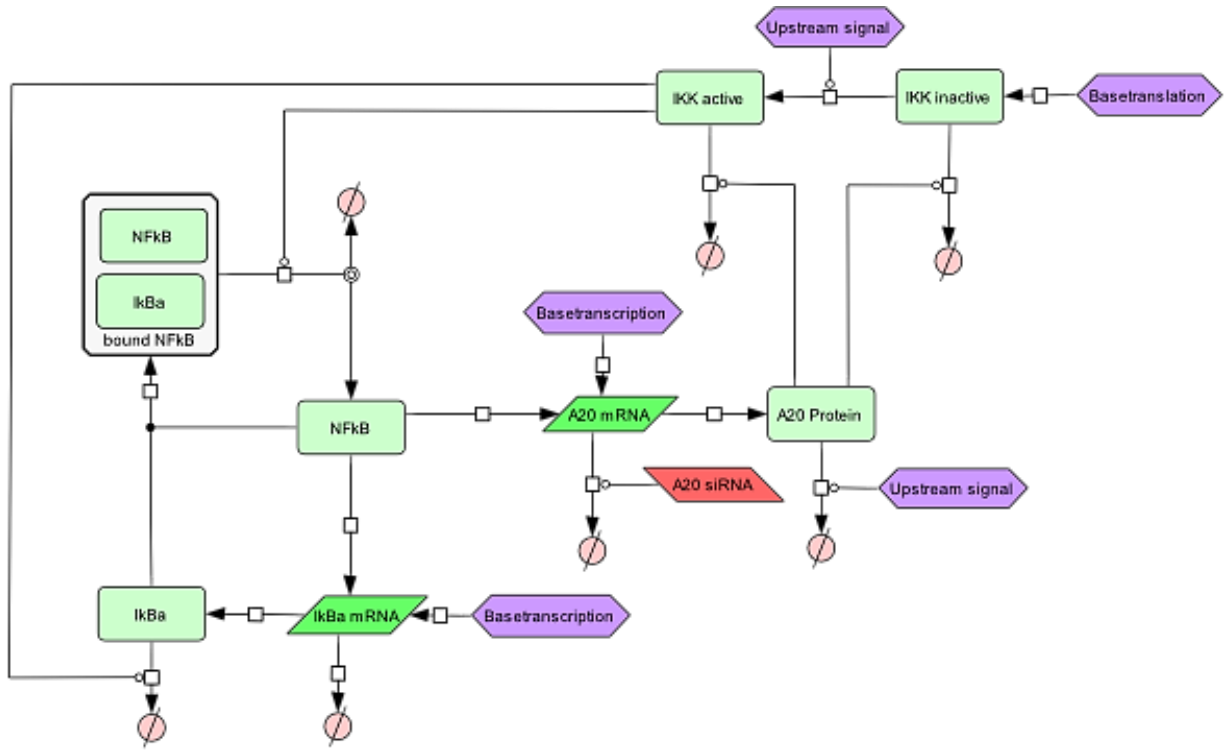


Figure 8: Initial complete model of the canonical NFkB pathway below the level of IKK activation
This model is the initial model. It is similar to the Lipniacki model, but does not include core import/export reactions.

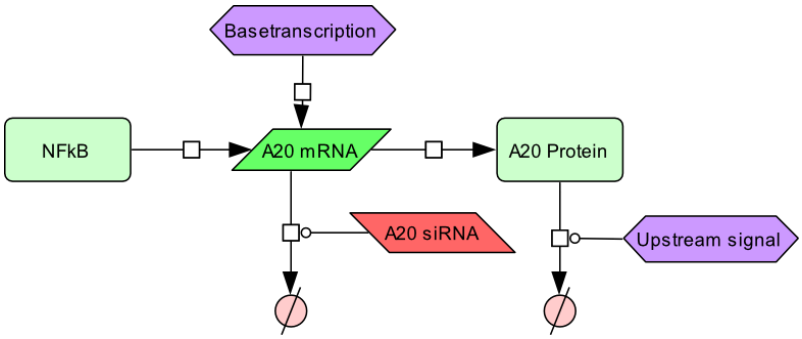


Figure 9: A20 regulation submodel.
This model is the submodel describing the regulation of A20 by NFkB and Upstream effects.

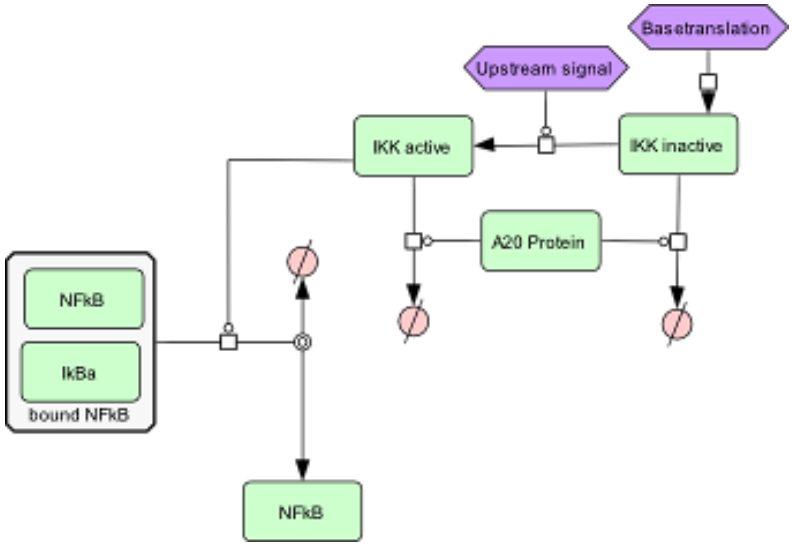


Figure 10: Version 1 of the IKK regulation submodel
This model is the submodel describing the regulation of IKK by A20 and Upstream effects.

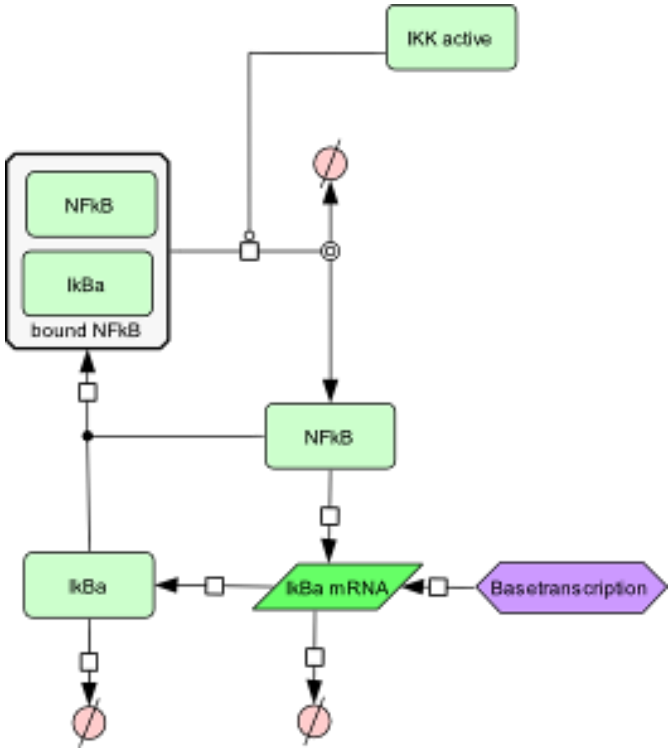


Figure 11: Version 1 of the IkBa regulation submodel.
This model is the submodel describing our initial understanding of the interplay between IkBa and NFkB, regulated by active IKK. It was later replaced by the model in the next figure.

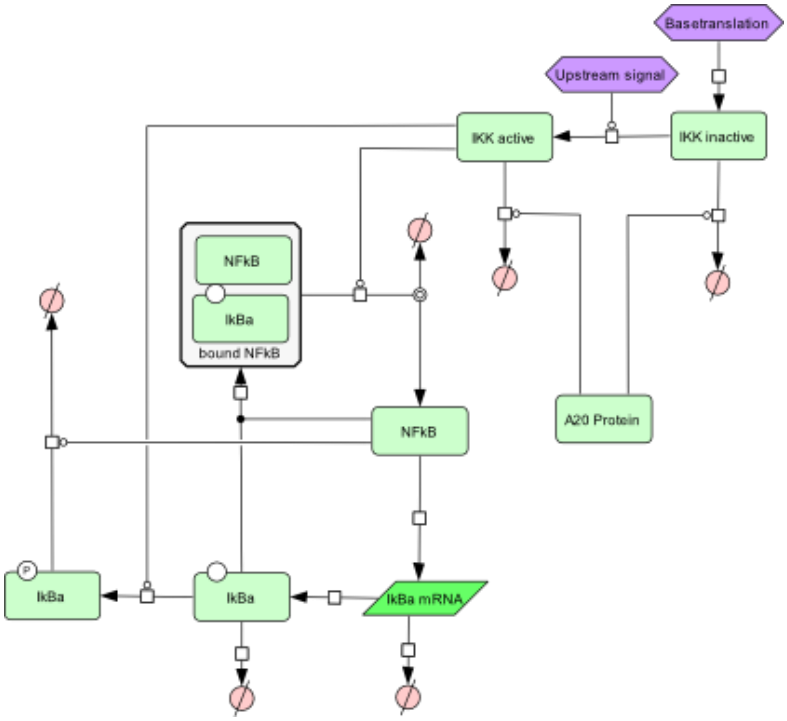


Figure 12: Merged IkBa and IKK submodel Version 2.
Based on insufficient fit quality, we decided to adapt the IkBa submodel. Phosphorylation of unbound IkBa and degradation of unbound, phosphorylated IkBa by NFkB were added as additional reactions.

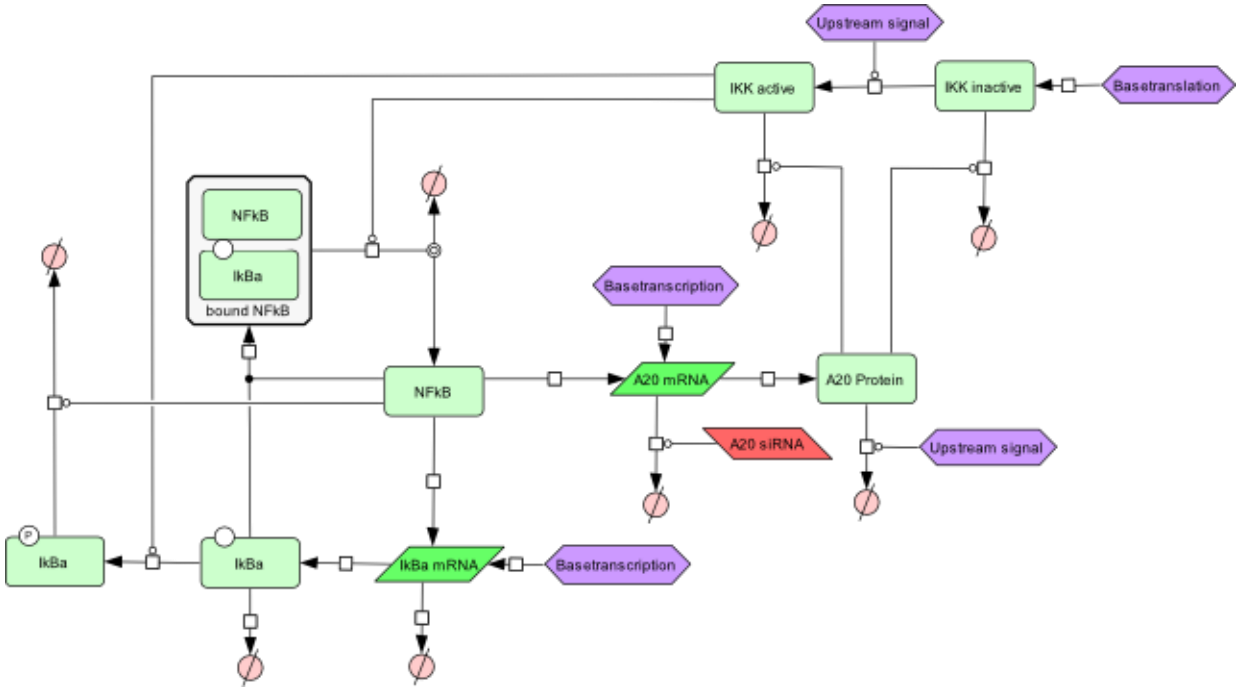


Figure 13: Complete, adapted model of the canonical NFkB pathway

6. Results

6.1. Modeling of NFkB activation in Jurkat T Cells

The utilization of computational models to analyze and predict the behavior of biochemical systems involves many challenges. One of these challenges is the adaption of an existing model to account for different data types or differences in the biological system analyzed. In this chapter, we will deal with this problem.

We have to consider several aspects when adapting the existing models of NFkB activation to our experimental setup. A major biological difference is that the models introduced in section 3.1 have been designed based on experiments performed on fibroblast cells, which are cells of the connective tissue. In contrast we analyze the behavior of Jurkat cells, a cell line of immortalized T cells. While this does not change the way the different molecules interact, we have to consider the possibility that molecule numbers, transcription rates and volumes of cellular compartments differ between the two cell types.

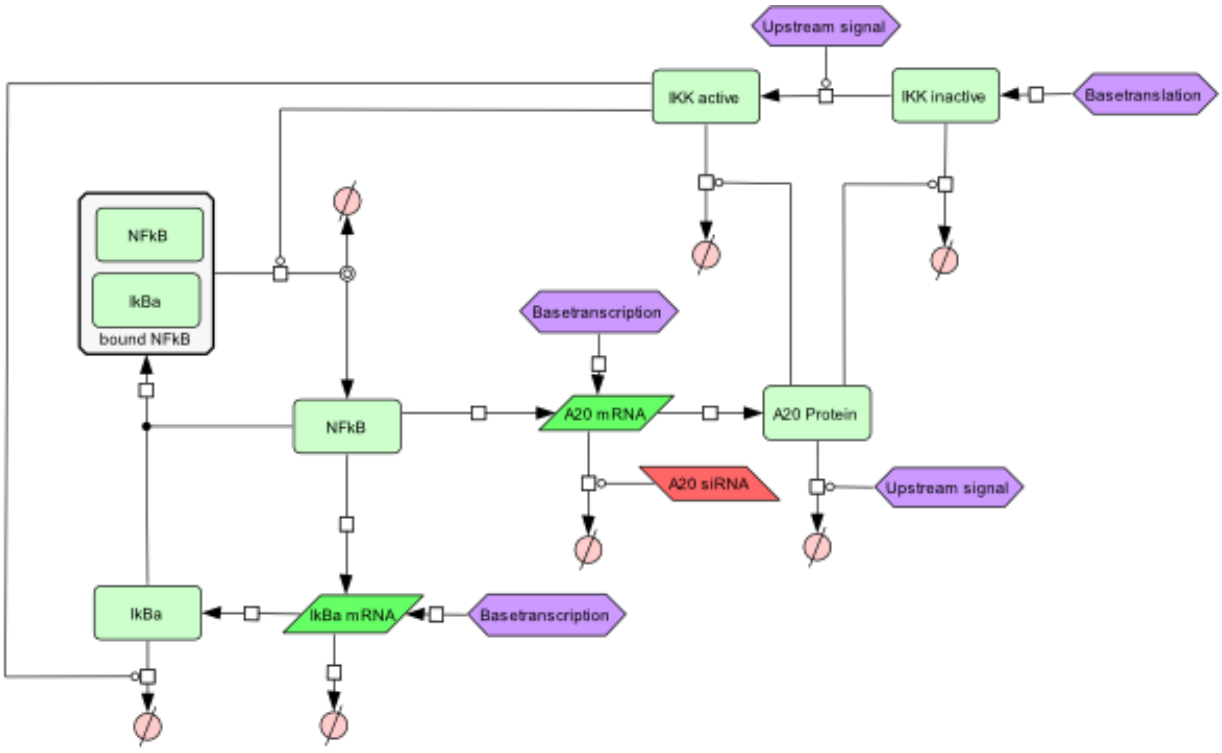


Figure 14: Complete topology of all reactions of the canonical NFkB pathway below the level of IKK activation, excluding core import/export reactions

In addition we utilize different types of experimental measurements. Currently, we quantify only the relative concentration of the proteins measured, not the absolute numbers. We also utilize whole cell extracts, which means that we can't estimate which fraction of the measured proteins is found in the nucleus and which in the cytoplasm. On the other hand, we analyze the effects of stimulation for long periods of time up to ten hours. This allows us to characterize not only the kinetics of the initial response (which is mostly characterized by the spiking IKK activity), but also the resulting long term reaction to continued stimulation.

6.1.1. General model topology

The general model topology of the NFkB pathway we utilize, as illustrated in Figure 14, is mostly based on the established model by Lipniacki discussed in section 0. Primarily, we changed the degradation of IkBa and IkBaNFkB from the explicit set of binding- and unbinding reactions found in most models into a single phosphorylation reaction. We found this to be necessary, as we had no biological prior knowledge regarding the absolute molecule numbers of IKK and IkBa.

We assessed that the advantage of modeling the degradation process in more detail, i.e. more detailed predictions, would have been mostly lost by the added degree of parameter under determination introduced by fitting the relation of the molecule numbers. This is especially true as preliminary fitting results suggested that even biologically invalid molecule ratios could result in acceptable data fits. This problem could have been handled by using a prior for the range of biologically acceptable IKK/IkBa ratios, however based on discussion with our experimental partner we decided that currently available data regarding the canonical pathway in Jurkat cells is insufficient to define such a prior without researcher bias.

Due to similar reasons we initially decided not to model core import/export processes. The modeling of core import/export processes requires seven new parameters: import and export ratios for NFkB, IkBa and bound IkBaNFkB complex, as well as the ratio of nuclear volume compared to cytoplasmic volume. Hoffmann et.al. performed several intricate experiments with mouse fibroblasts to estimate these ratios and deal with the parameter under determination they introduce. However, we could not reproduce these analyses, as we utilized only whole cell extract data and not separate measurements for cytoplasm and nucleoplasm.

We have reasons to believe that simply transferring the established estimations of core import/export dynamics to our system would result in biologically questionable assumptions. For example it is known that the ratio of nucleoplasm and cytoplasm is significantly different between Jurkat and fibroblast cells. Initial modeling attempts suggested that even biologically questionable assumptions (e.g. a very small rate of NFkB import) could explain our experimental observations. To avoid a model which could potentially produce biologically contradictory predictions, we decided to ignore the mechanisms of core import/export. This can be understood as an implicit assumption of a quasi-steadystate, i.e. that all import/export reactions happen on a

timescale so fast compared to the rest of reactions that they can be assumed to be instantaneous. Under quasi steady state condition, all concentrations can be replaced by their equilibrium concentrations, which can be calculated explicitly. For a system of non-saturated core import/export processes such an equilibrium is defined by a fixed ratio of distribution of the total molecules between both compartments. In turn, such a ratio could be integrated into the existing kinetic constants, resulting in an implicit model that recapitulates the events of both compartments in a single compartment. However, as we discuss in section 6.1.4.2, dropping core import/export processes leads to additional problems, which raise the question whether such a quasi-steadystate assumption is justified.

6.1.2. Detailed choice of reaction kinetics

Beyond the large topological decisions to model the activation of NFkB without core import/export dynamics and to replace sequences of enzyme binding, processing and unbinding with one-step reactions, several aspects of the ODE model can be implemented in multiple ways.

The first aspect is the degradation of IKK by A20. The exact mechanism of this process is currently unknown. In the established model by Lipniacki as introduced in section 3.1.2., A20 degrades only the assembled IKK complex, but not the unassembled precursor IKK inactive. This also makes sense from a biological point of view. If A20 degraded the inactive precursor, this would result in a high precursor turnover in unstimulated cells. An A20 knockdown should therefore result in an increase of the inactive precursor that has so far not been noticed.

We decided to implement the assumption that A20 targets only active IKK. Note that this decision is motivated by biological considerations; based on preliminary simulations, degradation of inactive IKK or indiscriminate degradation of active and inactive IKK could also explain the observed experimental measurement.

The second aspect is the degradation of IkBa and IkBaNFkB by IKK. Again, this topic is still subject to biological discussions. While early studies reported that the IKK complex has a higher specificity for the IkBaNFkB complex compared to the IkBa monomer, current studies by [14:Mathes, E. et. al.] dispute this claim. They state that IKK phosphorylates both IkBa and IkBaNFkB indiscriminately, but that phosphorylation results only in increased decay for the IkBaNFkB complex. In contrast the IkBa monomer is not further destabilized upon phosphorylation. We will discuss further implications of the assumed phosphorylation dynamics in section 0.

The final aspect we consider is which type of reaction kinetics should be utilized to model the enzymatic reactions of the pathway (degradation of IkBaNFkB by IKK, degradation of IKK by A20 and degradation of A20 by upstream effects) and the NFkB induced transcription of A20 and IkBa mRNA. Enzymatic reactions can be modeled as either first order mass action kinetics or

Michaelis Menten kinetics. The possibilities for mRNA transcription dynamics are either mass action kinetics and hill kinetics of varying orders.

As already discussed, both Michaels Menten kinetics and sigmoid activation functions have a region of ***approximately linear response***. This allows us to understand mass action kinetic as a specialized case of the other type of dynamics. Michaelis Menten kinetics respectively sigmoid activation functions are generalizations. The more complex generalizations should only be favored above mass action kinetics if a simple linear response is not sufficient to explain the experimental observations; otherwise the utilization of more complex dynamics results in the introduction of unnecessary model under determination and less certain predictions. We discuss these decisions in detail in section 6.1.4.1 to 6.1.5, as they can have significant impact on the model based predictions. We believe it is important to justify the choice of kinetics in detail, especially as some of our findings run contrary to the practices utilized in the established models.

6.1.3. Analysis of submodels based on conditional independence

Analyzing a model with a large number of parameters is often problematic, as the performance of parameter estimation and optimization techniques decays rapidly with an increasing number of parameters. This is an aspect of the "curse of dimensionality", and has been discussed as such in numerous publications. This problem is often tackled by splitting models into independent submodels if the model topology allows this and analyzing these submodels separately. This approach, while a potent method to handle problems of high dimensions, is often limited in applicability. If a model includes feedback mechanics, none of the states in a feedback loop can be assumed to be independent from each other. This is the case in the model topology illustrated in Figure 14, with the exception of inactive IKK, all states are subject to feedback by NFkB.

However, we can adapt the concept of conditional independence to deal with the existing feedback loops in our model by finding submodel which are ***conditionally independent*** and analyzing them separately. For example, if we change the behavior of IkBa, this will influence the behavior of A20 only via the intermediate node NFkB. If we find that a change to the concentration of IkBa does not influence the value of NFkB, then for this change, A20 and IkBa are conditionally independent. Of course, this approach is limited by the states that we either directly observe or at least can infer up to a constant. States that can neither be observed nor inferred are unsuitable as cutting points to separate submodels.

We analyze these submodels by replacing the input states of the submodels with splines of our experimental observations and analyzing the input/output behavior for different reaction kinetics and parameterizations. This approach allows us to reject some invalid reaction kinetics without the need for parameter optimization runs of the entire model. Overall, we analyze the following submodels:

A20 activity submodel

This submodel describes the regulation of A20 by NFkB, A20 siRNA and upstream signals. It includes the transcription (both based and induced) and degradation of A20 mRNA and the translation and degradation of A20 protein. The upstream signal and the degradation of A20 siRNA are on/off "switch" variables, which have a fixed value based on experimental conditions. The NFkB input is realized using a spline of our experimental data.

IkBa activity submodel.

The IkBa activity submodel models the synthesis, decay and binding processes of the free IkBa monomer. Note that the amount of bound IkBaNFkB is not part of the model; rather, the inferred value of the IkBa NFkB binding reaction is treated as an model out. The only variable input of the model is the NFkB activity profile.

IKK activity submodel

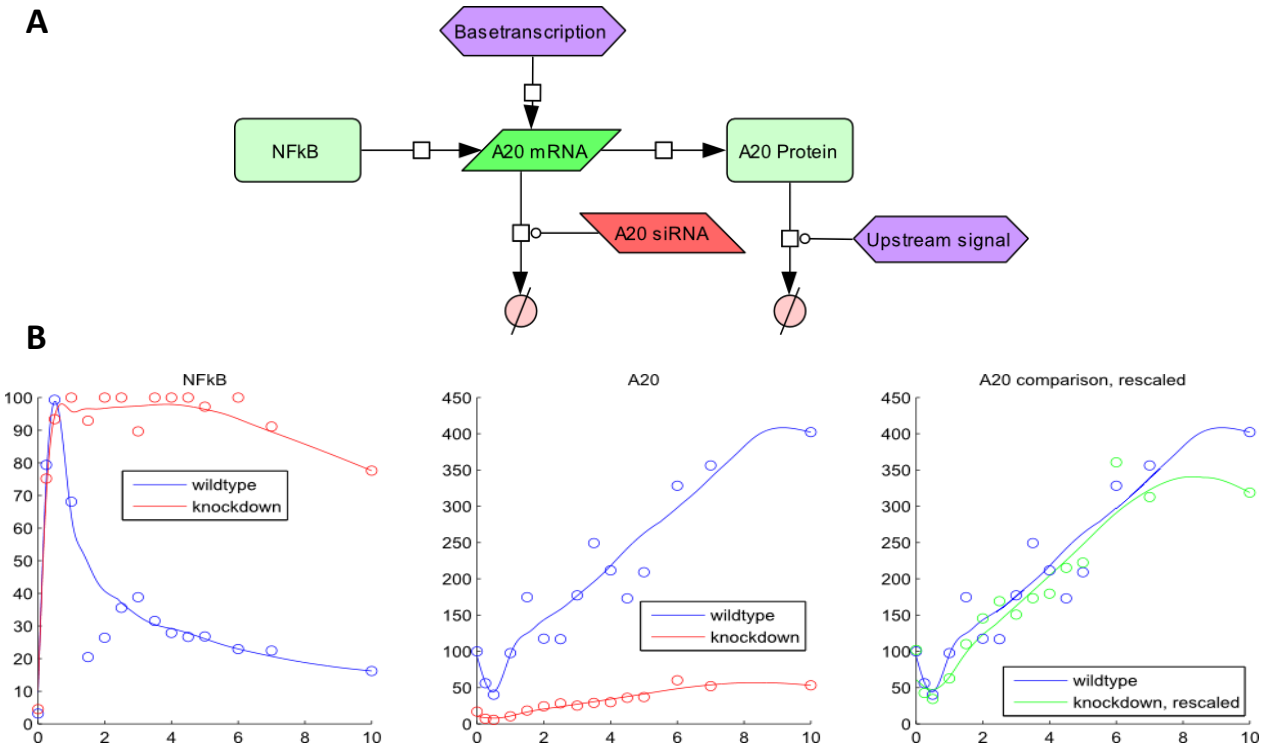
The IKK activity submodel models the regulation of IKK by A20 and upstream signals. A20 is the only variable input of the submodel and again realized using a spline of our experimental measurements. The output of the model is an IKK activity profile similar to the model inputs utilized by Hoffmann et.al. It is compared to the inferred IkBa NFkB binding profile analyzed in the IkBa activity submodel.

6.1.4. Analysis of the A20 activation submodel

In this section, we will focus on a theoretical analysis of qualitative properties of the A20 submodel, the observed data and the resulting implications for choosing suitable reaction kinetics. We avoid numerical analysis as well as possible, to confirm that our decisions are indeed based on qualitative properties of the data and not result of numerical optimization.

Our major result is that the induction of A20 mRNA should preferably be modeled using hill kinetics, as we find noticeable saturation effects when comparing the reaction magnitude of A20 decay, degradation and synthesis. We show in detail that linear or polynomic transcription dynamics are not an option and we are justified to model the transcription of NFkB target genes using sigmoid saturation based kinetics.

The A20 submodel has two internal states which change over time according to the assumed reaction kinetics, A20 protein and A20 mRNA. The inputs of the model are an spline estimate of NFkB and the two experimental "switches" A20 degradation and presence of A20 siRNA.



Change of A20 protein and A20 mRNA:

$$\frac{d}{dt}(A20) = +R_{A20}synth - R_{A20}decay|base - R_{A20}decay|induced$$

$$\frac{d}{dt}(A20_{mRNA}) = +R_{A20mRNA}transc|base + R_{A20mRNA}transc|induced - R_{A20mRNA}decay|base - R_{A20mRNA}decay|induced$$

Established reaction kinetics:

$R_{A20mRNA}transc\|base = k_{A20mRNA}transc\|base$	Basal transcription is a assumed to be a fix value.
$R_{A20mRNA}decay\|base = k_{A20mRNA}decay\|base * A20\ mRNA$	Basal decay of mRNA is based on first order MA kinetics
$R_{A20mRNA}decay\|induced = k_{A20mRNA}decay\|base * \mathrm{A20_{mRNA}} * \mathrm{A20_{siRNA}}$	A20 mRNA decay induced by siRNA is modeled as a linear increase in A20 mRNA basaldecay
$R_{A20}synth = k_{A20}synth * A20mRNA$	Synthesis/Translation of new protein is based on first order MA kinetics
$R_{A20}decay\|base = k_{A20}decay\|base * A20$	Basal decay of proteinss is based on first order MA kinetic

Variable reaction kinetics:

$R_{A20}decay\|induced$	Induced A20 degradation can be either mass action or Michealis Menten kinetic
$R_{A20mRNA}transc\|induced$	Induced A20 mRNA transcription can be either n_{th} order hill- or mass action kinetic

Figure 15: A20 submodel visualization, A20 submodel time course data with splines and overview of reaction equations and kinetics

The model contains seven reactions:

- Basal- and NFkB induced synthesis of A20 mRNA
- Basal- and induced decay of A20 mRNA
- Basal- and induced decay A20 protein
- Translation of A20 mRNA to synthesize new A20 protein

Of these reactions, the induced transcription of A20 mRNA and the induced decay of A20 protein can be modeled in multiple ways. An overview of the submodel, the reaction equations and the kinetics is provided in Figure 15C.

We will use several algebraic expressions repeatedly in our analysis. These are summarized in the following section:

Common formulas and estimations in the analysis of the A20 submodel

The value of A20 at time t expressed as the integral over the first derivative + initial value:

$$(Eq: 5.2.2.1.1)\ A20_t = A20_0 + \int_{x=0}^{t} [R_{A20}synth - R_{A20}decay|base - R_{A20}decay|induced]$$

The time course data for the experimental data we utilize is visualized in Figure 15B. If we compare the value of the A20 splines, we find that the A20 knockdown time course is almost exactly 1/6 of the value of the A20 wildtype time course (third graph). We will use this observation multiple times in the following chapter by approximating:

$$(Eq: 5.2.2.1.2)\ A20^{wt} = 6 * A20^{kd}$$

As the basal decay of A20 is directly proportional to the value of A20, we can also approximate:

$$(Eq: 5.2.2.1.3)\ R_{A20}^{wt}decay|base = k_{A20}decay|base * A20^{wt}$$
$$= 6 * k_{A20}decay|base * A20^{kd} = 6 * R_{A20}^{kd}decay|base$$

In addition, we note that from approximately one hour onward, both A20 time courses are monotonously increasing. This means that we have net synthesis, allowing us to approximate:

$$(Eq: 5.2.2.1.4)\ for\ t \geq 1h:$$
$$R_{A20}synth(t) \geq R_{A20}decay|induced(t) > R_{A20}decay|induced(t) + R_{A20}decay|base(t)$$

6.1.4.1. Analysis of A20 steady state conditions

An essential step in modeling a biochemical system is to analyze the initial conditions that exist prior to experimental stimulation. For this system, we assume that at t=0min we start from a steady state behavior both in A20 wildtype and knockdown. We already mentioned that the steady state concentration of A20 protein in knockdown cells is approximately 1/6 of the concentration in wildtype cells $(Eq: 5.2.2.1.2)$. As biologically, the decay of A20 is not directly influenced by the siRNA knockdown, in order for steady state conditions to hold the A20 protein synthesis under steady state conditions is 1/6 of the protein synthesis in wild type cells.

$$(Eq: 5.2.2.2.1.1) \quad R_{A20}^{kd} synth_{SS} \approx \frac{1}{6} R_{A20}^{wt} synth_{ss}$$

As the synthesis rate is proportional to the amount of A20 mRNA, we conclude that the effect of the A20 knockdown is a reduction of the A20 mRNA to 1/6. Again, as the transcription of mRNA is assumed to be unaffected by A20 siRNA due to the underlying biological processes. This implies that the reduction of the A20 mRNA steady state level in the knockdown cell line is due to an increased basal decay of A20 mRNA. We can estimate that the siRNA induced decay of A20 mRNA under steady state conditions is approximately five times the basal mRNA decay (normalized by the value of the A20siRNA switch), as this results exactly in a reduction of both mRNA and protein steady state concentrations to 1/6.

Initial considerations

Based on the basic definitions and observation we find that, under the assumption of first order decay and degradation:

$$\begin{aligned} R_{A20}^{wt} decay|base &= k_{A20} decay|base * A20^{wt} \\ &= k_{A20} decay|base * 6 * A20^{kd} = 6 * R_{A20}^{kd} decay|base \end{aligned}$$

$$\begin{aligned} R_{A20}^{wt} decay|base &= k_{A20} decay|induced * A20^{wt} \\ &= k_{A20} decay|induced * 6 * A20^{kd} = 6 * R_{A20}^{kd} decay|induced \end{aligned}$$

hold naturally. This implies that equation $(Eq: 5.2.2.1.2)$ will always be satisfied under the assumption of mass action kinetics, if:

$$R_{A20}^{wt} synth = 6 * R_{A20}^{kd} synth$$

As the degradation of A20 mRNA is increased by a factor of 6 in knockdown cells, this suggests that the transcription of A20 mRNA is the same in both knockdown and wildtype cells. This would explain the above observations by scaling A20 mRNA and A20 protein level exactly down to 1/6 as we observe.

However, the assumption that the value of A20 mRNA synthesis is the same in knockdown and wildtype cells has to be justified, as the NFkB level in wildtype cells is between 40% and 20% of the activation of NFkB in knockdown cells. This implies that A20 mRNA synthesis is saturated at very low levels of NFkB activity. Such an observation runs contrary to established models, which use kinetics that allow for no saturation effects. However, we will show that, based on the behavior of the stimulated model, such an assumption is justified.

6.1.4.2. Analysis of the A20 time course under stimulation

Transcriptional delay in the induction of A20 synthesis

The first problem when analyzing the dynamic behavior of the A20 submodel rather than only steady state condition arises from dropping core import/export dynamics. This results in a significantly reduced ability to model delay in responses. This is most obviously observable in the behavior of A20 and NFkB during the first hour as visualized in Figure 15B. A20 is rapidly degraded during the first thirty minutes of the experiment and subsequently recovers to the original concentration at about one hour. This new synthesis is induced by NFkB which is activated even more rapidly. During the first fifteen minutes, more than 80% of the maximum NFkB activity is achieved. NFkB activity peaks at 30 minutes with maximum activity, and is subsequently re-inhibited by newly synthesized IkBa. This illustrates the noticeable lag between activation of NFkB and NFkB induced synthesis; at the time when NFkB activity is at a maximum, A20 just starts to recover from the initial degradation.

Qualitatively, this could be explained by a low mRNA synthesis rate, which could contribute to delaying the peak of A20 synthesis; however the numeric analysis visualized in Figure 16 shows that such a low mRNA synthesis rate does allow the rapid recovery observed, but would rather result in a far slower recovery. This illustrates the need include some kind of transcriptional delay between NFkB activity and resulting synthesis. We treated the actual value of the transcriptional delay as another model parameter with a range from 15-30 minutes, based on the observed lag between peak NFkB activity and maximum A20 recovery.

Analysis of A20 dynamics in initial transcription lag phase

In the time frame 0-0.5 hours the amount of A20 decays monotonously both in wildtype and knockdown cells. After this the observed amount of A20 starts to increase again due to newly synthesized A20 proteins. The relative magnitude of the decay is similar for both wildtype and knockdown cells. At thirty minutes, the last time point in the lag phase, 60% of A20 have been degraded in wildtype cells and 66% in knockdown cells. If we approximate to be the relative degradation to be 60% in both cell lines, based on equations $(Eq: 5.2.2.1.1)$ and $(Eq: 5.2.2.1.2)$ we find:

$$\frac{1}{6} * (\mathbf{A20_0^{wt}} + \int_{x=0}^{t} [\mathbf{R_{A20}^{wt}synth} - \mathbf{R_{A20}^{wt}decay|base} - \mathbf{R_{A20}^{wt}decay|induced}]) = \frac{1}{6}\mathbf{A20^{wt}}$$
$$= \mathbf{A20^{kd}}$$
$$= \mathbf{A20_0^{kd}} + \int_{x=0}^{t} [\mathbf{R_{A20}^{kd}synth} - \mathbf{R_{A20}^{kd}decay|base} - \mathbf{R_{A20}^{kd}decay|induced}]$$

Utilizing the steady state conditions $(Eq: 5.2.2.2.1.1)$ $R_{A20}^{kd} synth_{SS} \approx \frac{1}{6} R_{A20}^{wt} synth_{ss}$ and the approximation $(Eq: 5.2.2.1.3)$ $R_{A20}^{wt} decay|base = 6 * R_{A20}^{kd} decay|base$, we find::

$$\frac{1}{6} * \int_{x=0}^{t} R_{A20}^{wt} decay|induced = \int_{x=0}^{t} R_{A20}^{kd} decay|induced$$

and by differentiation

$$(\text{Equ: } 5.2.2.1.2.1) \quad \frac{1}{6} * R_{A20}^{wt} decay|induced \approx R_{A20}^{kd} decay|induced$$

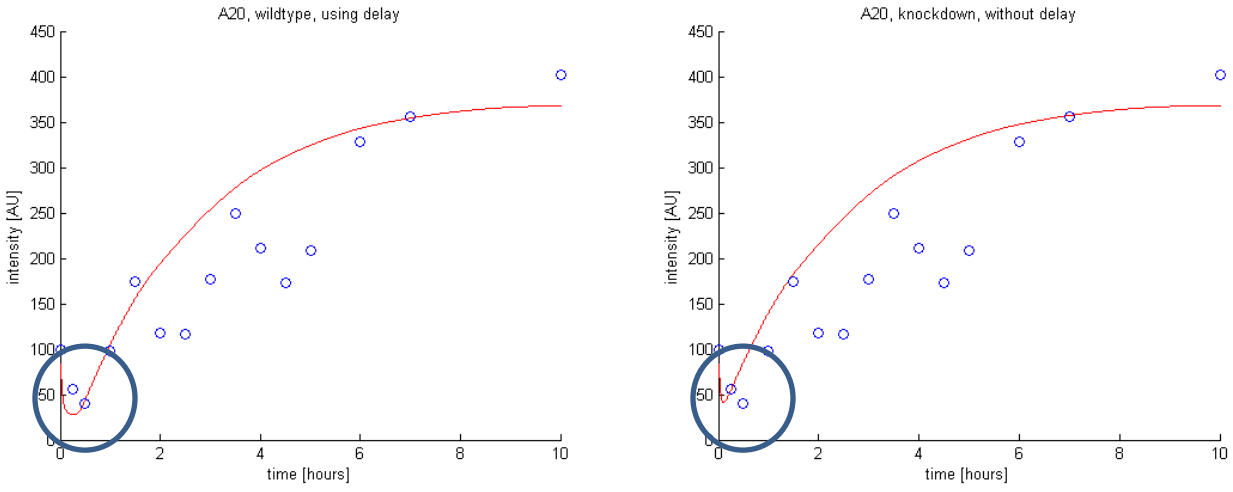


Figure 16: Analysis of delay in NFkB induced A20 synthesis

Simulated annealing fit of the A20 submodel with (left) and without (right) transcriptional delay. The model without transcriptional delay cannot explain the observations at t = 30 minutes (highlighted with a circle) without severe decrease in fitting quality. In contrast, the model with transcriptional delay is shows a dynamic that allows the recovery of A20 to start at ~ 30 minutes.

As discussed, there are two possible kinetics we consider for the degradation of A20 caused by the upstream signal. We can now check both for their compatibility with (Equ: 5.2.2.1.2.1).

(A) $R_{A20}decay|induced = k_{A20}decay|induced * A20$ First order Mass Action kinetic

(B) $R_{A20}decay|induced = k_{A20}decay|induced * \frac{A20}{A20+K_{MM}}$ Michaelis Menten kinetic

As we see, (**Equ**: **5**. **2**. **2**. **1**. **2**. **1**) holds for mass action kinetics for all parameters:

$$\frac{1}{6} * \mathbf{A20_{wt}} * \mathbf{k_{A20}decay|induced} \approx \mathbf{A20_{kd}} * \mathbf{k_{A20}decay|induced}$$

In contrast for Michaelis Menten kinetics we find that:

$$\frac{1}{6} * \mathbf{k_{A20}decay|induced} * \frac{\mathbf{A20_{wt}}}{\mathbf{A20_{wt}} + \mathbf{K_{MM}}} = \mathbf{k_{A20}decay|induced} * \frac{\mathbf{A20_{kd}}}{\mathbf{A20_{wt}} + \mathbf{K_{MM}}}$$
$$\approx \mathbf{k_{A20}decay|induced} * \frac{\mathbf{A20_{kd}}}{\mathbf{A20_{kd}} + \mathbf{K_{MM}}}$$

only holds for $K_{MM} \gg A20_{wt}$. during most of the lag phase, implying $K_{MM} \gg 50$. For such large K_{MM} values, the entire reaction dynamics of the lag phase happen in the linear region of the Michaelis Menten reaction.

Analysis of first order mass action kinetics

Note that we have not yet shown that mass action kinetics can explain the observations during the lag phase, but only illustrated that if wild type decay can be explained using mass action kinetics, so can the decay in knockdown cells. To show that Mass Action kinetics are indeed sufficient, we consider again (***Eq***: **5**. **2**. **2**. **1**. **1**):

$$\boldsymbol{A20_t = A20_0 + \int_{x=0}^{t} R_{A20}synth - A20 * [k_{A20}decay|base + k_{A20}decay|induced]}$$

Note that $R_{A20}^{wt}synth$ is a constant value during the lag phase. Therefore, the above equation is of the type:

$$x(t) = c_1 + \int_0^t c_2 + x * c_3$$

This is equation is the antiderivative of the differential equation $\frac{dx}{dt} = c_2 + x * c_3$. This differential equation can be explicitly solved: $x = \left(\frac{c_2}{c_3}\right) + K * \exp(c_3 * t)$. K is determined by the initial condition: $K = x_0 - \left(\frac{c_2}{c_3}\right)$. Therefore we can solve:

$$\mathbf{A20_t} = \frac{\mathbf{R_{A20}synth}}{\mathbf{k_{A20}decay|base} + \mathbf{k_{A20}decay|ind}} + \left(\mathbf{A20_0} - \frac{\mathbf{R_{A20}synth}}{\mathbf{k_{A20}decay|base} + \mathbf{k_{A20}decay|ind}}\right) * \mathbf{exp}\left(\mathbf{t} * (\mathbf{k_{A20}decay|base} + \mathbf{k_{A20}decay|ind})\right)$$

We replace $(k_{A20}decay|base + k_{A20}decay|induced) = h_1$ for better readability.

$$\boldsymbol{A20_t} = \frac{\boldsymbol{R_{A20}synth}}{\boldsymbol{h_1}} + \left(\boldsymbol{A20_0} - \left(\frac{\boldsymbol{R_{A20}synth}}{\boldsymbol{h_1}}\right)\right) * \boldsymbol{exp}(\boldsymbol{t} * \boldsymbol{h_1})$$

The existence of a parameter set that explains our observations during the lag phase is now given due to the form of the explicit solution. An equation of the type $x = \left(\frac{c_2}{c_3}\right) + K * \exp(c_3 * t)$ can be fitted through any two data points without constraints. Based on this analysis, we find that first order mass action kinetics are sufficient to explain the initial decay phase. Michaelis Menten kinetics with a large K_{MM}value could also be utilized, should mass action kinetic proof insufficient to explain observations outside of the initial lag phase.

Analysis of A20 dynamics after initial lag phase

After the induced transcription and subsequent translation starts to kick in at approximately 30 minutes, A20 recovers rapidly both in wildtype and knockdown cells due to NFkB induced transcription and translation. Due to approximation $(Eq: 5.2.2.1.2)\ A20^{wt} = 6 * A20^{kd}$, we find that:

$$(\text{Equ: } 5.2.2.1.4.1) \quad \int \left[R_{A20}^{kd} synth - R_{A20}^{kd}\, decay|base - R_{A20}^{kd}\, decay|induced\right] \approx \frac{1}{6}\int \left[R_{A20}^{wt} synth - R_{A20}^{wt}\, decay|base - R_{A20}^{wt}\, decay|induced\right]$$

Based on approximation $(Eq: 5.2.2.1.3)\ R_{A20}^{wt} decay|base = 6 * R_{A20}^{kd} decay|base$, in order for (Equ: 5.2.2.1.4.1) to hold, we require

$$\int \left[R_{A20}^{kd} synth - R_{A20}^{kd}\, decay|induced\right] \approx \int \left[\frac{1}{6} R_{A20}^{wt} synth - \frac{1}{6} R_{A20}^{wt}\, decay|induced\right]$$

by differentiation

(Equ: 5.2.2.1.4.2)

$$\left[R_{A20}^{kd} synth - R_{A20}^{kd}\, decay|induced\right] \approx \left[\frac{1}{6} R_{A20}^{wt} synth - \frac{1}{6} R_{A20}^{wt}\, decay|induced\right]$$

Both in wildtype and knockdown cell lines, NFkB is activated to maximum activity during the first thirty minutes. After this initial peak, the NFkB activity drops strongly in wildtype cells, but stays approximately the same in knockdown cells. This suggests that the effective transcription rate in knockdown can be assumed to be constant after the initial lag phase, implying that the resulting synthesis rate $R_{A20}^{kd} synth$ is also constant. In contrast, NFkB activity in wildtype cells ranges from between approximately 40% of maximum activity at three hours to 20% at 10 hours. Again, we assume $(Eq: 5.2.2.1.2)\ A20^{wt} = 6 * A20^{kd}$.

We will now analyze the possible A20 transcription dynamics:

(A) n_{th} order Mass Action

$$R_{A20mRNA} transc|induced = k_{A20mRNA} transc|induced * NFkB^n$$

(B) n_{th} order Sigmoid function

$$R_{A20mRNA} transc|induced = k_{A20mRNA} transc|induced * \left(\frac{NFkB}{NFkB + k_{hill}}\right)^n$$

Case differentiation 1: Mass action transcription function

Based on the observation that at 3 hours the NFkB activity is 40% of the constant activity NFkB activity in knockdown cells, the assumption of mass action kinetics implies that accounting for transcriptional delay, from 3.5 hours onward:

$$\begin{aligned} & R_{A20}^{wt} transc|induced \\ & = k_{A20mRNA} transc|induced * NFkB_{wt}^n \\ & \leq k_{A20mRNA} transc|induced * (0.4 * NFkB_{kd})^n \\ & = 0.4^n * k_{A20mRNA} transc|induced * NFkB_{wt}^n \\ & = 0.4^n * R_{A20}^{kd} transc|induced \end{aligned}$$

In order to analyze the protein synthesis and degradation dynamics, we assume that $R_{A20}^{kd}synth \sim R_{A20mRNA}^{kd}transc$ and $R_{A20}^{wt}synth \leq \frac{0.4^n}{6} R_{A20}^{kd}synth$ holds for the late time points of 5hours and later, the implication being that the initial peak in mRNA synthesis caused by the NFkB peak during the initial stimulation does not last through the entire experimental period. This is motivated by the observation that mRNA half life time is relatively short, which is biologically motivated. Using equation(Equ: 5.2.2.1.4.1):

$$\left[R_{A20}^{kd}synth - R_{A20}^{kd}\, decay|induced\right] \approx \left[\frac{1}{6} R_{A20}^{wt}synth - \frac{1}{6} R_{A20}^{wt}\, decay|induced\right]$$

We find

$$\left[R_{A20}^{kd}synth - R_{A20}^{kd}\, decay|induced\right] \leq \left[\frac{0.4^n}{6} R_{A20}^{kd}synth - \frac{1}{6} R_{A20}^{wt}\, decay|induced\right]$$

Which can be restated as:

$$\left[\left(1 - \frac{0.4^n}{6}\right) R_{A20}^{kd}synth - R_{A20}^{kd}\, decay|induced\right] \leq \left[-\frac{1}{6} R_{A20}^{wt}\, decay|induced\right]$$

For further analysis, we have to make a case differentiation for the different types of induced A20 decay:

Case differentiation 1.1: Mass action transcription function , Mass action induced decay

$$\left[\left(1 - \frac{0.4^n}{6}\right) R_{A20}^{kd}synth - k_{A20}^{kd}\, decay|induced * A20_{kd}\right] \leq \left[-\frac{1}{6} k_{A20} decay|induced * A20_{wt}\right]$$

which implies

$$\left(1 - \frac{0.4^n}{6}\right) R_{A20}^{kd}synth \leq 0$$

Which is obviously false for n > 0 (note that n = 0 is technically a valid solution, however a 0_{th} order mass action function, i.e. a constant function makes no sense as a an induced synthesis rate).

Case differentiation 1.2: Mass action transcription function, Michaelis Menten based decay

$$\left[\left(\frac{0.4^n}{6} - 1\right) * R_{A20}^{kd}synth + k_{A20} decay|induced * \frac{A20_{kd}}{A20_{kd} + K_{MM}}\right] \geq \left[k_{A20} decay|induced * \frac{A20_{kd}}{A20_{wt} + K_{MM}}\right]$$

Based on approximation $(Eq: 5.2.2.1.4)\ t \geq 1h: R_{A20}synth(t) \geq R_{A20}decay|induced(t)$

$$\left[\left(\frac{0.4^n}{6}\right) * R_{A20}^{kd}synth\right] \geq \left[k_{A20}decay|induced * \frac{A20_{kd}}{A20_{wt} + K_{MM}}\right]$$

Which can be restated:

$$\left[\left(\frac{0.4^n}{6}\right) * \frac{A20_{wt} + K_{MM}}{A20_{kd} + K_{MM}}\right] \geq \frac{\left[k_{A20}decay|induced * \frac{A20_{kd}}{A20_{kd} + K_{MM}}\right]}{R_{A20}^{kd}synth}$$

Based on approximation $(Eq: 5.2.2.1.4)\ t \geq 1h: R_{A20}synth(t) \geq R_{A20}decay|induced(t)$

$$\left[\left(\frac{0.4^n}{6}\right) * \frac{A20_{wt} + K_{MM}}{A20_{kd} + K_{MM}}\right] \geq \frac{\left[k_{A20}decay|induced * \frac{A20_{kd}}{A20_{kd} + K_{MM}}\right]}{\left[k_{A20}decay|induced * \frac{A20_{kd}}{A20_{kd} + K_{MM}}\right]}$$

Finally we find:

$$\left[\left(\frac{0.4^n}{6}\right) * \frac{6 * A20_{kd} + K_{MM}}{A20_{kd} + K_{MM}}\right] \geq 1$$

Which is again only true for n = 0, and therefore not of biological relevance.

Case differentiation 2: Sigmoid transcription function

For sigmoid activation functions, the effective rate of transcription in wild type cells can range from a rate of 1.0 * knockdown transcription rate (under the assumption that transcription factor saturation occurs in wildtype cells during the entire time interval, $k_{hill} \ll 20[AU]$) to 0.2^k (if the activation pattern in wildtype cells is in a linear region of the sigmoid function, implying $k_{hill} >> 20[AU]$). If we assume that transcription of A20 is saturated, we find that $R_{A20}^{kd}synth \approx \frac{1}{6}R_{A20}^{wt}synth$ holds for all parameter values. Therefore (Equ: 5.2.2.1.4.1) becomes:

$$\int R_{A20}^{kd}\ decay|induced \approx \int \frac{1}{6}R_{A20}^{wt}\ decay|induced$$

We conclude that, if we assume that A20 transcription is maximally activated at very low NFkB levels, the kinetics of A20 degradation require and support no saturation effects. For higher k_{hill} values, implying little saturation effects we find that

$$R_{A20}^{kd}synth \gg \frac{1}{6}R_{A20}^{wt}synth$$

and in turn

$$\int R_{A20}^{kd}\ decay|induced \gg \int \frac{1}{6}R_{A20}^{wt}\ decay|induced$$

has to hold. This implies that if we model the transcription of A20 using a reaction kinetic and parameterization without clear saturation effects, modeling A20 degradation as a linear reaction is not an option and the alternative of Michaels Menten kinetic based decay has to be utilized to compensate the reduced A20 synthesis in wildtype cells. The smaller the assumed saturation effects in mRNA induction, the stronger non-linear the Michaelis Menten equation hast to behave, implying small K_{MM}values. Note however, that the K_{MM} value of the Michaelis Menten kinetic is limited to relatively large values by the observed lag phase dynamics. We conclude that modeling the induced mRNA transcription indeed required a sigmoid activation function with strong saturation effects at low NFkB levels.

6.1.4.3. Summary of transcriptional and kinetic dynamics of A20 activation

Based on the previous section, we summarize our analysis regarding the modeling of A20 activation:

- A20 siRNA increases the decay of A20 mRNA by the factor six in a linear fashion.
- In knockdown cell lines, the steady state values of A20 mRNA and A20 protein are reduced to 1/6.
- Transcription of A20 mRNA induced by NFkB is modeled with a transcriptional delay between 15min and 30min. This is implemented using a delay chain to avoid technical issues of DDE simulation.
- Induced transcription of A20 mRNA approximates a sigmoid activation function. This sigmoid function has a low k_{hill} value, implying maximum A20 transcription at NFkB levels of 20[AU].
- Modeling of induced transcription with mass action kinetics is not possible.
- Degradation of A20 protein due to Upstream effects can be modeled using mass action kinetics without introducing qualitative problems.

The major difference to previous publications is that we illustrated that it is not sufficient to model induced mRNA transcription using simple mass action kinetics, but that it is required to employ a more complex model. This observation is critical, as it is very intuitive that the assumption of saturation based transcription kinetics can have massive influence on model based predictions compared to the assumption of a linear response without saturation.

Numerical simulation of the A20 submodel

Note that until now we have mainly focused on the analysis of qualitative properties, such as the existence of saturation effects. However, we have not taken the quantification of the error score into account. This is done numerically, in order to confirm that our chosen submodel topology does not only show qualitatively correct behavior, but also a good quantitative fit. This fit is

visualized in Figure 17. As we see, modeling the A20 submodel using both an sigmoid activation function and Michaelis Menten based decay results overall in a better qualitative fit (panel B). However, this is to be expected due to the additional degree of freedom the Michaelis Menten equation introduces. We decided that, given its simpler reaction kinetic, the fitting quality of the submodel with linear decay (panel A) is sufficient for now.

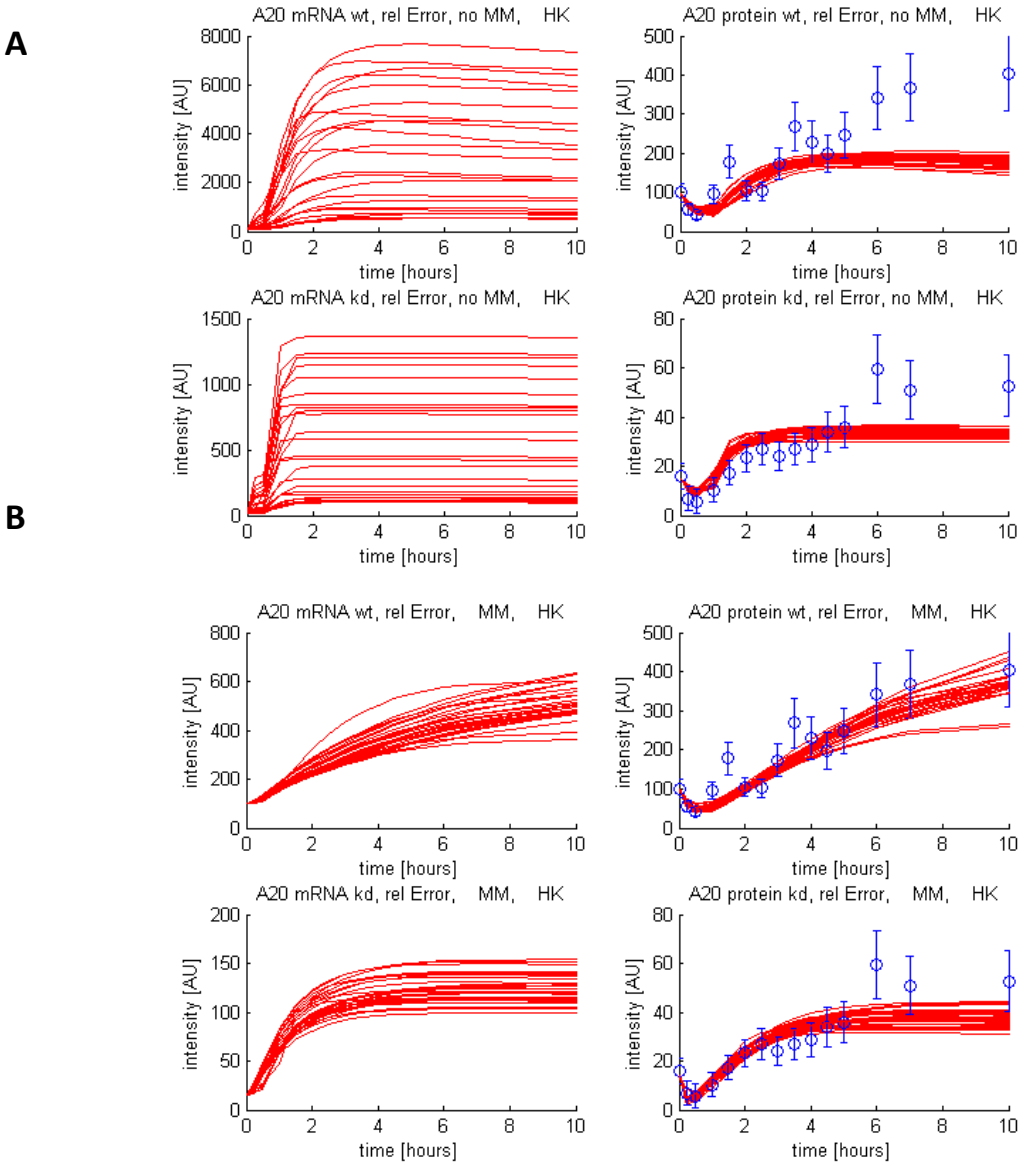


Figure 17: Fits of the A20 submodel with and without Michaelis Menten based A20 degradation
A. Multiple parameters fits of the A20 submodel without Michaelis Menten kinetics. Hill kinetic mRNA induction was used, error was calculated based on the assumption of an relative error. The average error value was around 45 units.
B. Multiple parameters fits of the A20 submodel with Michaelis Menten kinetics. Hill kinetic mRNA induction was used, error was calculated based on the assumption of an relative error. The average error value was around 25 units.

6.1.5. Analysis of the IKK and the IkBa submodels

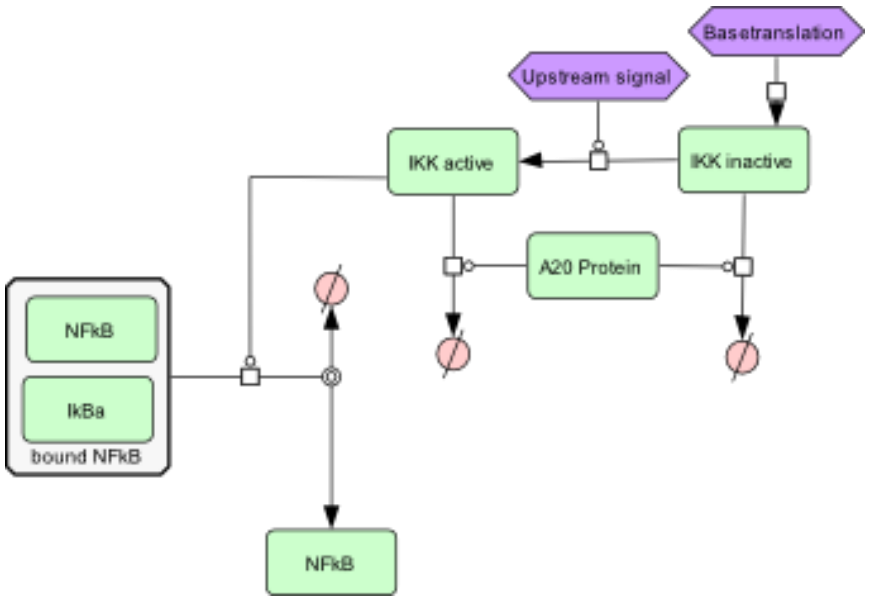
While the A20 submodel can be disconnected from the rest of the model by supplementing our observations as artificial submodel inputs, this is not possible for the IKK activity and the IkBa regulation submodel. The natural "breaking point" would be to separate the model at the state active IKK, using it as output for the IKK submodel and as input for the IkBa regulation submodel. However, we have no direct measurements of IKK, making this direct approach impossible. It is however possible to approximate the activity of IKK based on considerations introduced in the next sections.

The IkBa submodel deals with the regulation of the IkBa monomer. This includes the basal- and NFkB induced transcription of IkBa mRNA, the translation of IkBa mRNA to IkBa protein, the natural decay of the IkBa monomer, the binding processes between IkBa and NFkB and finally the regulation of bound and unbound IkBa by IKK.

The potential degradation of IkBa by IKK has been subject of discussion and experimental studies for quite some time. Early reports state that the IkBa monomer binds to IKK with a very low affinity compared to the IkBaNFkB dimer. In contrast more recent reports state that while IKK binds and phosphorylates IkBa, this does not result in an increase in IkBa decay.

Beyond the issue of IkBa phosphorylation by IKK, we also have to consider the transcriptional dynamics of IkBa mRNA. This is similar to the issue of A20 transcription; the two possibilities are mass action kinetics and sigmoid activation functions with the potential of saturation. Established models by Lipniacki and Hoffman use mass action kinetics of first to third order.

Compared to the IkBa submodel, the IKK activation submodel can be modeled relatively straight forward without contradicting experimental evidence; however this is in part due to the current lack of detailed understanding of the involved reaction mechanics. If recapitulates the synthesis and activation of inactive IKK and the inhibition of IKK by A20. Synthesis and activation steps are modeled using mass action kinetic. For the inhibition of IKK by A20, Michaelis Menten and mass action kinetics are possible candidates.



Change of IKK active and IKK inactive

$$\frac{d}{dt}(IKK_{inactive}) = +R_{IKK}tramslation - R_{IKK_{inactive}}decay - R_{IKK}activation$$

$$\frac{d}{dt}(IKK_{active}) = +R_{IKK}activation - R_{IKK_{active}}decay|base - R_{IKK_{active}}decay|induced$$

Established reaction kinetics:

$R_{IKK}tramslation = k_{IKK}tramslation$	IKK translation is not induced by NFkB and therefore assumed constant
$R_{IKK_{inactive}}decay = k_{IKK_{inactive}}decay * IKK_{inactive}$	IKK basal decay is modeled using first order mass action kinetics
$R_{IKK}activation = k_{IKK}activation * IKK_{inactive} * Upstream$	IKK activation is modeled using second order mass action kinetics
$R_{IKK_{active}}decay\|base = k_{IKK_{active}}decay\|base * IKK_{active}$	IKK_{active} basal decay is modeled using first order mass action kinetics

Variable reaction kinetics:

$R_{IKK_{active}}decay\|induced$	The degradation of active IKK by A20 can be modeled using either first order mass action kinetics or saturation based Michaelis Menten kinetics

Inferred model output:

$R_{IkBa\|NFkB}decay\|induced$	The output of this submodel is not a directly observed state, but rather an inferred reaction value. The degradation of IkBaNFkB complex can be modeled using either linear or mass action kinetics

Figure 18 : IKK submodel, initial version: Visualization and overview of reaction equations

Analyzing different alternative kinetics of IkBa degradation

The degradation of IkBa by different mechanisms is a process that has been studied for more than twenty years. Multiple studies estimated the half life time of unbound IkBa to be between 30-60 min, with some minor cell type specific variance. In general, the unbound IkBa monomer has always been found to be rather unstable [13:Krappmann, D. et. al.].

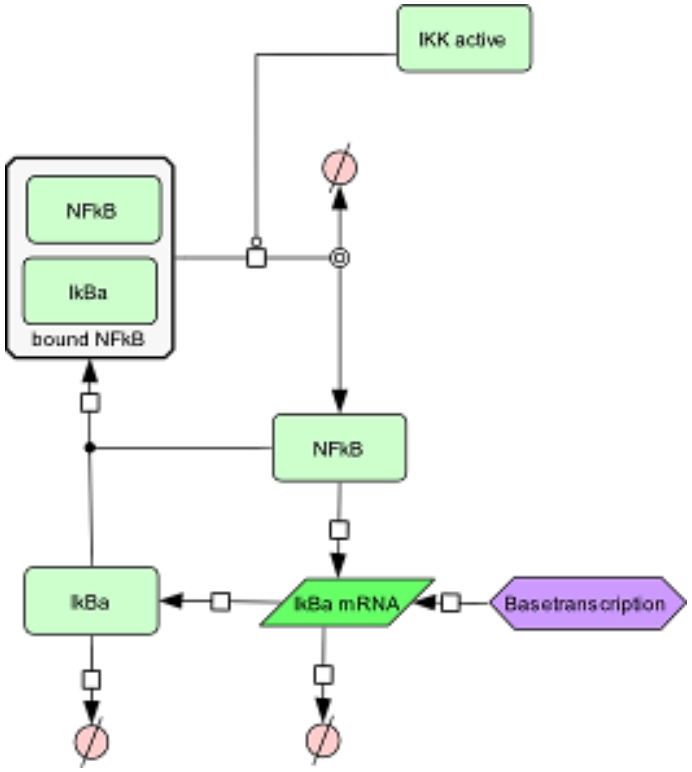
In contrast, the IkBaNFkB dimer shows significantly more susceptibility to the cellular context. The IkBaNFkB dimer has been observed to be very stable (with almost no decay detected after 7 hours of observation) when IkBa and NFkB were transfected into COS-7 cells using expression vectors. In contrast, an experiment using [^{35}S]-labeled IkBa determined the turnover of the IkBaNFkB dimer in various other cell types (among them B-cells, pre B-cells and Jurkat cells) to be similar to the turnover of the IkBa monomer [14:Mathes, E. et. al.].

The phosphorylation of IkBa by IKK adds another layer of complexity to this problem. Both bound and unbound IkBa have been shown to be substrate for IKK. Phosphorylation of bound IkBa leads to its rapid degradation within minutes. However, the stability of the unbound IkBa does not seem to be affected by the phosphorylation [14:Mathes, E. et. al.]. It is known that, upon binding of IkBa to NFkB the IkBa molecule goes through some conformational changes that include refolding of several domains. It has been speculated that only the combination of these refolding processes together with phosphorylation results in the observed degradation of the IkBa subunit, but either conformational changes or phosphorylation alone are insufficient [14:Mathes, E. et. al.].

This implies that IkBa is regulated at a minimum of two levels. In a significant number of cells a pathway that leads to the degradation of both bound and unbound IkBa with a half-life time of about 30 min to 2 hours seems to be active. This pathway causes a significant turnover of the normally stable IkBaNFkB dimer. Most notably, this turnover has been observed to occur in Jurkat cells. We assume that this process is not directly affected by stimulation. Therefore, we model the basal decay of bound and unbound IkBa to be identical.

The second level of regulation is IkBa phosphorylation. Here, the best way to model the process is not clear. We considered the following possibilities and analyzed the degree to which they might impact the model accuracy:

- (1) IkBa monomer is not phosphorylated
- (2) IkBa monomer is degraded by phosphorylation
- (3) IkBa monomer is phosphorylated, but phosphorylation does not increase the decay of it. The phosphorylated IkBa monomer…
    - (3.1) …is not able to bind NFkB
    - (3.2) …is able to bind NFkB, but in doing so is degraded as both phosphorylation and refolding by binding are present in the IkBa monomer.



Change of IkBa mRNA and IkBa monomer:

$$\frac{d}{dt}(IkBa_{mRNA}) = +R_{IkBa}tramscription|base + R_{IkBa}tramscription|induced - R_{IkBa_{mRNA}}decay$$

$$\frac{d}{dt}(IkBa) = +R_{IKK}synth - R_{IkBa}decay - R_{IkBa|NFkB}bind$$

Established reaction kinetics:

$R_{IkBa}tramscription\|base = k_{IkBa}tramscription\|base$	The basal transcription rate of IkBa mRNA is a constant value.
$R_{IkBa_{mRNA}}decay = k_{IkBa_{mRNA}}decay * IkBa_{mRNA}$	The decay of IkBa mRNA happens according to first order mass action kinetics.
$R_{IkBa}synth = k_{IkBa}synth * IkBa_{mRNA}$	The translation of IkBa happens according to first order mass action kinetics
$R_{IkBa}decay = k_{IkBa}decay * IkBa$	The decay of IkBa protein happens according to first order mass action kinetics.
$R_{IkBa\|NFkB}bind = k_{IkBa\|NFkB}bind * IkBa * NFkB$	The binding of IkBa and NFkB is modeled using second order mass action kinetics.

Variable reaction kinetics:

$R_{IkBa}tramscription\|induced$	Induced A20 mRNA transcription can be moeled using either n_{th} order hill- or mass action kinetic

Figure 19: IkBa submodel, initial version: Visualization and overview of reaction equations and kinetics

We start by analyzing variant (1) and (2), the complete lack of phosphorylation and indiscriminate phosphorylation and degradation respectively. Comparing the amount of unbound NFkB in our different experimental setups, we find that knockdown cells show barely any drop in unbound NFkB until about 7 hours, implying that bound NFkB is degraded as fast as it is produced. In contrast, the amount in of bound NFkB in wildtype cells increases steadily, which means that the binding reaction in wildtype cells exceed the corresponding degradation reaction. We can therefore estimate

$$R^{kd}_{IkBa|NFkB} decay|induced \approx R^{kd}_{IkBa|NFkB} bind$$

and

$$R^{wt}_{IkBa|NFkB} decay|induced < R^{wt}_{IkBa|NFkB} bind$$

Based on the binding dynamics as summarized in Figure 19 we estimate the relative rate of IKK activity between knockdown- and wildtype cells for mass action kinetics:

$$\frac{IKK^{kd}}{IKK^{wt}} > \frac{IkBa|NFkB^{wt}}{IkBa|NFkB^{kd}}$$

If we assume complete saturation of substrate, i.e. $R^{wt}_{IkBa|NFkB} decay|induced = v_{max} * IKK$, we can instead estimate:

$$v_{max} * IKK^{kd} \approx k_{IkBa|NFkB}\ bind * IkBa^{kd} * NFkB^{kd}$$
$$> k_{IkBa|NFkB} bind * IkBa^{wt} * NFkB^{wt} > v_{max} * IKK^{wt}$$

With an estimated relative activation rate:

$$\frac{IKK^{kd}}{IKK^{wt}} > \frac{IkBa^{kd} * NFkB^{kd}}{IkBa^{wt} * NFkB^{wt}}$$

We find that

$$\frac{IkBa|NFkB^{wt}}{IkBa|NFkB^{kd}} > \frac{IkBa^{kd} * NFkB^{kd}}{IkBa^{wt} * NFkB^{wt}}$$

holds for all measurements after two hours. This means we can use

$$\frac{IkBa^{kd} * NFkB^{kd}}{IkBa^{wt} * NFkB^{wt}}$$

as a lower boundary for the IKK activity ratio, independently from the actual phosphorylation kinetics utilized. The pool of both IkBa monomer and unbound NFkB is significantly increased in knockdown cells compared to wildtype cells. NFkB is found in excess by a factor of 2.5 to 5, IkBa levels are comparable at around two hours and reach an excess of about 5 times at ten hours. Based on the assumed binding kinetics, this implies that the IKK activity ratio ranges from a

factor of 2.5 at 2 hours to approximately 25 at 10 hours in wildtype cells. This ratio is visualized in Figure 20.

We find that this ratio does conflict with our analysis of the IKK activation submodel. Keep in mind that the effect size of our A20 knockdown is only a decrease to one sixth. As A20 is assumed to degrade IKK in at most a linear fashion, it seems intuitively unlikely that the A20 knockdown can explain an increase in IKK activity by a factor of 12.5-25. To confirm this, we performed multiple parameter optimization attempts with the IKK submodel with linear activation dynamics (as Michaelis Menten kinetics would only have decreased the difference between A20 knockdown and wildtype cells). We tried to optimize the average difference between knockdown- and wildtype activity in the time interval 2-10 hours. However we found that this ratio can't be increased significantly above a ratio of approximately ~6 fold stronger activation in knockdown cells. Therefore, we find that both completely ignoring the phosphorylation of IkBa and modeling direct IkBa degradation by IKK results in a dynamic behavior which does not fit to the effect size of the A20 knockdown as we measured it.

As we see, our inability to explain the phosphorylation and degradation process for mechanisms (1) and (2) with satisfactory results stems from the increase in the binding reaction value in knockdown cells compared to wildtype cells, which is too large to be compensated by the decrease in A20 activity and the resulting increase in IKK activity. This could be compensated in several ways:

- Assuming that a minimum A20 activity is required in order regulate IKK by A20; this would imply that the A20 induced degradation of IKK in knockdown cells is less than 1/6 of the degradation in wildtype cells.
- Assuming that the rate of IkBa and NFkB binding is limited to be less than linear to the total pool of IkBa times NFkB.

The assumption of a minimum A20 activity currently has no biological support. While it is in general possible, as the process of IKK degradation by A20 is not yet well understood, we believe that our data is insufficient to propose such a radically different reaction kinetic. Experimental exploration of this possibility would require comparative quantification of IKK phosphorylation activity between knockdown and wildtype cells after the initial peak phase. If the result would indeed show that IKK activity is increased by more than a factor of 6 in knockdown cells, this hypothesis would be significantly strengthened.

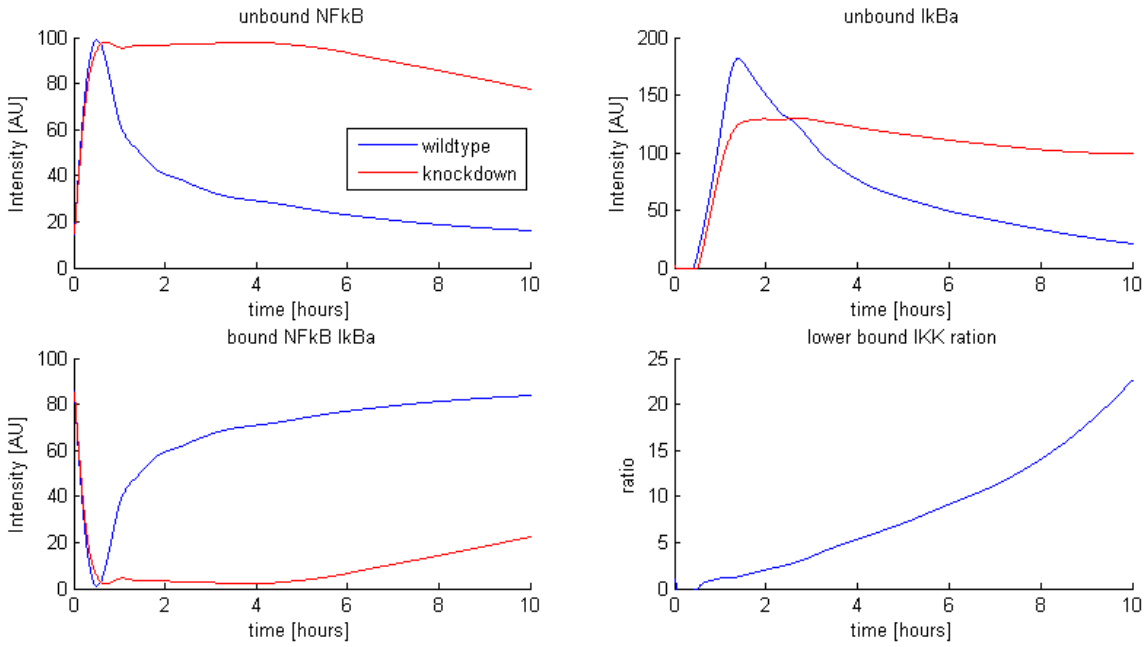


Figure 20: Concentrations relevant for IKK/IkBa submodel and lower boundary for IKK activity ratio between knockdown- and wildtype cells.

A limiting factor for IkBa and NFkB binding processes could be the distribution into different compartments which we decided not to model explicitly due to our resolution being limited to whole cell extract data. If the distribution of IkBa and NFkB happens at unequal ratios in knockdown- and wildtype cells, this could limit the amount of binding activity in knockdown cells. In addition, phosphorylation of IkBaNFkB complexes can only happen in the cytoplasm, as IKK is not found in the nucleus. Therefore, if a larger fraction of the bound complex is found in the nucleus in the wildtype cell line compared to the knockdown cell line, this could explain a reduction of effective phosphorylation in the wildtype cells.

The remaining alternative variants of IkBa phosphorylation, (3.1) and (3.2) offer alternative explanations for an IKK activity ratio that is proportional to the A20 knockdown size. If phosphorylation of the IkBa monomer is assumed to prevent the binding of IkBa and NFkB, the effective value of $IkBa_{kd}$ would have to be reduced by a factor that is determined by the fraction of IkBa currently phosphorylated. The same is true for variant 3.2, where binding of IkBa and NFkB results in immediate degradation of the IkBa molecule; however in this case an additional degradation of the phosphorylated IkBa would have to be introduced. Both variants require us to introduce a separate state pIkBa. This weakens our experimental observations, value of the state free IkBa is no longer uniquely determined by our experimental observations. Instead we can only compare our experiment to a sum of states (which is not unique):

$$IkBa_{mono}^{observed} = IkBa_{mono}^{simulated} + pIkBa_{mono}^{simulated}$$

We limited us to the analysis of the model version (3.2), as we believe it best fits the experimental observations as noted in [14:Mathes, E. et. al.]; however, we believe it is necessary to perform further experiments with more detailed measurements to explore the process of IkBa phosphorylation in relation to A20 in more detail. The adapted model is visualized in Figure 21.

Analyzing the adapted IkBa phosphorylation model

After deciding on an adaption for the phosphorylation topology, we still have the consider which kinetic to utilize for IkBa phosphorylation and whether or not a sigmoid transcription function is required.

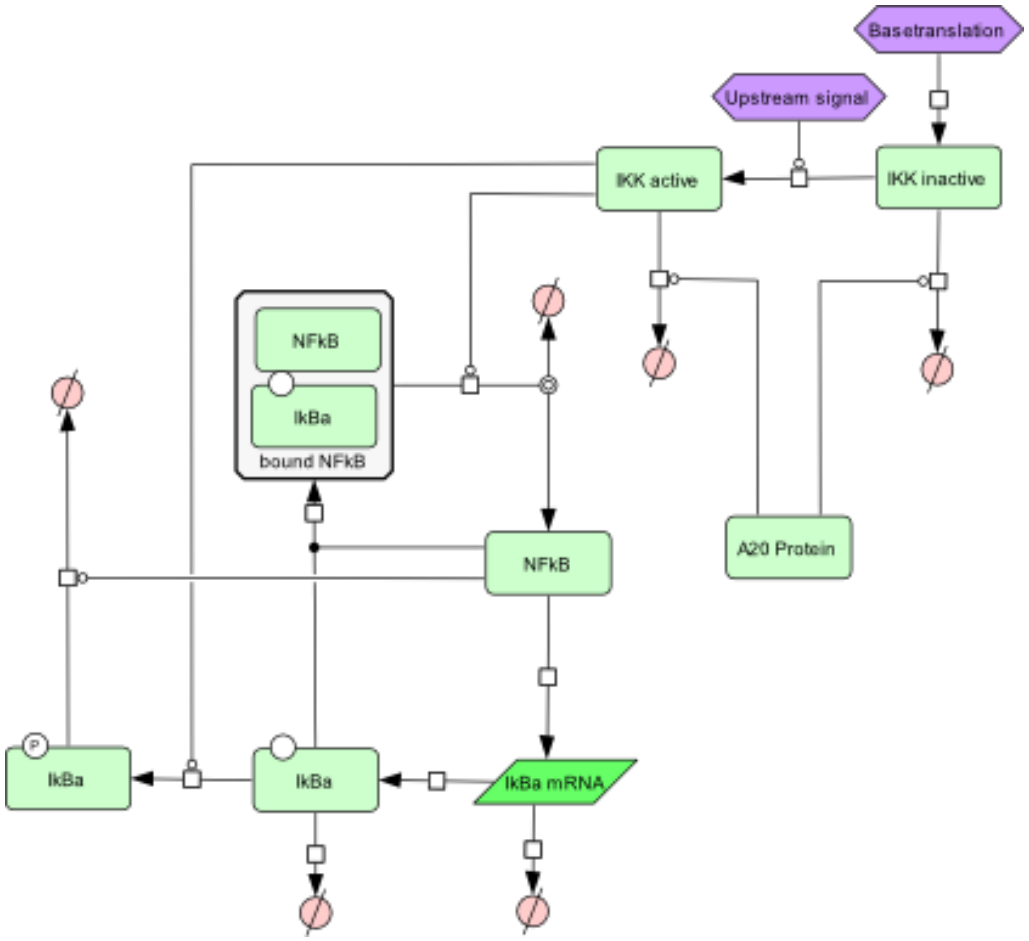


Figure 21: IkBa and IKK submodel, adapted versions: Visualization

Based on insufficient fit quality, we decided to adapt the IkBa submodel. Phosphorylation of unbound IkBa and degradation of unbound, phosphorylated IkBa by NFkB were added as additional reactions.

In regards to the phosphorylation kinetic, we decided that Michaelis Menten kinetics would be unsuitable, as our model has two different substrates for IKK. In such a situation, the basic assumptions behind the standard Michaelis Menten kinetic become violated, and adaptions are required to account for the concurrence between different substrates. To avoid these even more complex kinetics we implemented the model using Mass Action Kinetics and later confirmed its ability to explain the observed experimental measurements numerically.

The analysis of the transcription function can be done rather informally. The total amount of IkBa is comparable in both wildtype and knockdown cells. As both bound- and unbound IkBa is assumed to have a similar basal decay, this implies that the amount of regular decaying IkBa is similar in both cell types. However, as discussed in the previous section, we can estimate that the effective rate of phosphorylation is significantly increased in knockdown cells.

In addition, we can estimate that the amount of IkBa that is degraded due to phosphorylation is significantly large than the amount of IkBa that is degraded by the basal pathway. This implies that the total IkBa decay is increased in knockdown cells by an unknown, but significant factor. It is conceivable that this factor justifies a 5 fold overexpression of IkBa in knockdown cells to reach a similar total equilibrium amount. However, based on purely theoretical considerations, we are unable to confirm this. Therefore, we performed a numerical analysis by implementing and comparing the fit quality of both variants. The results (not shown) suggest that both variants are viable. Therefore, we opted for the variant with fewer parameters, i.e. linear induction dynamics.

6.1.6. Calculation of parameters based on data normalization

The model we implemented based on the assumptions introduced previously in this chapter has a total of 21 parameters. This is a significant number if we desire to perform Bayesian parameter sampling. Sampling of an 21 dimensional space can take significant computational effort if uncorrelated samples are desired (as they generally are). However, several parameters can be indirectly inferred based on normalizing assumptions. We already discussed that the siRNA induced decay of A20 mRNA can be estimated:

$$k_{A20mRNA}induceddecay \; = \; 5 \; * \; k_{A20mRNA}basedecay;$$

Rescaling of basal mRNA levels

The first rescaling step we perform is rescaling of the mRNA levels during the steady states. W.l.o.g. the initial amount of mRNA for both A20 and IkBa can be rescaled to 100[AU]. Due to steady state conditions ($R_{mRNA}decay = R_{mRNA}synth$), this allows us to either calculate the basal transcription based on the mRNA decay or vice versa. We decided to rescale the transcription

rate, as decay rates can generally be understood intuitively as inferred half-life times; therefore sampling of decay rates and calculation of transcription rates results in sampled parameters that are easier interpreted. We find:

$$k_{IkBa}transcription|base \; = \; mRNA_{IkBa}^{ss} * \; k_{IkBa_{mRNA}}decay$$
$$k_{A20}transcription|base \; = \; mRNA_{A20}^{ss} * \; k_{A20_{mRNA}}decay$$

Rescaling of basal synthesis level / basal concentrations

By rescaling total inactive IKK, total IkBa and total A20 to a fixed value, we can once again normalize synthesis parameters based on steady state conditions.

$$k_{IKK}synth \; = IKK^{ss} \; * \; k_{IKK}basedecay;$$

$$k_{IkBa}synth \; = \frac{IkBa_{total}^{ss}}{mRNA_{IkBa}^{ss}} \; * \; k_{IkBa}basedecay;$$

$$k_{A20}synth \; = \frac{A20_{total}^{ss}}{mRNA_{A20}^{ss}} \; * \; k_{A20}basedecay;$$

Normalization of initial levels of bound and unbound IkBa and NFkB

Finally, we can estimate the IkBa NFkB binding ratio, as the decay of bound IkBa equals the binding reaction of new IkBa under steady state conditions:

$$k_{IkBaNFkB}bind \; = \frac{IkBaNFkB^{ss}}{IkBa^{ss} * NFkB^{ss}} \; * \; k\_IkBa\_basedecay;$$

6.1.7. Definition of an ODE model of the canonical pathway

We implemented the ODE model of the canonical pathway as previously discussed in this chapter. The resulting model has 17 states, out of which 9 are biological states and 8 are part of the delay chain utilized to model transcriptional delay. The model has 21 parameters, out of which the value of 8 can be inferred, mostly based on steady state conditions. The remaining 13 parameters have to be fitted. In addition, the model has three constant inputs, degradation of A20 by stimulation, activation of IKK by stimulation and degradation of A20 mRNA by A20 siRNA.

The model has 21 reactions describing biological processes. Most of these reactions have been modeled canonically. Deviating from the established norm, A20 mRNA transcription has been modeled using a 2^{nd} order sigmoid activation function. In contrast, IkBa mRNA synthesis uses a polynomial activation function. All enzymatic reactions have been modeled using 2^{nd} order mass action. The complete model is visualized in Figure 22. This figure also gives an exemplary fit of the ODE model to the experimental data. All information required to implement the model as discussed have been summarized in Table 1 to Table 3.

Table 1: Overview of all parameters and their min-max values during parameter optimization

Parameter name	type	min value	max value	comments
k_IkBa_mRNA_basesynth	inferred	n/a	n/a	-
k_A20_mRNA_basesynth	inferred	n/a	n/a	-
k_A20_mRNA_induceddecay	inferred	n/a	n/a	-
k_IKK_synth	inferred	n/a	n/a	-
k_IkBa_synth	inferred	n/a	n/a	-
k_A20_synth	inferred	n/a	n/a	-
k_IkBa_NFkB_bind	inferred	n/a	n/a	-
k_A20_basedecay;	fitted	$4*10^{-4}$ s^{-1}	$3*10^{-5}$ s^{-1}	half life 30 min – 6 hours
k_A20_induceddecay;	fitted	10^{-3} $s^{-1}[AU]^{-1}$	10^{-6} $s^{-1}[AU]^{-1}$	wide range as no biological prior exists
k_IKK_basedecay;	fitted	$4*10^{-4}$ s^{-1}	$3*10^{-5}$ s^{-1}	half life 30 min – 6 hours
k_IKK_a_induceddecay;	fitted	10^{-3} $s^{-1}[AU]^{-1}$	10^{-6} $s^{-1}[AU]^{-1}$	wide range as no biological prior exists
k_IKK_activation;	fitted	10^{-3} $s^{-1}[AU]^{-1}$	10^{-6} $s^{-1}[AU]^{-1}$	wide range as no biological prior exists
k_IkBa_basedecay;	fitted	~$4*10^{-4}$ s^{-1}	$9*10^{-5}$ s^{-1}	half life 30 min – 2 hours
k_IkBa_phosphorylate;	fitted	10^{-3} $s^{-1}[AU]^{-1}$	10^{-6} $s^{-1}[AU]^{-1}$	wide range as no biological prior exists
k_A20_mRNA_basedecay;	fitted	~$8*10^{-4}$ s^{-1}	$9*10^{-5}$ s^{-1}	half-life 15 min – 2 hours
k_IkBa_mRNA_basedecay;	fitted	~$8*10^{-4}$ s^{-1}	$9*10^{-5}$ s^{-1}	half-life 15 min – 2 hours
k_A20_mRNA_ind	fitted	1	100	relative increase over basal transcription rate
k_IkBa_mRNA_ind	fitted	1	100	relative increase over basal transcription rate
kHill_A20;	fitted	1 [AU]	100 [AU]	half maximum induction 1-100 [AU]
delay_NFkB;	fitted	1200 s	2400 s	20-40 min transcriptional delay

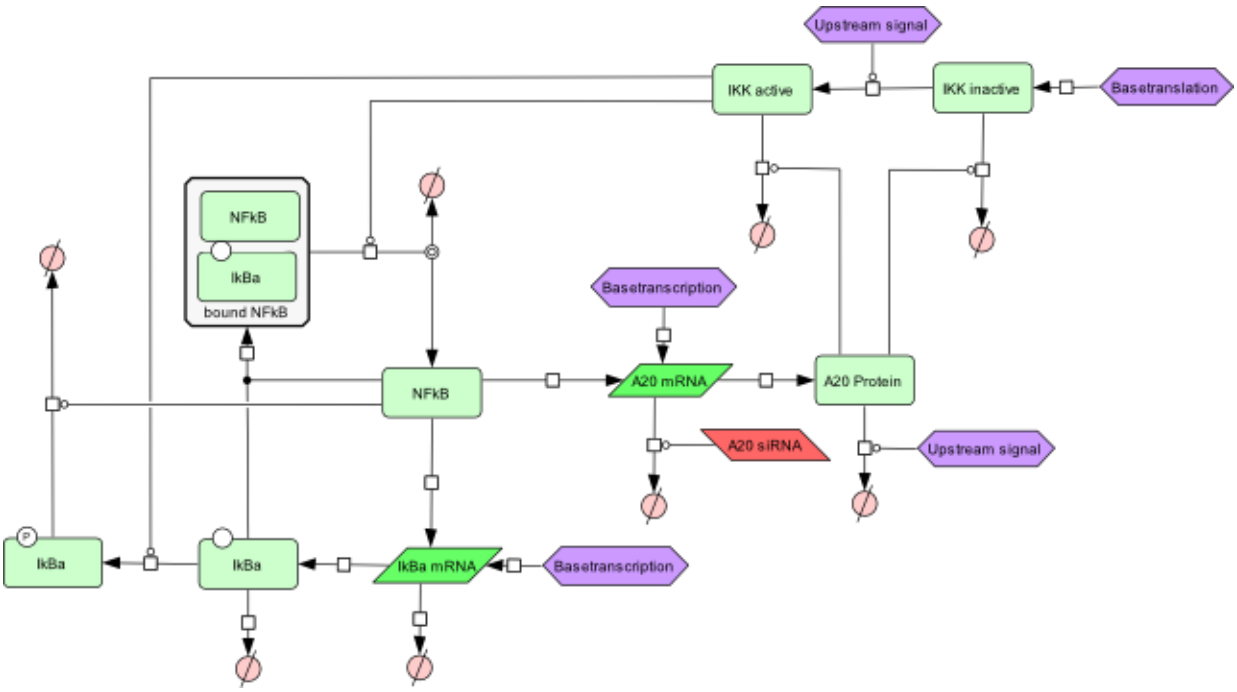
Table 2: Overview of all states of the final model

State name	biological description	produced by	degraded by	other
A20_prot	A20 Protein	R_A20_synth	R_A20_basedecay R_A20_induceddecay	n/a
IKK_i	inactive IKK	R_IKK_synth	R_IKK_i_basedecay R_IKK_activation	n/a
IKK_a	activated IKK	R_IKK_activation	R_IKK_a_basedecay R_IKK_a_induceddecay	n/a
IkBa	unbound IkBa	R_IkBa_synth	R_IkBa_phos R_IkBa_basedecay R_IkBa_NFkB_bind	n/a
pIkBa	phosphorylated, unbound IkBa	R_IkBa_phos	R_pIkBa_basedecay R_pIkBa_bind_and_decay	n/a
NFkB	unbound NFkB	R_IkBa_NFkB_basedecay R_IkBa_NFkB_phosphorylate	R_IkBa_NFkB_bind	n/a
IkBa_NFkB	bound IkBa NFkB dimer	R_IkBa_NFkB_bind	R_IkBa_NFkB_basedecay R_IkBa_NFkB_phosphorylate	n/a
A20_mRNA	A20 mRNA	R_A20_mRNA_synth	R_A20_mRNA_basedecay R_A20_mRNA_induceddecay	n/a
IkBa_mRNA	IkBa mRNA	R_IkBa_mRNA_synth	R_IkBa_mRNA_basedecay	n/a
NFkB_delaychain_1	technical state to realize delay	n/a	n/a	8/delay_NFkB * (NFkB - NFkB_delaychain_1)
NFkB_delaychain_n	technical state to realize delay	n/a	n/a	8/delay_NFkB * (NFkB_delaychain_(n-1)- NFkB_delaychain_n)

Table 3

Reaction Name	Reaction type	Associated parameters	Reaction equation
R_A20_synth	1st order mass action	k_A20_synth	$k_{A20}synth * A20_{mRNA}$
R_A20_induceddecay	2nd order mass action	k_A20_induceddecay	$k_{A20}induceddecay * A20_{prot} * Upstream_{deg}A20$
R_A20_basedecay	1st order mass action	k_A20_basedecay	$k_{A20}basedecay * A20_{prot}$
R_IKK_synth	constant reaction	k_IKK_synth	$k_{IKK}synth$
R_IKK_i_basedecay	1st order mass action	k_IKK_basedecay	$k_{IKK}basedecay * IKK_i$
R_IKK_a_basedecay	1st order mass action	k_IKK_basedecay	$k_{IKK}basedecay * IKK_a$
R_IKK_activation	2nd order mass action	k_IKK_activation	$k_{IKK}activation * IKK_i * Upstream_{act}IKK$
R_IKK_a_induceddecay	2nd order mass action	k_IKK_a_induceddecay	$k_{IKKa}induceddecay * IKK_a * A20_{prot}$
R_IkBa_synth	1st order mass action	k_IkBa_synth	$k_{IkBa}synth * IkBa_{mRNA}$
R_IkBa_NFkB_basedecay	1st order mass action	k_IkBa_basedecay	$k_{IkBa}basedecay * IkBa_{NFkB}$
R_IkBa_basedecay	1st order mass action	k_IkBa_basedecay	$k_{IkBa}basedecay * IkBa$
R_pIkBa_basedecay	1st order mass action	k_IkBa_basedecay	$k_{IkBa}basedecay * pIkBa$
R_IkBa_phos	2nd order mass action	k_IkBa_phosphorylate	$k_{IkBa}phosphorylate * IkBa * IKK_a$
R_IkBa_NFkB_phosphorylate	2nd order mass action	k_IkBa_phosphorylate	$k_{IkBa}phosphorylate * IkBa_{NFkB} * IKK_a$
R_pIkBa_bind_and_decay	2nd order mass action	k_IkBa_NFkB_bind	$k_{IkBaNFkB}bind * pIkBa * NFkB$
R_IkBa_NFkB_bind	2nd order mass action	k_IkBa_NFkB_bind	$k_{IkBaNFkB}bind * IkBa * NFkB$
R_A20_mRNA_synth	sigmoid transcription function	k_A20_mRNA_indsynth k_A20_mRNA_basesynth	$k_{A20mRNA}basesynth * \left(1 + k_{A20mRNA}ind * \left(\frac{NFkB_{delay8}}{NFkB_{delay8} + \mathrm{kHill}_{\mathrm{A20}}}\right)\right)$
R_A20_mRNA_basedecay	1st order mass action	k_A20_mRNA_basedecay	$k_{A20_{mRNA}basedecay} * A20_{mRNA}$
R_A20_mRNA_induceddecay		k_A20_mRNA_induceddecay	$k_{A20_{mRNA}induceddecay} * A20_{mRNA} * A20_{siRNA}$
R_IkBa_mRNA_synth	polynomial transcription function	k_IkBa_mRNA_indsynth k_IkBa_mRNA_basesynth	$k_{IkBamRNA}basesynth * (1 + k_{A20mRNA}ind * NFkB)$
R_IkBa_mRNA_basedecay	1st order mass action	k_IkBa_mRNA_basedecay	$k_{IkBamRNA}basedecay * IkBa_{mRNA}$

A



B

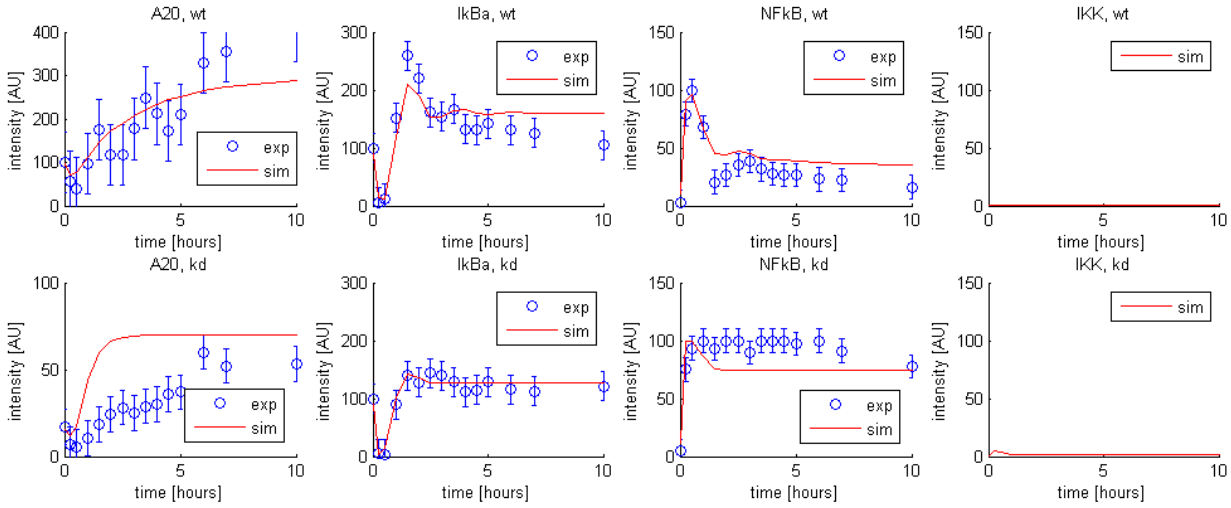


Figure 22: Complete model of the canonical NFkB pathway, final version and fit to experimental data
A. Final model of the canonical NFkB pathway
B. Exemplary data fit of the model to the observed experimental data.

6.2. Sampling parameter ensembles using the Metropolis Hastings Monte Carlo algorithm

Due to the limited scope of this thesis, we did not perform an exhaustive MCMC based analysis of the parameter space of our model. We believe that currently several weaknesses of the experimental data, such as the small number of replicates, have to be improved before we are able to make accurate Bayesian predictions. However, we performed a run with four parallel chains, generating ~100.000 samples each, to illustrate that our model is generally accessible to MCMC approaches. We used a log-normally distributed proposal function with a variance of 2% of the log interval for each parameters (e.g. for a parameter ranging from 10^{-4} to 10^{-6}, the log-based standard deviation was 2*0.02). This resulted in an acceptance rate of approximately 8%, which is below the ideal acceptance rate, but high enough to allow efficient generation of samples. Each chain took about 20 minutes to compute on a standard office PC. Each chain was started from an log-uniform randomly chosen starting point. We analyzed the cross correlation within each chain to confirm that we did indeed generate independent samples and the Gelman Rubin statistic (comparison of inter- vs. intra- chain variance) to confirm that all chains sampled from the same distribution.

6.2.1. Prior for kinetic parameters

Protein decay and mRNA decay

We chose priors for protein- and mRNA decay based on the half-life times that they inferred. For A20 and IKK proteins, a half-life time log-uniformly distributed between 30 min and 6 hours was used. For IkBa, this was reduced to 30 min to 2 hours, based on the reported high turnover of IkBa. All mRNA half-life times are assumed to be between 15 min and 2 hours. These parameter borders all include the values used by Lipniacki et.al. We confirmed that, should new biological evidence arise, these border can be increased by multiple orders of magnitude to accommodate different reaction kinetics.

Enzyme based degradation of IkBa and A20

For all enzyme based we assumed an log-uniform prior between 10^{-3} and 10^{-7}. These priors cover a significantly wider range than those of the other reactions, as we have no intuitive way of interpreting them. The upper boundary was chosen based on the observation that for a rate of 10^{-3} and faster rates, we can simply assume that any reaction happens instantaneous. This is illustrated by the observation that higher rates frequently result in numerical problems, with the solver being forced to pick extremely small step widths. The lower boundary was chosen based on a trial-and-error basis. We fixed both rates separately at a value of 10^{-7} and lower and then optimized all other reaction rates. We found that the smallest value at which any satisfactory data fits were possible was a kinetic rate of about ~10^{-6}.

7. Discussion

7.1. Discussion of fit quality and model complexity

We find that our model fits very well for IkBa both in wildtype and knockdown cells and NFkB/A20 in wildtype cells, but has some deficits for NFkB and A20 in knockdown cells. The amount of free NFkB in knockdown cells is frequently underestimated by our model for time points after the initial peaking phase. In addition, the recovery of A20 in knockdown cells happens slightly too fast and often stabilizes 10-20[AU] above the observed value. However, given the high noise of our observations, resulting in part from the small number of experimental replicates, in part from the difficulty of quantifying blot experiments, neither effect is strong enough to put our general modeling approach into question.

In general, the quantitative fit can be improved further by choosing Michaelis Menten kinetics over mass action kinetics for both degradation of A20 and phosphorylation of IkBa. However these more complex dynamics introduce additional degrees of freedom. We do not feel that these additional degrees of freedom are justified by the qualitative improvement in the error score if model predictions are our goal. On the other hand, for a descriptive model using Michaelis Menten kinetics is less of a problem as these are well established in biochemistry and their utilization can be considered to cause no controversy. This has to be seen in contrast to the assumption of complex transcription dynamics, where each parameter has complex biological implications such as the number of binding sites, their binding specificity for the transcription factor and the interaction of multiple binding sites among each other.

NFkB fit in knockdown cells

Experimental data suggests that in knockdown cells, almost no NFkB is sequestered by IkBa after the initial stimulation. However, our model suggests that about 20-40% of the total NFkB is bound to IkBa after the initial peaking phase. We believe the following aspects of both our experimental data and our modeling approach might be responsible:

One possible explanation is our normalization of the NFkB concentration. Based on biological considerations, we declare the initial peak to be the intensity associated with 100% free NFkB and normalize the remaining data accordingly. Any later time point in our measurements that exceeds this intensity is considered to simply be at maximum intensity. However, the number of points that actually show higher intensities illustrates that this approach to normalization has to be considered problematic.

An additional aspect is the lack of core import/export processes in our model. We already discussed that this could explain that the effective rate of phosphorylation in knockdown cells seems to be increased more than linear to the A20 knockdown efficiency.

The last explanation could be insufficient or incomplete modeling of IKK/A20 interactions. We already mentioned that the exact biochemical mechanisms of these interactions are unknown. Therefore it is possible that A20 influences the activity of IKK in more than one way. If this is the case, a more than linear influence of an A20 knockdown on the IKK level seems conceivable.

A20 fit in knockdown cells.

The stabilization of A20 in knockdown cell lines is predicted to happen slightly faster than observed by our model. This is the same effect we noticed in our analysis of the A20 submodel as illustrated in Figure 15. As discussed, this effect can be reduced by modeling the degradation of A20 by upstream effects with a Michaelis Menten reaction equation instead of a mass action equation. However, alternative explanations are also possible, as the exact mechanisms behind the degradation of A20 are currently unknown. We decided against the inclusion of these more complex kinetics in this model; however we believe that further experimental studies should explore these aspects more, to allow better characterization of the degradation of A20.

Model complexity

In general, we can note that our model significantly simplifies the established signaling models without loss of the ability to explain the experimental observations. While our model has a total of 13 free parameters, established models by Hoffmann and Lipniacki have 20-30 free parameters, depending on whether biological assumptions are counted as free parameters or as inferred parameters. Of course, it has to be kept in mind that researches that manually adapt parameters will often implicitly utilize considerations like those we described in section 6.1.6; therefore it is difficult to estimate the exact degrees of freedom of the established, manually optimized models.

We achieved this reduction in complexity mainly by adapting the model to our measurement data. This included dropping several biological properties we could not estimate based on our measurements. We decided against modeling of absolute molecule numbers, as these could not be estimated from the relative concentrations measured by western blotting. The second example for this is our decision to utilize a single one compartment model instead of the established multi compartment models.

These decisions allow us to reduce the model complexity to a point where Bayesian parameter sampling becomes computationally feasible to make ensemble based predictions on the model behavior. Of course, we can't predict anything regarding the model aspects that have been removed due to simplification. In contrast to the established models we make less detailed predictions (e.g. we are not able to predict perturbations regarding core import/export), that are in exchange statistically better founded due to being ensemble based.

We believe that to some degree more detailed measurements would allow us to handle more complex model topologies, like those in the models of Hoffmann and Lipniacki. An important aspect to consider for future approaches to analyze more complex topologies is our approach of splitting the large signaling model into submodels that can be considered independent under experimental conditions. This allows us to reduce the number of parameters that have to be sampled as a group to small subgroups, but is limited by the availability of smartly chosen "cutting points" in the model, i.e. states that are measured with a resolution that allows interpolation of time course data. We believe that choosing measured states according to their value as cutting points could potentially allow us to split even large, detailed submodels of the canonical pathway into small submodels that would be accessible to Bayesian analysis

Introduction of new transcription kinetic into models of the NFkB pathway

A major result of our work is the introduction of a new, saturation based transcription kinetic for target genes of NFkB compared to the previously established polynomial transcription kinetic. This has far reaching implications:

- Predictions based on our model are massively influenced by the introduction of saturation mechanisms. The introduction of a sigmoid activation function almost automatically introduces ***regions of qualitatively varying stimulus response***.

- Such regions are very interesting from a biological point of view, as it is very conceivable that they are ***associated with the different qualitative responses of T Cells***. Therefore, it is important to ascertain that their existence is indeed required by biological observations and not just a residual effect of the chosen modeling kinetics.

- Connecting models with saturation based dynamics to ***existing Boolean models of gene regulation*** is much more intuitive than connecting models without observable saturation. Boolean models generally only consider the transcription of a gene to be "on" or "off" without further consideration for the actual intensity of gene transcription. For models with saturation effects, an intuitive way to map the response of a qualitative model to a Boolean state is to consider a gene "on" whenever it is closer to its saturation value than its baseline value and "off" otherwise. In contrast, cutoff value to map models without saturation is more difficult to assess, often introducing an additional level of researcher bias. Introduction of biologically justified saturation based transcription dynamics might therefore be a first step in the direction of connecting quantitative modeling of NFkB activity (as done by Hoffmann, Lipniacki and our work) to qualitative regulation of gene networks during T Cell activation.

7.2. Summary of assumptions and design of further experiments

Biological assumptions in the design of our model

In order to keep the degrees of freedom in our model limited, which is necessary to allow efficient sampling, we used a significant number of biological assumptions to justify replacing simple model topologies with more complex ones. It is important to keep these assumptions in mind for future work, as they are likely to be the subject of future experimental studies.

Degradation of IkBa

A significant number of our assumptions focus on the processes of decay, degradation and phosphorylation of bound and unbound IkBa.

- Basal degradation of bound and unbound IkBa is similar, justified by the assumption of an underlying, cell type specific mechanism of degradation that is not influenced by stimulation of the canonical pathway as reported by [14:Mathes, E. et. al.].
- The time it takes for phosphorylated, bound IkBa to decay can be neglected, as it happens on a much faster time scale than the other processes we modeled. This assumption is based on our interpretation of the available data, especially the rapid degradation of IkBa during the first 15-30 minutes of stimulation.
- Phosphorylation of unbound IkBa is not sufficient for fast degradation. This is again based on the work of [14:Mathes, E. et. al.].

Single compartment model

A central assumption to our modeling approach is the assumption that the division of the cell in multiple reaction compartments can be neglected. This implies the assumption that all core import and export processes happen at a fast rate, so that an qssa (quasi steady state approximation) based equilibrium assumption regarding the distribution of both bound and unbound NFkB and IkBa is possible. Experimentally, this could be verified by measuring the distribution of all three molecular species in both wildtype- and knockdown cells at multiple time points. For each species C, our assumption can be considered justified if

$$\forall i,j: \frac{C_i^{nuc}}{C_i^{cyt}} = \frac{C_j^{nuc}}{C_j^{cyt}}$$

where i and j are the experimental indices across wildtype and knockdown experiments. If time points within one experiment strongly violate this relation, we have to assume that for some time region this assumption does not hold and predictions might therefore suffer from our assumption. If the same time points in different experiments violate it, we have to assume that for these time points our assumption is generally invalid.

Shared kinetic rates between different reactions

We assume multiple times that different but similar reactions in the model happen with the same rate constant. This usually implies that we assume that some kind of modification, e.g. phosphorylation of IkBa, does not change the substrate specificity or binding rate to some other protein. To summarize, we assume that:

- Unbound IkBa, unbound pIkBa and bound, unphosphorylated IkBaNFkB complex are all degraded at the same basal level.
- Active IKK and inactive IKK are degraded at the same basal level.
- Bound and unbound IkBa are phosphorylated at the same rate.
- Unbound pIkBa and unbound IkBa bind NFkB at the same rate.

Goals of future experiments

When designing future experiments to elucidate the dynamics of the canonical NFkB pathway, it is sensible to start by evaluating the shortcomings of the current model iteration. There are several points where we had to adapt the model topology or reaction kinetics based on our intuition of the underlying biological processes. The design of future experiments to better justify these decisions seems therefore a natural step.

We feel the most important open point in our analysis is estimating the exact effect of the A20 knockdown on IKK. As stated multiple times, the result of an A20 knockdown to 1/6 of the original amount of A20 is an increase of apparent ***effective IKK induced degradation*** of IkBa by a super linear factor of 12-20, which does not fit well with either our model or the established model by Lipniacki. Biologically, three possible explanations exist.

(1) The effect of an A20 knockdown translates to an ***increase of IKK activity*** in a super-linear fashion. Two alternative explanations for such a super linear relation exist.

- The first would be that the ***interaction of A20 and active IKK*** is more complex than a sequence of binding → modification → unbinding reactions and involves e.g. ***cooperative effects*** between multiple A20 molecules. This explanation seems biologically questionable, as A20 has never been reported to act in homo multimeric complexes.
- The second explanation would be that A20 does not only target active IKK, but also ***additional upstream effectors*** of the activation of IKK. In such a case, the increase of the A20 activity by a factor of 6 would not only reduce IKK activity by ***direct inhibition of IKK***, but also by further ***weakening the signal for IKK activation***. This would likely result in the observed super linear relation between A20 activity and reduction in IKK activity. We already mentioned that several other molecules of the NFkB pathway, such as TRAF and RIP, are ***potential targets of A20***. On the other hand, our current

experimental data does not allow us to propose a potential target, but only to reason that an additional upstream target might exist. Further research would be required to identify such a target, should the hypothesis that an A20 knockdown results in a super linear increase of IKK activity be confirmed by further experiments.

(2) The increase of IKK activity in a linear fashion results in an ***increase of availability of IKK substrate***.

- If the canonical pathway of NFkB regulation is modeled using a ***multi compartment model***, it is possible that the distribution of bound IkBa between cytoplasm and nucleoplasm is connected to the activity of IKK. If we assume that the ***ratio of bound IkBa in the cytoplasm*** is larger for the knockdown cell line compared to the wild type cell line, this could be a possible explanation for the apparently larger rate of IKK activity in the knockdown cell line. As only bound IkBa in the cytoplasm can be phosphorylated, retention of a significant portion of bound IkBa in the nucleoplasm would result in a ***decreased exposition of IkBa to IKK***. This in turn would reduce the effective rate of IkBa phosphorylation in the knockdown cell line and explain why the apparent rate of IKK activity seems to be increased in knockdown cells by a super linear factor. However, not only does this require a multi compartment model, it also implies that the ***export of bound IkBaNFkB*** complex from the nucleoplasm ***happens at a rather slow rate***. We decided not to implement this interpretations, as this assumption is currently not supported by biological evidence.

(3) The increase of IKK activity in a linear fashion results in an ***reduction of successful IkBa – NFkB binding processes***. As we estimate the effective rate of IKK induced IkBa reduction based on the inferred IkBa NFkB binding rate, an decrease of the binding rate would also result in a reduction of the estimated concentration of active IKK.

- The explanation we implemented is the ***phosphorylation of the IkBa monomer without immediate degradation***. This results in what can basically be considered "failed binding attempts" where phosphorylated IkBa binds to NFkB but is immediately degraded due to the additional conformational changes. This results in an effective reduction of the IkBa-NFkB binding process. It is important to note the requirement that ***unbound IkBa is not immediately degraded upon phosphorylation, but significantly more stable than phosphorylated, bound IkBa***.
- We believe that the experimental studies published in [14:Mathes, E. et. al.] favor this explanation.
- As we see in our experimental data fits, this explanation alone ***is not sufficient to explain the entire difference*** between knockdown- and wildtype cells.

An additional aspect that we consider worth exploring is the exact dynamic of the ***transcription of NFkB target genes***. Studies in this regard have been performed, but have currently found little consideration in regards to systematic modeling of the NFkB activation pathway. We believe that our justification of saturation based dynamics for A20 regulation illustrates that these two ***currently separate lines of research*** should be combined. Not only would this allow to better estimate both kinetics and magnitude of transcriptional feedback involved in the canonical pathway of NFkB regulation (for studying the transcriptional dynamics of IkBa and A20), it would also allow better interpretation of the physiological effects of NFkB activation (if known downstream genes of T Cell differentiation are studied). Aspects we consider important to study are ***cooperativity*** between binding sites, ***saturation concentration*** of transcription factors, and ***maximum induction of target genes*** in relation to the biologically relevant concentrations of NFkB found during the stimulation of the NFkB pathway.

Suggested experiments

Based on the discussion of the current state of the model, we want to suggest possible further experiments and discuss how they could influence further modeling and application of our model to the analysis of the NFkB pathway.

Experiment 1: IKK activity profiles

We believe that, given the current state of our experimental data and our model, the generation of IKK activity profiles would be a major step in better characterizing the A20 feedback processes. The current version of our model shows deficits in explaining the extremely low concentration of bound IkBa NFkB complex in knockdown cell lines. While we believe that the effect is not large enough to generally put our model into question, it is noticeable enough to warrant further analysis. The generation of IKK activity profile that can quantitatively be compared between knockdown and wildtype cell lines would allow a clear answer on the question whether the observed super linear increase of IkBa degradation in knockdown cell lines happens due to effects Upstream or Downstream of IKK.

In addition, an IKK activity profile that allows interpolation of data points would allow us to completely split the IKK and IkBa submodel in regards to parameter sampling similar to the way we split the A20 submodel from the rest of the model. This would allow for further numerical improvements in regards to simulation speed.

Experiment 2: qPCR of both A20 and IkBa mRNA

Another worthwhile experiment to perform could be the quantification of both A20 and IkBa mRNA levels. This would not only help to further determine which kinetics should be used to model mRNA transcription, but would also help with estimating the overall magnitude (i.e. the increase in protein turnover) of IkBa and A20 degradation which is currently not well determined.

8. Conclusion and final perspectives

In conclusion, we believe that the work we report in this thesis can contribute to the analysis of the canonical NFkB pathway in multiple different ways. We illustrate that multi compartment models are not required to produce qualitatively correct model fits, and might not even be needed for quantitative fits. Compared to established models, numerical handling of our model is simpler, allowing the application of automated, non-biased parameter optimization and sampling techniques.

By justifying the introduction of saturation based transcription kinetics, we illustrate that it is indeed possible to study details of the NFkB pathway based on the general ODE modeling approach. Hill kinetic transcription kinetics might also serve as a bridge different lines of research, e.g. the "macroscopic" view on the complete regulatory NFkB pathway and detail focused analysis of binding site structures, affinity and cooperativity.

However we also illustrate that several key points in the regulation of IKK by A20 are still open and cannot be explained by the current model, while at the same time the available experimental data does not allow us to expand the model in a non-arbitrary way. Here, the main question is whether the knockdown of A20 affects multiple targets Upstream of IKK (hypothesis regarding the role of A20). We are happy to report that additional experiments that will hopefully help to understand the situation are currently being planned.

Reference List

1. **Hoffmann,A., Levchenko,A., Scott,M.L., & Baltimore,D. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* 298, 1241-1245 (2002).**

2. **Cheong,R., Hoffmann,A., & Levchenko,A. Understanding NF-kappaB signaling via mathematical modeling. *Mol. Syst. Biol.* 4, 192 (2008).**

3. **Courtois,G. & Gilmore,T.D. Mutations in the NF-kappaB signaling pathway: implications for human disease. *Oncogene* 25, 6831-6843 (2006).**

4. **Hiscott,J., Kwon,H., & Genin,P. Hostile takeovers: viral appropriation of the NF-kappaB pathway. *J. Clin. Invest* 107, 143-151 (2001).**

5. **O'Neill,L.A. & Kaltschmidt,C. NF-kappa B: a crucial transcription factor for glial and neuronal cell function. *Trends Neurosci.* 20, 252-258 (1997).**

6. **Hayden,M.S. & Ghosh,S. Signaling to NF-kappaB. *Genes Dev.* 18, 2195-2224 (2004).**

7. **Gilmore,T.D. & Herscovitch,M. Inhibitors of NF-kappaB signaling: 785 and counting. *Oncogene* 25, 6887-6899 (2006).**

8. **Bonizzi,G. & Karin,M. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends Immunol.* 25, 280-288 (2004).**

9. **Oeckinghaus,A. *et al.* Malt1 ubiquitination triggers NF-kappaB signaling upon T-cell activation. *EMBO J.* 26, 4634-4645 (2007).**

10. **Duwel,M. *et al.* A20 negatively regulates T cell receptor signaling to NF-kappaB by cleaving Malt1 ubiquitin chains. *J. Immunol.* 182, 7718-7728 (2009).**

11. **Sachdev,S., Hoffmann,A., & Hannink,M. Nuclear localization of IkappaB alpha is mediated by the second ankyrin repeat: the IkappaB alpha ankyrin repeats define a novel class of cis-acting nuclear import sequences. *Mol. Cell Biol.* 18, 2524-2534 (1998).**

12. **Gilmore,T.D. Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene* 25, 6680-6684 (2006).**

13. **Krappmann,D., Wulczyn,F.G., & Scheidereit,C. Different mechanisms control signal-induced degradation and basal turnover of the NF-kappaB inhibitor IkappaB alpha in vivo. *EMBO J.* 15, 6716-6726 (1996).**

14. **Mathes,E., O'Dea,E.L., Hoffmann,A., & Ghosh,G. NF-kappaB dictates the degradation pathway of IkappaBalpha. *EMBO J.* 27, 1357-1367 (2008).**

15. **Heyninck,K. & Beyaert,R. A20 inhibits NF-kappaB activation by dual ubiquitin-editing functions. *Trends Biochem. Sci.* 30, 1-4 (2005).**

16. **Basak,S. *et al.* A fourth IkappaB protein within the NF-kappaB signaling module. *Cell* 128, 369-381 (2007).**

17. **Lipniacki,T., Paszek,P., Brasier,A.R., Luxon,B., & Kimmel,M. Mathematical model of NF-kB regulatory module. *Journal of Theoretical Biology* 228, 195-215 (2004).**