Modelling and analysis of the NF-κB pathway in Bio-PEPA

Federica Ciocchetta[1], Andrea Degasperi[2], John K. Heath[3] and Jane Hillston[4]

[1] Microsoft Research - University of Trento Centre for Computational and Systems Biology, Trento, Italy
ciocchetta@cosbi.eu
[2] Department of Computing Science, University of Glasgow, G12 8QQ, UK
andrea@dcs.gla.ac.uk
[3] School of Biosciences, University of Birmingham, B15 2TT, UK
J.K.Heath@bham.ac.uk
[4] Laboratory for Foundations of Computer Science, University of Edinburgh, EH8 9AB, UK and Centre for Systems Biology at Edinburgh (CSBE)[‡], Edinburgh, Scotland
jeh@inf.ed.ac.uk

Abstract. In this work we present a Bio-PEPA model describing the Nuclear Factor κB (NF-κB) signalling pathway. In particular our model focuses on the dynamic response of NF-κB to an external stimulus. Each biochemical species in the pathway is represented by a specific Bio-PEPA component and the external stimulus is abstracted by time-dependent Bio-PEPA events describing the start and the end of the signal.
The Bio-PEPA model is a formal intermediate representation of the pathway on which various kinds of analysis can be performed. Both stochastic and deterministic simulations are carried out to validate our model against the experimental data and *in-silico* experiments in the literature and to verify some properties, such as, the impact of the duration of the external stimulus and of the total NF-κB initial amount on the behaviour of some species of interest. Furthermore we use stochastic simulation to compare the behaviour of the single cell against the average behaviour of a population of cells. Finally, sensitivity analysis is considered to investigate the most influential parameters of the model. Importantly, the approach taken suggests that the sensitivity of some parameters alters with the time evolution of the pathway.

Keywords. Process algebras, NF-κB pathway, modelling, analysis

1 Introduction

Nuclear Factor κB (NF-κB) is a protein complex that regulates numerous genes that play important roles in inter- and intra-cellular signalling, cellular stress response, cell growth, survival and apoptosis [1,2]. The investigation of the specific mechanisms that

[‡] The Centre for Systems Biology at Edinburgh is a Centre for Integrative Systems Biology (CISB) funded by the BBSRC and EPSRC in 2006.

govern NF-κB activities is essential for the understanding of various biological processes and for the potential use of NF-κB as a drug target. In the literature there are numerous models describing different aspects of the NF-κB pathway [3,4,5,6,7,8]. These models are based on various assumptions about the biochemical mechanisms involved and describe specific subsets of species and interactions. Most of them are defined in terms of Ordinary Differential Equations (ODEs) and the validation and analysis are based on numerical integration of ODEs. Furthermore, in [9,10] Ihekwaba *et al.* proposed a *Gepasi* [11][1] model of the pathway and analysed it using deterministic (i.e. ODE-based) sensitivity analysis.

Recently, there have been various applications of process algebras for the modelling and analysis of biochemical networks [12,13,14,15,16,17]. These formalisms were originally defined in the context of concurrent systems in computer science and are useful in the field of systems biology too. In particular, among their various advantages and properties, they offer a formal model of a system in terms of interacting components, and support a compositional approach to model construction. They exemplify *algorithmic* or *executable* systems biology [18,19], an approach in which the intention is to construct models which are more than simply a mathematical function which recreates the mapping from system input to system output.

In this work we consider the process algebra Bio-PEPA [20,21] and define a Bio-PEPA model for the NF-κB pathway earlier described by Lipniacki *et al.* [5] in terms of a system of ODEs (we will refer to this as the *Lipniacki model*). Our choice is motivated by the fact that, at the time of our study, the Lipniacki model was the most recent concerning the NF-κB pathway and it made more realistic assumptions than previous models. Specifically, the model is characterized by explicit handling of compartments, with sizes derived from biological knowledge [22], and by the transport of species between compartments. Furthermore, the model takes into account some experimental constraints on important species at the steady state (for instance, the level of free IκBα is less than 15% of the total IκBα [22]). The pathway considered, and the data used for the validation, derive from mouse fibroblasts [23,3].

The aim of this work is twofold. Firstly, we demonstrate the power of Bio-PEPA as a modelling language for biochemical networks. In particular, we show how to capture some features of these networks in Bio-PEPA, such as static compartments and the presence of external stimuli which cause the activation of some reactions, using locations and temporal events, respectively. Furthermore, the model also benefits from Bio-PEPA's support for generic kinetic laws by means of functional rates, the explicit definition of stoichiometry and recording the role of each species in a reaction.

Secondly, we use some of the analysis techniques defined for Bio-PEPA in order to extend the existing analysis of the model. Whilst previous work focused on a set of ODEs which were subject to numerical integration and deterministic sensitivity analysis, we perform stochastic simulation to verify the possible variability across several

[1] Gepasi is a software package for modelling biochemical systems. It provides a number of tools to fit models to data, optimize any function of the model, perform metabolic control analysis and linear stability analysis. Gepasi translates the language of chemistry (reactions) to mathematics (ODEs).

runs of the model[2]. We first consider the average of several runs in order to obtain the average behaviour of a population of cells and validate our results against the experimental data, derived from a population of mouse fibroblasts [23,3]. Then we focus on single simulation runs and compare them with the average of several simulation runs. We are particularly interested in the oscillatory behaviour of nuclear NF-κB as it seems to have an important role in essential activities of the cell [7,8]. With this in mind, we perform some in silico experiments to investigate the impact of the duration of the signal and the influence of two inhibitors of the NF-κB, the proteins A20 and IκBα, on the oscillations. Finally, we use sensitivity analysis on that stochastic model to isolate the most influential parameters. This sensitivity analysis is applied using a novel algorithm, based on the definition of histogram distance over the simulation runs [24], implemented in the version of the Dizzy simulator developed at the University of Edinburgh [25]. Our analysis is complementary to the previous work in [9]; in particular the model properties are analysed from a different point of view and the differences between the two approaches are discussed.

A preliminary study of the pathway in Bio-PEPA was presented at the Dagstuhl seminar "Formal Methods in Molecular Biology" [26]. Here we give more details about the pathway and the approach used and we report new analysis results and experiments, especially with respect to the sensitivity analysis.

The rest of the paper is structured as follows. The NF-κB pathway and the Bio-PEPA model of the pathway are described in Sect. 2 and Sect. 5, respectively. Sect. 3 report an overview of related work. Bio-PEPA is introduced in Sect. 4. In Sect. 6 the validation of our model and some analysis results are presented. Finally, in Sect. 7 we report some concluding remarks.

2 The NF-κB pathway

In the following we describe the pathway as captured by the Lipniacki model since it is our reference for this work. A general schema of the pathway is reported in Fig. 1. The main species involved in the pathway are the IκB kinase (IKK), NF-κB, the protein A20, the protein IκBα, their complexes, mRNA transcripts of A20, IκBα and a hypothetical control gene (cgen). The species cgen represents a control gene, regulated by NF-κB, distinct from the genes corresponding to A20 and IκBα. In the absence of an external stimulus (i.e. normal condition), NF-κB is bound to the inhibitor protein IκBα and remains in the cytoplasm. When an upstream stimulus (SIGNAL), such as the Tumor Necrosis Factor (TNF) or the interleukin-1α (IL-1α), is received, the IKK protein in the neutral form[3](IKKn) is transformed into its active phosphorylated form (IKKa) and then it is modified, under the influence of the stimulus and the protein A20, into another inactive form (IKKi). The inactive form IKKi is different from IKKn as it is overphosphorylated. The activation of IKK is enabled only when the stimulus is

[2] Note that each run can represent the behaviour of a single cell. If we assume that each cell is independent of the others, multiple simulation runs can approximate the average behaviour of a population of cells.

[3] Neutral refers to IKK in absence of any extracellular stimuli. Neutral IKK does not interact with IκBα and therefore does not trigger the cascade of the NF-κB pathway.

present whereas its inactivation is possible both in the presence of the stimulus (in this case it is activated by A20) and also in the normal condition. When activated, IKKa can trigger the degradation of IκBα, which has the effect of releasing free cytoplasmic NF-κB. This enters the nucleus and upregulates the transcription of the two inhibitors, A20 and IκBα, and a large number of other genes (represented by cgen in Fig. 1). The newly synthesized IκBα again inhibits NF-κB while A20 can inhibit IKKa by catalysing its transformation into IKKi, which is no longer able to trigger the degradation of IκBα.

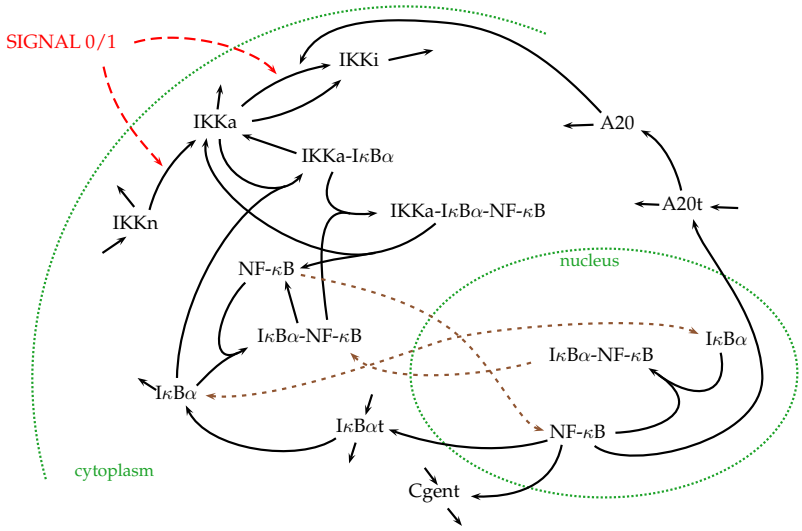

Fig. 1. Schematic depiction of the NF-κB signalling pathway considered in the paper. Signal 0 corresponds to the absence of the external stimulus, signal 1 corresponds to the presence of the stimulus. The red long dashed arrows are the interactions triggered by the signal, the brown short dashed arrows the transport reactions between compartments, black continuous arrows represent all the other kinds of interaction (association and dissociation reactions, translation of mRNAs into proteins). Compartments are delimited by green dotted lines.

The pathway is characterized by the following main features.

1. There are *two compartments*, the nucleus and the cytoplasm. Realistic compartment sizes have been obtained from experiments and this information is taken into account in the derivation of rates and concentrations. These data relate to mouse fibroblasts [22]. Cytoplasmatic NF-κB, the complex nuclear IκBα-NF-κB, nuclear and cytoplasmatic IκBα, can move from one compartment to the other.
2. mRNA transcripts move from the nucleus to the cytoplasm as soon as they are created. Therefore, the translation of mRNA in the associated proteins happens in the cytoplasm. Whilst transcription is stimulated by NF-κB (the *inducible term* in

the ODE representation), a basic rate of transcription will be observed at all times (the *steady term* in the ODE representation).

3. The *external stimulus* is represented by a *signal*. In [5] Lipniacki *et al.* assume that the stimulus is of long duration (persistent); indeed it is active for 6 hours, starting after 1 hour. Short pulse-like stimuli can also be considered to describe specific kinds of inflammation [3,8]. The effect of the signal is to enable some reactions (i.e. the rate becomes non-zero); specifically, the activation of IKKn and the transformation of IKKa into IKKi.
4. There are two *regulatory feedback loops*: the former involves IκBα and the latter A20. Indeed nuclear NF-κB upregulates the transcription of both proteins and these, in turn, inhibit the activity of NF-κB. In the latter feedback loop the action of A20 on the regulation of NF-κB is not direct: A20 inactivates IKKa, this stops the degradation of IκBα and, consequently, there is an increase in the inhibition of NF-κB. We do not consider the third feedback loop in Ashall *et al.* [8] as in our model we consider just one of the IκB isoforms.

Some simplifications are made in the construction of the model (the most important ones are reported below). In particular, it involves a restricted number of species and reactions. As explained above, the model closely follows the Lipniacki model, reflecting the available information and making a number of assumptions in order to simplify the analysis [5].

- First, NF-κB and IKK are protein complexes, but the details of their structure and the complicated kinetics leading to their formation are neglected. Specifically, NF-κB proteins are small groups of dimeric transcription factors which consist of different members (for instance, in mammals these are RelA, Rel, RelB, p50, p52).
- Second, the inhibitory proteins A20 and IκBα mimic the common activity of groups of inhibitors. For instance, IκBα is just one of the possible IκB isoforms involved in the pathway. The choice to consider only IκBα and not all the other isoforms (see for instance Hoffmann's model [3]) reflects the fact that this isoform is the most active and abundant in the cell and its absence, in contrast to the other isoforms, is lethal [27].
- Third, all the other proteins which are not considered in the model remain at their normal (i.e. in the absence of signal) levels.
- Fourth, IKK has three different forms. Each of them undergoes degradation with the same rate and the normal form IKKn is the only one that can be synthesised. We can obtain the inactive form only from the active form IKKa and this is in part independent from the external stimulation.

The Lipniacki model describes the pathway reported above in terms of a system of ODEs, where variables stands for species concentrations (in micro molars, μM). All the kinetic laws associated with the various interactions are simply mass-action.

In order to validate the model, Lipniacki *et al.* analyse its ability to reproduce the data from experiments on mouse fibroblasts. Hoffmann *et al.* report measurements involving wild type cells (i.e. without external stimulation) in response to persistent and pulse-like TNF activation [3]. On the other hand, Lee *et al.* measure the response of wild type and A20-deficient cells to a persistent TNF signal [23]. The experimental

data concern the average behaviour of a cell population (10^6 cells in [3]). We consider these experimental data to validate our model as well.

Due to the large number of unknown parameters, Lipniacki *et al.* proposed the following parameter estimation approach. They started from a reasonable set of parameters obtained from the literature (for instance from the experiments in [22,23]) able to produce a correct steady state in the absence of a TNF signal. Then they proceeded with the signal initiated by TNF and iterated until the fit to all the available data is satisfactory. In the definition of the parameters, the difference between the nuclear and the cytoplasmatic volumes and constraints following from experimental data are taken into account. In particular, as the original system of ODEs is in terms of concentrations, Lipniacki *et al.* took the proportion factor between the two compartments $k = V_c/V_n = 5$ into account in the kinetic laws of the reactions involving reactants and products in two different compartments, such as the transport of a species from the nucleus to the cytoplasm. Indeed, for a given number of molecules, the corresponding concentration of the species in the cytoplasm is k times less than the corresponding concentration in the nucleus and both the concentrations are present in the differential equations in the terms representing transport and interactions between compartments. The explicit and realistic use of compartments leads to more realistic constant parameters than in the previous model by Hoffmann *et al.* [3], especially with respect to the translation and the transport rates.

3 Related work

There is a vast literature of models for the NF-κB signalling pathway [3,4,5,6,7,8], and each model focuses on particular aspects of the pathway and refers to different sets of available experimental data.

The first attempt to model the IκB–NF-κB signalling module was made by Hoffmann *et al.* [3]. They defined an ODE model describing the interplay between three isoforms of the inhibitor proteins IκB (IκBα, IκBβ and IκBϵ) and NF-κB, under both a persistent and short pulse-like TNF stimulus. They proved that IκBα is responsible for a strong negative feedback loop that allows a fast turn-off of the NF-κB whereas the other two isoforms reduce the system's oscillations and stabilize NF-κB responses during longer stimulation. In contrast to the Lipniacki model, they assumed that the two compartments (i.e. the nucleus and the cytoplasm) have the same volume. This assumption simplified the definition of the system of ODEs, but leads to unrealistic transport rates. Their model is fitted against the experimental data reported in the paper and was able to reproduce the expected behaviour. However, the model does not satisfy all the experimental constraints; for instance constraints on the quantity of free IκBα are not considered. Moreover it does not consider some keys species and mechanisms, such as the protein A20 and the IKK activation/disactivation. These two proteins and the related processes have an essential role in the pathway; indeed the knockout of A20 in mice dramatically alters the cell response to TNF stimulation due to persistent IKK activity and causes A20-deficient cells to die prematurely [23]. These shortcomings are addressed in the Lipniacki model and, therefore, in our work too.

Subsequent work by Ihekwaba *et al.* conducts a sensitivity analysis of the model presented by *Hoffmann et al.* [9,10]. In this case the model was defined in the notation of Gepasi [11], but it is a close representation of Hoffmann *et al.*'s ODE model. Thus as with that earlier model, Ihekwaba *et al*'s model fails A20 and IKK activation/inactivation, and supposed equal compartment sizes for the cytoplasm and the nucleus. In translating the model, the authors found some discrepancies in the Hoffmann's supplementary material and proposed a new set of parameters for ten of the reactions of the pathway in order to fit the available experimental data. These parameters concern the synthesis of the various IκB isoforms and association/dissociation of complexes involving NF-κB, IKK and IκB isoforms[4]. The Gepasi model was then mapped to ODEs in order to obtain the temporal evolution of the species and analysed using the parameter scan capability in order to identify those parameters in the IκB–NF-κB system (containing only the IκBα isoform) that most affect the oscillatory concentration of nuclear NF-κB, in terms of *period* (time taken for one oscillation), *phase* (the timing of the beginning of the period) and *amplitude* (the range values attained during an oscillation). Parametric sensitivity analysis was performed on all the system's parameters: each parameter was considered in isolation and varied in order to see the impact of its variation on the behaviour of nuclear NF-κB. Of the 64 parameters in the model, just nine exerted significant influence and these mainly involved IκBα and IKK. A more advanced analysis of the parameters was reported in [10], where pairwise modulation of the nine parameters found in the previous study was carried out. Synergistic effects were observed: the effect of one of the parameters was strongly dependent on the values of another parameter, proving a very strong non-linearity in the system, as expected.

A hybrid variant of the Lipniacki model is proposed by the same authors in [6]: ordinary differential equations, used for description of fast reaction channels of processes involving a large number of molecules, are combined with a stochastic switch to account for the activity of the genes involved. In this way the authors improved simulation efficiency and defined a model able to appropriately handle the small numbers of transcripts. This model was defined in MATLAB and an *ad hoc* algorithm for the simulation was implemented. In the other models defined in terms of ODEs these species are approximated by continuous variables with very low concentrations. This simplification is not realistic. However, for this pathway, it gives correct average behaviours.

Recently a new model for the NF-κB pathway has been proposed by Ashall *et al.* [8]. The authors define a hybrid model, similar to the one proposed in [6] and an associated simulation algorithm, able to capture the behaviour of the pathway under repeated short pulses of TNF at various intervals, mimicking pulsatile inflammatory signals. In order to reproduce the new experimental results based on pulsatile stimulation they modified the Lipniacki model. Specifically they proposed a new set of reactions to describe the IKK activation/disactivation. Differently from the Lipniacki model, inactive IKK (IKKi) can be transformed into neutral IKK (IKKn) and A20 is an inhibitor of this interaction. Furthermore, a third feedback loop is added describing the influence of the isoform IκBϵ on the pathway. Ashall *et al.* used their model to investigate the possible role of the NF-κB oscillations in the systems and how these oscillations are affected, for instance, by the various TNF pulse frequencies and the presence of feedback loops. In

[4] Note that most of these reactions are not present in the Lipniacki model.

the Lipniacki model this third loop is not considered because the focus is on the most abundant and important isoform of IκB.

Cho *et al.* [4] focused on just the first part of the signalling cascade up to the activation of NF-κB, without considering the translocation into the nucleus and the activation of the transcription of the various genes by NF-κB. The authors proposed an ODE model describing in detail the various interactions that leads to the activation of NF-κB.

All the models reported above are defined directly in terms of systems of ODEs, as hybrid models or as Gepasi models (and then translated into ODEs). In the literature there are just a few applications of process algebras for the modelling and study of this pathway [28,29]. In both these papers the focus was on the modelling of the pathway using process algebras and just the validation of the model against the literature was reported. Larcher *et al.* [29] represented the pathway previously described by Hoffmann *et al.* [3] in *BetaWB*, a language based on Beta binders [15]. A clear mapping exists from biochemical entities, such as species and reactions, into the BetaWB language. Each species is abstracted by a bio-process, a box with an interface representing its interaction capabilities. The interactions among species are represented by various kinds of actions such as the formation of complexes and decomplexation, the communication between species and the change of the interface, the deletion or creation of a box, the join and split of boxes. The analysis is based on stochastic simulation and on the comparison with the results in the literature. Note that in [3] the two compartments are assumed to have the same size. The compartments are not considered explicitly in the BetaBW model but their equal size is reflected in the derivation of the number of molecules and stochastic rates. In [28] Hillston and Duguid defined a PEPA [30] model for the pathway presented in Cho *et al.* [4]. The level of abstraction is higher than the one in the work by Larcher *et al.*, in particular, species are abstracted by processes and all the reactions by interactions between processes. Hillston and Duguid considered the reagent-centric style of PEPA, on which Bio-PEPA is based. However, in PEPA it is not possible to defined some features of biochemical systems, such as stoichiometry and events. Furthermore, in the standard version, PEPA does not support functional rates. In [28], the map from the PEPA model into the associated ODE model is presented and ODE numerical integration results are shown.

4 Bio-PEPA

In this section we give a short description of Bio-PEPA [20,21], a language that has been developed for the modelling and analysis of biological systems. Recently Bio-PEPA has been extended to incorporate events [31] and to support biological locations [32], two features which will be useful in developing our model of the NF -κB pathway.

The main descriptive components of a Bio-PEPA system are the *species (or sequential) components*, describing the behaviour of each species, and the *model component*, describing the interactions between the various species. The species initial amounts are given in the model component.

The syntax of the Bio-PEPA components is defined as:

$$S ::= (\alpha, \kappa)\ \mathsf{op}\ S \mid S + S \mid C \mid C@L \qquad \text{with } \mathsf{op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

$$P ::= P \underset{\mathcal{H}}{\bowtie} P \mid S[x]$$

where S is the *species component* and P is the *model component*. We assume a countable set of model components C, a a countable set of locations $\mathcal{L}$ and a countable set of action types $\mathcal{A}$. These three sets are disjoint. In the prefix term (α, κ) op S, the action type α is ranged over by the set $\mathcal{A}$ and abstracts a reaction of the network, $\kappa \in \mathbb{N}$ is the *stoichiometry coefficient* of species S in reaction α and the *prefix combinator* "op" represents the role of S in the reaction. Specifically, $\downarrow$ indicates a *reactant* (i.e. the amount of the species decreases), $\uparrow$ a *product* (i.e. the amount of the species increases), $\oplus$ an *activator* (i.e. the species activates the reaction without modifying its amount), $\ominus$ an *inhibitor* (i.e. the species inhibits the reaction without modifying its amount) and $\odot$ a generic *modifier*. The general modifier operator is useful to indicate species that are involved in a reaction without changing their concentration but which cannot be classified as activators or inhibitors (e.g. a gene during transcription). We can use "(α, κ) op " as an abbreviation for "(α, κ) op S". The operator "+" expresses the choice between possible actions, and the constant C is defined by an equation $C \stackrel{def}{=} S$. The notation $C@L$ indicates that the species represented by the component C is in the location L.

The process $P \underset{\mathcal{H}}{\bowtie} Q$ denotes synchronisation between components P and Q, the set $\mathcal{H} \in 2^{\mathcal{A}}$ determines those action types on which the operands are forced to synchronise, with $\underset{*}{\bowtie}$ denoting a synchronisation on all common action types. Note that the synchronization of components on a given action type α represents the participation of the corresponding species in the same reaction (abstracted by the action type α). In the model component $S[x]$, the parameter $x \in \mathbb{R}^+$ represents the initial value. The reader is referred to [21] for further details of the language and its semantics.

In addition to species and model components, a Bio-PEPA system is characterised by a context containing the constant parameters, the functional rates, the locations, the possible events and auxiliary information about the species.

The parameters are defined in the model by means of a set of parameter definitions $\mathcal{K}$. Each parameter is defined by "k_{name} = value unit", where "$k_{name} \notin C$" is the parameter name, "value" denotes a positive real number and the (optional) "unit" denotes the unit associated with the parameter.

In order to collect the information about the dynamics of the system, we associate a functional rate f_α with each action type α. The set of functional rates is denoted $\mathcal{F}_R$. The function f_α can depend on parameters, names of species components and possibly on simulation time and represents the kinetic law of the associated reaction as a mathematical expression. The mathematical expressions are defined in terms of mathematical operators or predefined functions, expressing well-known kinetic laws such as mass-action, Hill kinetics and Michaelis-Menten. In the former case the names of the parameters and the names of the species components involved in the reaction must be given whereas with the predefined kinetic laws the components/species are derived from the context.

Locations represent both biological compartments (such as nucleus, cytoplasm, $\cdots$) and membranes. Membranes represent the boundaries of compartments and may or may not be explicitly included. Every model must have at least one compartment. Each location is described by "L : s *unit*, *kind*", where L is the (unique) location name, "s" expresses the size and can be either a positive real number or a more complex

mathematical expression depending on time t; the (optional) "*unit*" denotes the unit of measure associated with the location size, and "*kind*" $\in \{\mathbf{M}, \mathbf{C}\}$ expresses if it is a membrane or a compartment, respectively. Although the relative position of locations is assumed to be static, their size may change with time by expressing the volume or area as a function of time. In this latter case only specific kinds of analysis are supported (i.e. numerical integration of ODEs).

A key reaction involving location is the translocation of a species S from one location L_i to the location L_j. This is simply abstracted by a reaction $S\,@\,L_i \rightarrow S\,@\,L_j$, where $S\,@\,L_i$ is the reactant and $S\,@\,L_j$ is the product.

Events are constructs that represent changes in the system due to some triggering conditions. This allows biochemical perturbations to the system to be represented, such as the timed introduction of reagents or the modulation of system components by external stimuli. A Bio-PEPA event has the form (*id, trigger, event_assignment, delay)*, where:

- *event_id* is the event identifier,
- *trigger* is a mathematical expression that, when it evaluates to true, makes the event fire. It can be composed of one or more conditions involving the components of the Bio-PEPA model and/or time;
- *event_assignment_list* is a list of changes (assignments) to elements of the system in response to the event;
- *delay* is the length of time between when the event fires and when the event assignments are executed. *delay* is either 0 (*immediate events*) or a positive real value (*delayed events*). In the model we consider in this paper we consider just immediate temporal events, i.e., they are not delayed and their trigger involves time.

Note that events are added to the language as a distinct set of elements and the rest of the syntax is unchanged in order to keep the specification of the model as simple as possible. This approach is particularly useful when the same biochemical system is studied under different experimental regimes as the list of events can be modified without any changes to the rest of the system. Details of analysis supporting events are reported in [31].

Finally, a set $\mathcal{N}$ is defined in order to collect some auxiliary information about the species used in some kinds of analysis supported by Bio-PEPA.

In order to illustrate Bio-PEPA syntax, we show how a simple network can be specified in Bio-PEPA. This network is composed of the following two reactions (in chemical reaction form):

$$S + E \longrightarrow P + E\,; \qquad P \longrightarrow$$

The former is an enzymatic reaction describing the transformation of a substrate S into the product P with the help of the enzyme E and the latter is the degradation of the product P. The kinetic law for the enzymatic reaction is $f_E = \frac{v_M \cdot E \cdot S}{(K_M + S)}$ whereas for the degradation it is $f_{deg} = k \cdot P$. All the species are in the same location L.

The three species can be specified in Bio-PEPA by the following components:

$$S\,@\,L \stackrel{def}{=} (\alpha, 1)\downarrow \qquad P@L \stackrel{def}{=} (\alpha, 1)\uparrow + (\beta, 1)\downarrow \qquad E@L \stackrel{def}{=} (\alpha, 1)\oplus$$

The action type α abstracts the enzymatic reaction and β the degradation. The species S is only involved in the reaction α as a reactant (i.e. it is consumed). The enzyme E is only an activator for α. The product P can take part in both reactions. In the case of α it is a product (i.e. it is created) and in the case of β is a reactant (i.e. it is consumed). In all cases the stoichiometry is one.

The system is described by

$$S@L[x_{S,0}] \underset{\{\alpha\}}{\bowtie} E@L[x_{E,0}] \underset{\{\alpha\}}{\bowtie} P@L[x_{P,0}]$$

where $x_{S,0}$, $x_{E,0}$ and $x_{P,0}$ are the initial values of the three species and the functional rates are $f_\alpha = fMM(v_M, K_M)$, i.e. Michaelis-Menten kinetics with parameters v_M, K_M, and $f_\beta = fMA(k)$.

Bio-PEPA offers a formal intermediate compositional representation of biochemical systems, on which different kinds of analysis can be carried out, through defined mappings into continuous-deterministic and discrete-stochastic mathematical models. The Bio-PEPA language is supported by software tools (for instance the Bio-PEPA Workbench [33]) which automatically process Bio-PEPA models and generate other representations in forms suitable for different kinds of analysis [21,34]. In particular, the generated simulation model can be executed using MATLAB [35] and the Dizzy simulation tool [36], in which both stochastic simulation algorithms and differential equation solvers are implemented. Here we use a version of the Dizzy simulator developed at the University of Edinburgh [25], which extends the original tool with sensitivity analysis techniques and additional simulation methods. Some events, such as, for instance, time-dependent events, can be translated into time-dependent reaction rates in the Dizzy model (defined in terms of the step function *theta*, which is predefined in Dizzy). Most stochastic algorithms and ODE solvers in Dizzy support this function.

5 A Bio-PEPA model for the NF-κB pathway

In the following we illustrate the Bio-PEPA model describing the NF-κB pathway presented in Sect. 2. We show the mapping from each biochemical entity (species, reaction, $\cdots$) to Bio-PEPA. We report just the main ideas of the abstraction, the full model is reported in Appendix A and available from the Bio-PEPA web page [34].

The pathway is characterised by the presence of two compartments (and the transport of some species between them) and by the influence of an external signal. These features can be easily represented in Bio-PEPA using locations and events.

Compartments The *nucleus* and the *cytoplasm* are abstracted by *locations* in Bio-PEPA:

$$\begin{aligned} &\mathit{location\ nuc}\ :\ \mathit{kind} = \mathbf{C},\ \mathit{size} = 3.33 \cdot 10^{-13}\ l; \\ &\mathit{location\ cyt}\ :\ \mathit{kind} = \mathbf{C},\ \mathit{size} = 1.65 \cdot 10^{-12}\ l \end{aligned}$$

They are both of kind **C** (i.e. compartments) and their sizes are as given in [5].

Reactions Each *reaction* is associated with an *action type* and with a *functional rate*. For instance, in the case of degradation of the protein A20 we have the action type *a20_degradation* and the associated functional rate:

$$f_{a20\_degradation} = fMA(c5)$$

where $fMA(r)$ stands for mass-action with rate constant r and $c5$ is the constant degradation rate for $A20$.

Species Each *biochemical species* in the pathway is abstracted by a *species component*, describing its behaviour in terms of the interactions in which it is involved.
For instance the protein $A20$ is represented as:

$$\begin{aligned} A20@cyt \stackrel{def}{=} & (a20\_translation, 1) \uparrow + \\ & (a20\_degradation, 1) \downarrow + \\ & (transformation\_IKKa\_into\_IKKi\_by\_A20, 1) \oplus \end{aligned}$$

This species is in the cytoplasm and it is involved in three interactions: its translation, its degradation and as an activator for the transformation of IKKa into the inactive form IKKi. The abbreviation presented in Section 4 is used. In all the three cases mass-action kinetic laws are considered [5].

Model Component The species and their possible interactions are represented by the *model component*:

$$I\kappa B\alpha\text{–}NF\text{-}\kappa B@cyt[60000] \underset{*}{\bowtie} IKKn@cyt[0] \underset{*}{\bowtie} IKKa@cyt[0] \underset{*}{\bowtie} \cdots A20@cyt[0]$$

where the number between square brackets represents the initial number of molecules of each species.

Signal The *signal* (TNF stimulus) is abstracted in Bio-PEPA by two *time-dependent events*, representing the start and the end of the signal.

$$\begin{gathered} (begin\_signal,\ t = T_1,\ signal = 1,\ 0); \\ (end\_signal,\ t = T_2,\ signal = 0,\ 0) \end{gathered}$$

In our case $T_1 = 3600$ and $T_2 = 7 \cdot 3600 = 25200$ seconds (corresponding to 1 hour and 7 hours, respectively) and both events are immediate.

To simplify the handling of compartments for Bio-PEPA with locations we prefer to have the model expressed in terms of numbers of molecules, rather than concentration, for each species. Otherwise concentrations have to be converted to account for different volumes whenever molecules move between compartments. Considering the number of molecules is thus more convenient. Note that the rates and initial values for species in Lipniacki's model [5] are expressed in concentrations (μM). In order to derive a model in terms of molecule numbers for Bio-PEPA, the continuous concentration values are translated into discrete numbers of molecules and the rates are modified in order to take this transformation into account (see [37] for details).

In our model we have two compartments with different volume sizes, so we have to define two scaling factors for the transformation from concentrations to molecules and use them according to the location of each species and reaction. Specifically, the two scaling factors are:

$$nscale = V_n \cdot N_A \cdot 10^{-6} = 2 \cdot 10^5 \quad \text{molecules}/\mu\text{mol} \cdot \text{L}$$

[5] Note that in the reaction describing the transformation of IKKa into IKKi, the protein A20 remains constant and it is a sort of activator of the reaction. The reaction is described as $IKKa + A20 \rightarrow IKKi + A20$ with kinetic law $k_2 \cdot IKKa \cdot A20$.

$$cscale \;=\; V_c \cdot N_{\mathrm{A}} \cdot 10^{-6} \;\approx\; 10^6 \quad \text{molecules}/\mu\text{mol} \cdot \text{L}$$

where *nscale* and *cscale* are the scaling factors, $N_{\mathrm{A}} = 6.022 \cdot 10^{23}\ \text{mol}^{-1}$ is the Avogadro number and V_n and V_c, the volume sizes, for the nucleus and cytoplasm, respectively. As the concentrations are in terms of μM instead of M we have to multiply by 10^{-6}. As we consider numbers of molecules, we do not need to use the factor k, used to take into account the compartment volumes in the reactions involving species in different compartments when concentrations are used [5].

A list of all parameter values used in our model is reported in Table 1.

Table 1. Parameters and reactions of the model. For more details see [5,6].

Parameter	Value	Unit	Description
nscale	$2 \cdot 10^5$	molecules/μmol $\cdot$ L	scaling factor for the nucleus
cscale	10^6	molecules/μmol $\cdot$ L	scaling factor for the cytoplasm
k	5	-	cytoplasmic to nuclear volume
k_{prod}	0.000025·cscale= 25	molecules $\cdot$ s^{-1}	IKKn production rate
k_{deg}	0.000125	s^{-1}	IKKa, IKKn and IKKi degradation
k_1	0.0025	s^{-1}	IKK activation rate caused by TNF
k_2	0.1/cscale= 10^7	molecules^{-1} s^{-1}	IKK inactivation rate caused by A20
k_3	0.0015	s^{-1}	IKK spontaneous inactivation rate
c_1	$5 \cdot 10^{-7} \cdot k = 2.5 \cdot 10^{-6}$	s^{-1}	A20-inducible mRNA transcription
c_2	0	molecules $\cdot$ s^{-1}	A20-constitutive mRNA transcription
c_3	0.0004	s^{-1}	A20 mRNA degradation
c_4	0.5	s^{-1}	A20 translation rate
c_5	0.0003	s^{-1}	A20 protein degradation
t_1	0.1	s^{-1}	IKKa-IκBα catalysis
t_2	0.1	s^{-1}	IKKa-IκBα-NF-κB catalysis
c_{1_a}	$5 \cdot 10^{-7} \cdot k = 2.5 \cdot 10^{-6}$	s^{-1}	IκBα-inducible mRNA transcription
c_{2_a}	0	molecules $\cdot$ s^{-1}	IκBα-constitutive mRNA transcription
c_{3_a}	0.0004	s^{-1}	IκBα mRNA degradation
c_{4_a}	0.5	s^{-1}	IκBα translation rate
c_{5_a}	0.0001	s^{-1}	spontaneous, free IκBα degradation
c_{6_a}	0.00002	s^{-1}	IκBα degradation (complexed to NF-κB)
a_1	0.5/cscale= $5 \cdot 10^{-7}$	molecules^{-1} s^{-1}	IκBα-NF-κB association
a_2	0.2/cscale= $2 \cdot 10^{-7}$	molecules^{-1} s^{-1}	IKKa-IκBα association
a_3	1/cscale= $1 \cdot 10^{-6}$	molecules^{-1} s^{-1}	IKKa IκBα-NF-κB association
a_{1_n}	0.5/nscale= $2.5 \cdot 10^{-6}$	molecules^{-1} s^{-1}	IκBα-NF-κB association (nucleus)
c_{1_c}	$5 \cdot 10^{-7} \cdot k = 2.5 \cdot 10^{-6}$	s^{-1}	cgen inducible mRNA transcription
c_{2_c}	$0 \cdot k = 0$	molecules $\cdot$ s^{-1}	cgen constitutive mRNA transcription
c_{3_c}	0.0004	s^{-1}	cgen mRNA degradation
i_{1_a}	0.001	s^{-1}	IκBα nuclear import
e_{1_a}	$0.0005 \cdot k = 0.0025$	s^{-1}	IκBα nuclear export
e_{2_a}	$0.01 \cdot k = 0.05$	s^{-1}	IκBα-NF-κB nuclear export
i_1	0.0025	s^{-1}	NF-κB nuclear import

6 Validation and analysis

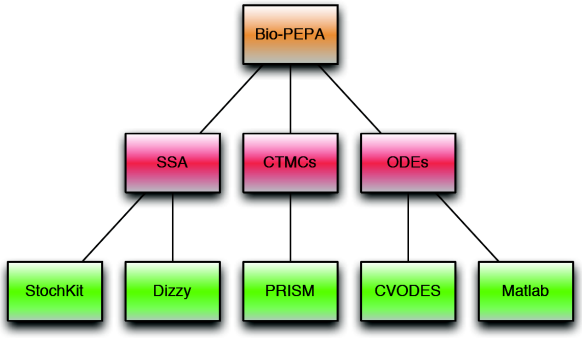

Fig. 2. Schematic view of the analyses available within the Bio-PEPA Workbench, from the Bio-PEPA description of the system (top level) three distinct classes of analysis are accessible (middle level): Stochastic Simulation (SSA), explicit state CTMC techniques such as numerical solution and stochastic model checking, and ordinary differential equations (ODEs); these analyses are supported using existing tools (bottom level).

The Bio-PEPA model of the NF-κB pathway was implemented in the Bio-PEPA Workbench [33] which supports a variety of analyses (see Fig. 2). We consider the ODE MATLAB model to provide validation against the results reported in [5] and the Dizzy model for stochastic simulation and sensitivity analysis. The simulations are carried out using Gillespie's direct method [37]. This choice was made on the basis of the efficiency of that algorithm with respect to our model and the fact that its implementation within Dizzy supports time-dependent events by means of the *theta* function, as discussed at the end of Sect. 4. Nevertheless we verified that the same results are obtained if other stochastic algorithms in Dizzy which support events are used (results not shown). Throughout the presented analyses the unit of time is one second.

6.1 Validation

As explained above, in order to validate our model against the original ODE model [5] and the available experimental data in [23,3] showing the behaviour of a cell population, we consider the ODE MATLAB model obtained from the Bio-PEPA Workbench [33].

In [5], Lipniacki *et al.* take the following approach to time series evolution of species. At time 0 just the complex cytoplasmic IκBα–NF-κB is present and all other species are zero. The simulation is run until the resting cell equilibrium state is reached (100 hours). The simulation is then run for a further seven hours, with the external signal enabled after one hour.

The ODE solver used is the MATLAB *ode23tb* [38], designed for stiff systems. Events are handled "by hand": stopping the simulation when the events occur and resetting the system of ODEs with the initial species values set to the final values obtained from the previous simulation and changing the rates of the two reactions enabled by the external stimulus (i.e. activation of IKK and the transformation of active IKK into inactive IKK by means of A20) in the appropriate way. Specifically, the rate constants of the two reactions are initially zero. At the start of the stimulus at time 1 hour these rates are set to values different from zero and, at the end of the stimulus at time 7 hours, they are reset to zero again.

In Fig. 3 we report the results obtained running our MATLAB ODE model with the approach described above. The figure corresponds to the Fig. 3 in [5], obtained using the ODE MATLAB Lipniacki model. The persistent TNF signal (subfigure [A]) causes a pulse activation of IKK (subfigure [C]). The pulse of active IKK (IKKa) initiates the cascade. First IκBα (subfigure [E]) and the complex IκBα–NF-κB (subfigure [F]) are degraded; the released NF-κB moved in the nucleus and, due to the low quantity of nuclear IκBα (subfigure [G]), it rapidly increases. Nuclear NF-κB (subfigure [H]) upregulates mRNA expression of both IκBα and A20 (subfigures [I] and [H]); the peaks of the associated transcripts are followed by peaks in the corresponding free cytoplasmatic proteins. The new synthesised IκBα binds to NF-κB and leads it out of the nucleus while A20 triggers IKK inactivation. The two negative loops and the movement of species between compartments are fundamental for NF-κB's oscillations and therefore have a large impact of the cell's activities.

The species in our MATLAB model are expressed in terms of number of molecules and, therefore, the species values in Fig. 3 are rescaled in terms of concentration in order to have a more direct comparison with the figures reported in [5]. The results obtained are identical to the ones shown in the paper. This also confirms that our conversion from concentration to number of molecules is appropriate. The same results are obtained when the Dizzy tool is used for the simulation.

6.2 Stochastic simulation: population vs single cell behaviour

In this section we consider the Dizzy model derived from our Bio-PEPA and we perform stochastic simulation in order to investigate the effect of stochasticity on the behaviour of the species in the pathway, in particular of nuclear NF-κB.

While deterministic models are good approximations of real biochemical systems when the number of molecules is sufficiently high, at low copy numbers the effect of random fluctuations becomes significant and so stochasticity needs to be taken into account to obtain a faithful representation of the real biochemical system [39]. This is particularly true when the activation of genes is involved, as generally there are few copies of each gene in the cell. For this reason, we decide to consider stochastic simulation for the following analyses of the model.

Note that a single simulation of the model is a specific realization of the underlying stochastic process and therefore can abstract the behaviour of a single cell. If we assume that the behaviour of the cells are independent (at least with respect to the reactions of the pathway under consideration) the comphrensive behaviour of a population of cells can be captured by the average of a number of simulation runs. Therefore, in the

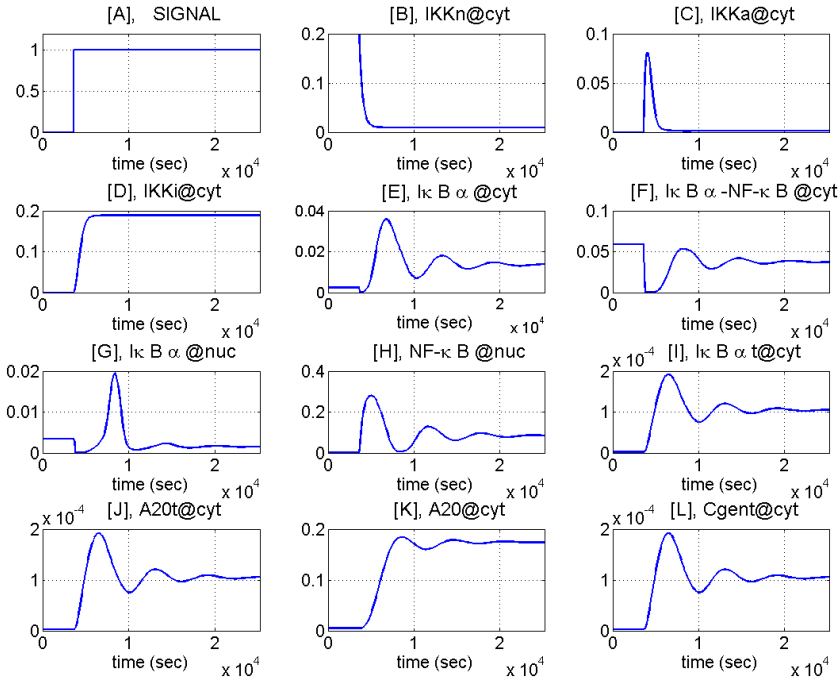

Fig. 3. Validation of the ODE MATLAB model obtained from Bio-PEPA (concentrations). The time (x value) is in seconds and the y values are concentrations (in μM).

following we report the graphs showing the average of several stochastic simulation runs as we want to compare them with the experimental data in [23,3], which refer to comphrensive behaviour of a population of cells. On the other hand, we consider the single simulation runs too, in order to show the possible differences between the behaviour of a single cell and a population of cells.

Particular attention must be given to species that are characterized by oscillatory behaviour, indeed the study of oscillations for NF-κB is fundamental as this mechanism seems to have an essential role on important activities of the cell, such as death and immunity. The oscillations need to be largely synchronous to be observed in a cell population, as, if they are out of phase, in the average they are damped, often to the point of invisibility [40,8]. From ODE models in the literature [3,5] (including the ones derived from Gepasi) we can only observe the average behaviour that, for some experiments, is not representative of the behaviour of a single cell, especially with respect to oscillations.

Fig. 4 reports the average of 100 simulation runs and the standard deviation of the species nuclear NF-κB (top-left), cytoplasmic IκBα (top-right), cytoplasmic IκBα–NF-κB (bottom-left) and IKKa (bottom-right). The average behaviour for these species is

very close to the deterministic solution. This is unsurprising since the amount of these species is indeed quite high. For nuclear NF-κB, IκBα and IκBα−NF-κB we can observe an average oscillatory behaviour, that is less evident after the first oscillations, when the variability between the different runs is greater. If we observe a single simulation run for the species (in Fig. 5 the results for two species are reported), the species present a persistent oscillatory behaviour, at least for the time interval considered. The oscillations in the various simulation runs are not completely in phase and, therefore, in the average they tend to be damped.

These results are in agreement with the literature [8]. Indeed experimental data concerning the pathway suggest a small degree of synchronization amongst the cell population when we have a persistent or long pulse TNF signal, with more evident synchronization in the first oscillations. However, in the case of shorter stimuli, the experimental results show that, under certain assumptions, a strong synchronization and the cell population behaviour is representative of the behaviour of the single cell [40,8].

Similar results, but with smaller standard deviations, are obtained with a larger number of simulation runs (results not presented). In this case, the simulation time increases.

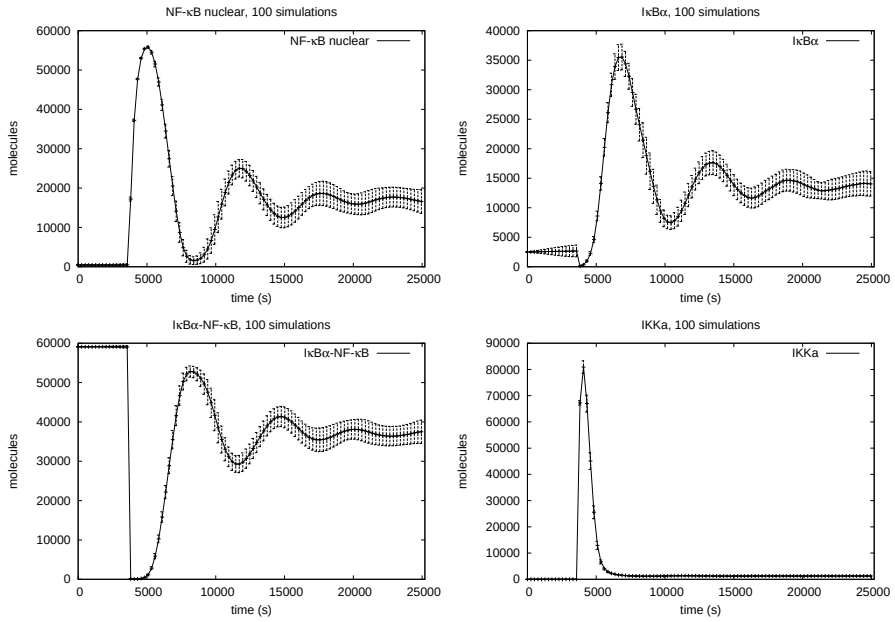

Fig. 4. Stochastic simulation (Gillespie's direct method) for nuclear NF-κB (top-left), cytoplasmic IκBα (top-right), cytoplasmic IκBα−NF-κB (bottom-left) and IKKa (bottom-right). For each of them the average value and the standard deviation are shown. 100 simulation runs are considered. These species correspond to the species in the subgraphs H, E, F and C of Fig. 3, respectively.

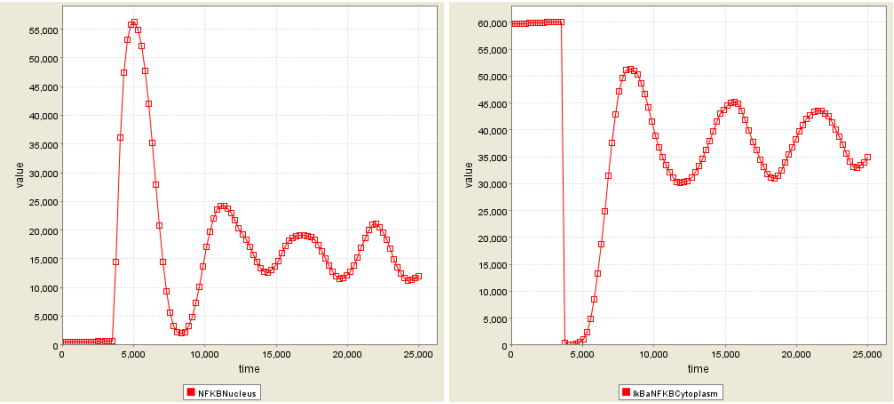

Fig. 5. Single stochastic simulation runs (Gillespie's direct method) for nuclear NF-κB (left) and cytoplasmic IκBα–NF-κB (right). These species correspond to the species in the subgraphs H and F of Fig. 3, and in the graphs top-left and bottom-left in Fig. 4, respectively.

6.3 In silico experiments

We use our model to study the behaviour of the system, particularly nuclear NF-κB, under various assumptions. In the following, we study how the duration of the TNF stimulus affects the behaviour of the nuclear NF-κB and we investigate the regulatory mechanisms, varying the initial total amount of the NF-κB and the inhibitory activity of A20 and IκBα. After that, we explore the behaviour of IκBα mRNA, IKKa and nuclear NF-κB in A20-deficient cells. Finally, we study the effect of varying both A20 and IκBα transcription rates. This last experiment is not present in [5].

The use of stochastic simulation within our analysis means that in addition to the cell population results, previously presented by [5], we can also analyse the behaviour of single cells. In particular, we compare the results obtained from single simulation runs (a single cell) to the average behaviour over many runs (a population of cells). The results of these single runs have yet to be validated against appropriate experiments but the data produced is ideally suited to live cell imaging techniques. This is in contrast to the population case (ODEs and multiple runs of stochastic simulations) which correspond to population approaches such as western blotting. However the results presented here, whilst not yet validated, do show agreement with present knowledge of the system.

Effect of the duration of TNF stimulus on nuclear NF-κB In the previous sections we have assumed that the TNF stimulus is persistent and lasts for 6 out of 7 simulation hours. It is interesting to investigate how a shorter duration stimulus can affect the nuclear NF-κB activity, in particular the oscillatory behaviour. Different durations describe specific kinds of inflammation [3,8].

Fig. 6 shows the average over 100 stochastic runs for IKKa and nuclear NF-κB when the stimulus lasts for 15 minutes (left), 60 minutes (middle) and 6 hours (original

value, right). The three graphs show the same behaviour for both IKKa and nuclear NF-κB for the first three hours. In particular, the pulse of nuclear NF-κB starts after one hour, has a peak at 90 minutes and lasts about one hour. This similarity between the three situations described above may be due to the fact that the pulse of NF-κB is strictly influenced by IKKa and the behaviour of IKKa is the same for all the three cases. Indeed IKK activation seems, at least in part, independent of the duration of the TNF stimulus. This can be explained by the fact that the inactivation rate under the stimulus is very high. Therefore, the inactivation of IKK is very fast and all IKKa is consumed soon after the beggining of the stimulus. However, the duration of the stimulus has an impact on the behaviour of NF-κB after the initial pulse: when the stimulus lasts for 15 minutes or 1 hour nuclear NF-κB drops to a very small amount (2 or 4 molecules, respectively). In contrast, with the longer stimulus we see a pronounced oscillation. The same results are obtained when we consider single simulation runs, as we can see in Fig. 7, where a single run for NF-κB with one-hour stimulus is reported. Therefore we can deduce that shorter stimuli do not induce oscillations for nuclear NF-κB either at the level of a single cell nor the level of a population of cells.

These results are in full agreement with the experimental data shown in [3] for wild-type cell under TNF stimuli of short duration.

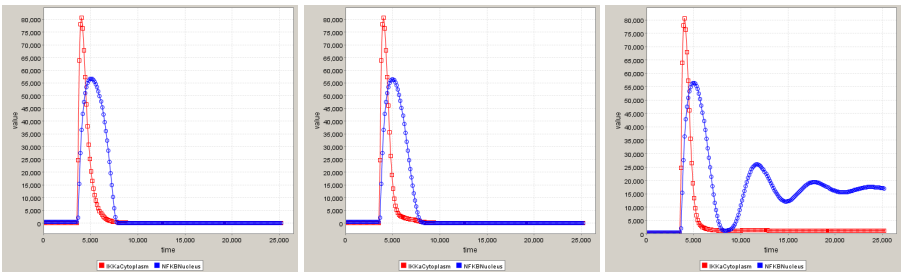

Fig. 6. IKKa (red line) and nuclear NF-κB (blue line) at and after 15 minute-long (left), 60 minute-long TNF stimulation (centre) and at and after 6 hours (right). The graphs show the average amount over 100 simulation runs.

Study of the regulatory mechanisms First, we consider various initial amounts for the total NF-κB. In Fig. 8 we report the average behaviour over 100 runs of the nuclear NF-κB when the initial total NF-κB is the original value, three times the original value and one third of the original value. These values have been chosen to reproduce the experimental data and in silico experiments in [5]. At the level of a population of cells, an increase in the amount of total NF-κB makes the oscillation more pronounced whereas, taken in the average, a decrease smooths out the oscillations.

In Fig. 9 we report single simulation runs (single cells) under the same assumptions. For the larger quantity of initial NF-κB the behaviour at the level of the population of cells and at the level of a single cell is identical. On the other hand, when smaller initial quantities are assumed, in the single cell we can observe an oscillatory trend, even

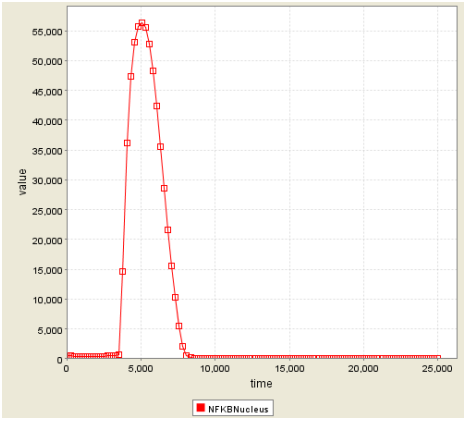

Fig. 7. A single stochastic simulation run for nuclear NF-κB with one-hour stimulus.

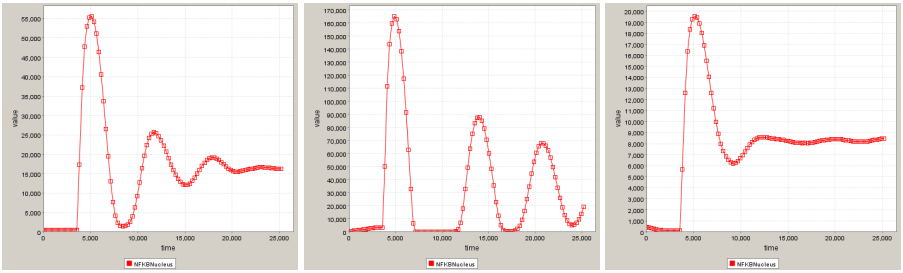

Fig. 8. Nuclear NF-κB when the initial total NF-κB is $6 \cdot 10^4$ (original value, left), is $1.8 \cdot 10^5$ (three times the original value, centre), and $2 \cdot 10^4$ (one third of the original value, right). The graphs show the average amount over 100 simulation runs.

though less regular than before, not observable when the average is considered. Therefore, the initial amount of NF-κB seems essential to obtain an oscillatory behaviour, especially when the population of cells is considered.

In order to see how the system behaves at different inhibitor levels, we elevate the A20 and IκBα mRNA transcription rates (c_1 and c_{1a}, respectively) three-fold with respect to the original value whereas the other parameters remain unchanged. As before, these values have been selected in order to compare our results with the experiments in [5]. The results are reported in Fig. 10. The higher level of A20 mRNA transcription leads to a higher level of the protein A20 and therefore to a lower level of IKKa. The resulting nuclear NF-κB has less damped oscillations. A similar result for nuclear NF-κB is obtained when we elevate the IκBα mRNA transcription rate three-fold with respect to the original value.

These results are in agreement with the ones in [5]. Concerning single simulation runs, also for these experiments we have more persistent and sustained oscillations.

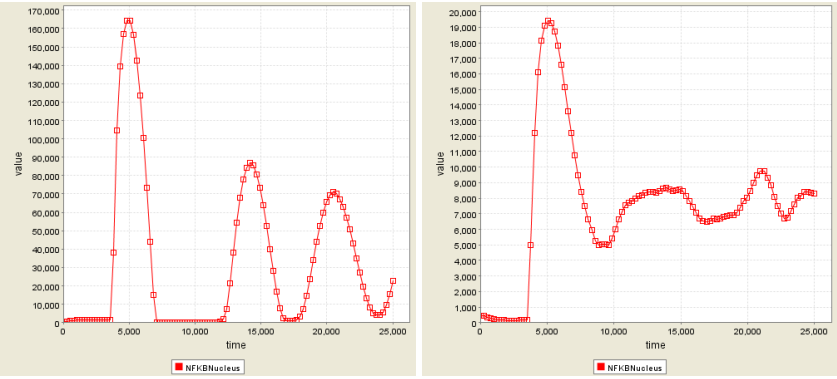

Fig. 9. Single simulation runs for nuclear NF-κB when the initial total NF-κB is $1.8 \cdot 10^5$ (three times the original value, left) and $2 \cdot 10^4$ (one third of the original value, right).

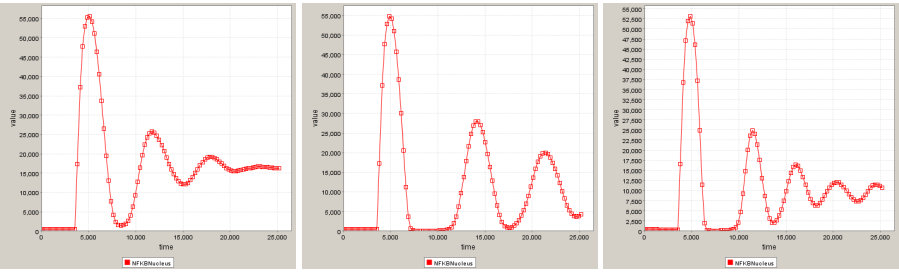

Fig. 10. Nuclear NF-κB when A20 and IκBα mRNA transcription rates have the original values (left), when A20 mRNA transcription rate is three-fold the original value (centre) and when IκBα mRNA transcription rate is three-fold the original value (right). The graphs show the average amount over 100 simulation runs.

A20 deficient cells The protein A20 is strongly NF-κB responsive and has an important inhibitory effect on the activation of NF-κB. The knockout of A20[6] in mouse dramatically alters the cell response to TNF stimulation due to persistent IKK activities and causes A20 deficient mice to die prematurely.

In Fig. 11 we report the stochastic simulation (average of 100 runs) for some species of the pathway. A20 deficient cells are simulated by setting A20 mRNA transcription rate to zero ($c1 = 0$). IKKa presents a level of activation in the tail after the peak. This is different from what happens in wild-type cells (see for instance Fig. 3) and it is mainly due to the fact that A20 is responsible for the inactivation of IKKa. This difference in the IKK activity influences the rest of the pathway response. In particular, this disregulates NF-κB, which then accumulates in the nucleus during TNF stimulation. As a consequence of this, the amount of IκBα mRNA increases and at the end of the stimulation we have twice the quantity with respect to the wild-type cells. These results are in agreement with the published experiments.

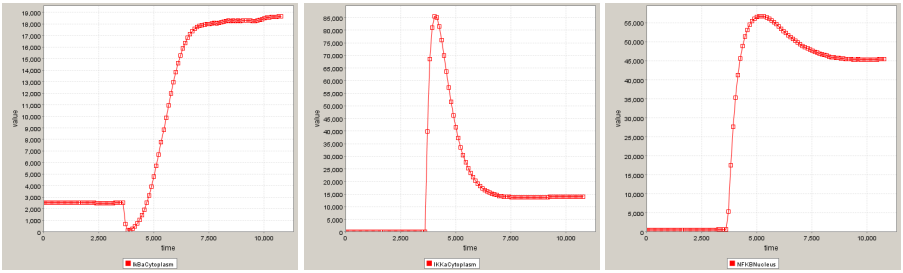

Fig. 11. A20 deficient cells. Stochastic simulation (average of 100 runs) for IκBα mRNA (left), IKKA (centre) and nuclear NF-κB (right).

Fig. 12 reports a single simulation run for nuclear NF-κB. For A20 deficient cells nuclear NF-κB does not have an oscillatory behaviour also at the level of a single cell. This confirms the importance of the A20 feedback loop on oscillations.

Study of varying both c_1 and c_{1a} In addition to the experiments reported above, where we vary each mRNA transcription rate (c_1 and c_{1a}) in isolation, in this work we use our model to investigate how the nuclear NF-κB is affected by varying both c_1 and c_{1a}. Indeed, from previous investigations [5,10], each of these two parameters influences the behaviour of nuclear NF-κB. Thus it is interesting to study their synergistic effect on the system. This experiment is not present in [5]. In Fig. 13 we report the results for some choices of the two parameter values. Specifically, we focus on the following cases: both c_1 and c_{1a} are three-fold the original values (top-left), c_{1a} is three-fold its original value and c_1 is one third of its original value (top-right), c_1 is three-fold its original value and c_{1a} is one third of its original value (bottom-left) and, finally, both c_1 and c_{1a} are one third of the original values (bottom-right). Sustained oscillations (in the

[6] We call such cells A20 deficient.

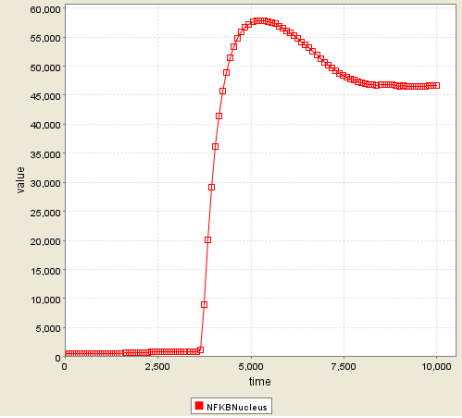

Fig. 12. A20 deficient cells. Single stochastic simulation for nuclear NF-κB.

average) are obtained for the first and the third cases, when c_1 has a higher value. On the other hand, when c_1 has lower and c_{1a} has higher than original values the oscillations are damped and disappear immediately after the first oscillation when both the parameters are low. The effect of varying c_{1a} when c_1 is high is on the amplitude and the period of the oscillations: the higher the c_{1a} value, the greater the amplitude and narrower the oscillations.

As observed for the previous experiments, for single simulation runs the oscillations are more sustained and persistent (see Fig. 14), especially with respect to the smaller values of the parameter c_{1a}. Indeed, when c_{1a} is a third of its original value, the oscillations observable in a single run disappear at the population level (graphs on the right in Fig. 14 and Fig. 13, respectively).

These results have yet to be validated against appropriate experiments but they are in agreement with the present knowledge of the system.

6.4 Sensitivity analysis

In order to improve our understanding of the influence of parameter perturbation on species concentration, we employed *Sensitivity Analysis* (SA) [41]. In general this is performed as follows:

1. An initial set of parameter values (*nominal parameters*) is identified. Usually these are the values that are considered the most likely or are the result of parameter fitting. The analysis takes place in the neighbourhood of this configuration. We will call the model with nominal parameters the *nominal model*.
2. A measure or index of the sensitivity is defined, based on the aspect of the model that is the subject of the analysis. An aspect might be the amount of a species S at a specific time or the amplitude of an oscillation, while the index might be the difference between the amount of S computed by the model with nominal parameters and the amount of S computed by the model with perturbed parameters. This

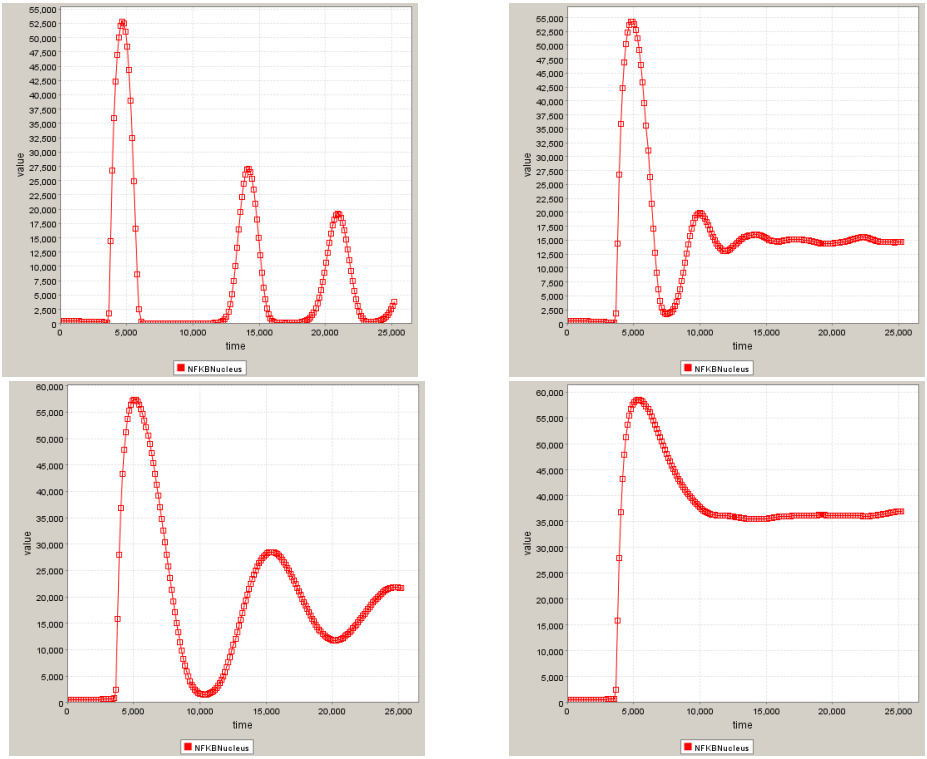

Fig. 13. Nuclear NF-κB under various assumptions about the A20 mRNA transcription rate c_1 and the IκBα mRNA transcription rate c_{1a}. Both c_1 and c_{1a} three-fold the original values (top, left), c_{1a} three-fold the original value and c_1 one third of the original value (top, right), c_1 three-fold the original value and c_{1a} one third of the original value (bottom, left) and both c_1 and c_{1a} one third of the original values (bottom, right). The graphs show the average amount over 100 simulation runs.

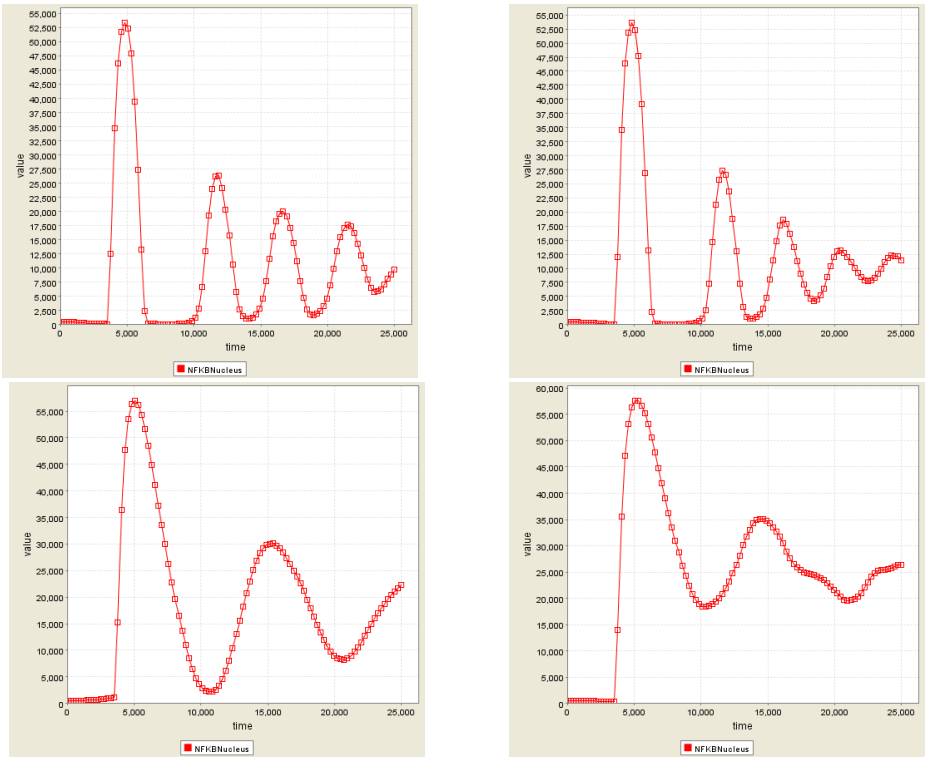

Fig. 14. Single simulation runs for nuclear NF-κB under various assumptions about the A20 mRNA transcription rate c_1 and the IκBα mRNA transcription rate c_{1a}. Both c_1 and c_{1a} three-fold the original values (top, left), c_{1a} three-fold the original value and c_1 one third of the original value (top, right), c_1 three-fold the original value and c_{1a} one third of the original value (bottom, left) and both c_1 and c_{1a} one third of the original values (bottom, right). These graphs correspond to the ones reported in Fig. 13 with a single simulation run instead of the average of 100 simulation runs.

difference is an example of a *sensitivity index* (SI), which quantifies the influence the parameters have on a particular aspect.

3. One or more parameters are modified by a fixed value or a percentage, and SIs are obtained. Reactions governed by parameters with high sensitivity indicate where the model is most susceptible to variations. On the other hand, low sensitivities indicate robustness.

A previous sensitivity analysis of the NF-κB pathway can be found in [9], where an extensive investigation of the influences of the parameters on the amplitude and period of the oscillations of nuclear NF-κB was presented. As described earlier, this analysis was based on a slightly improved version of the system of ODEs presented by Hoffmann *et al.* [3], described in terms of a Gepasi model. The "one-at-a-time" (*OAT*) method, which consists of perturbing only one parameter of the model at a time, was used and nine parameters out of a total of 64 where identified as the most important. This is based on numerical integration of ODEs. Since our stochastic version of the Lipniacki model has a lot of interactions in common with Hoffmann's, these nine parameters find a match in our model, allowing us to compare the results.

In our sensitivity analysis we also used an OAT approach, but three main differences with respect to [9] should be highlighted. First, we perform the analysis on a stochastic model, which permits a more informative analysis than in the case of ODE models, as we shall see shortly. Second, the model we consider includes only one IκB inhibitor (IκBα but not IκBβ and IκBϵ) and the additional A20 inhibitory effect on IKK. Third, we do not use the difference in amplitude or period of oscillations of nuclear NF-κB as a measure of sensitivity. Instead we consider the difference in amount of nuclear NF-κB and cytoplasmic IκBα at regular time intervals.

The parameters subjected to analysis are the already mentioned nine parameters identified in [9] (c_{3_a}, a_3, a_2, i_{1_a}, t_1, t_2, k_{deg}, c_{4_a} and c_{1_a}), with the addition of three parameters related to A20 activity (c_1, c_3 and c_4). For the details of these parameters see Table 1.

As sensitivity indices we use two complementary measures: the *average distance* and the *histogram distance* of stochastic simulations. The former consists of the simple difference in the amount of a species at a selected time, after the perturbation of one parameter. Since we use stochastic simulations, such an amount is the average over a certain number of runs. The latter has been introduced as a sensitivity measure in [24]. The idea behind this distance is that an estimated probability density function (edf) of the amount of a species at a time point can be constructed using a suitable number of simulation runs. Such an edf will have area equal to one, like a probability density function. The overlapping area of two edfs obtained from two different models is then an estimate of the likelihood that the two models will reproduce the same output (see Fig. 15). This measure was introduced originally in [42] to quantify the ability of an approximated version of Gillespie's SSA to replicate the original. We will call the results of the analysis *average sensitivities* if obtained using average distance, or *stochastic sensitivities* if obtained using histogram distance.

The complementarity of the two measures is evident. The average distance does not contain any notion of the distribution of the runs, but it becomes necessary if the distributions of the runs do not overlap at all (in which case the histogram distance

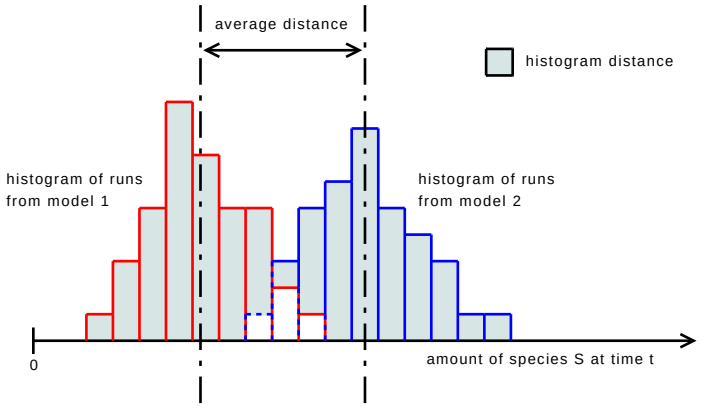

Fig. 15. Average distance and histogram distance between two sets of simulation runs at time t. The former represents the distance of the mean while the latter quantifies the likelihood of run overlapping.

is always two). Therefore, the stochastic OAT sensitivity analysis applies when one is interested in observing the change in the distribution of the amount of a particular species at a given time. This method has been implemented in the version of the Dizzy simulator developed at the University of Edinburgh [25].

In our analysis we observed the influence of a perturbation of 10% of the nominal value of the parameters (as in [9]) on the amount of nuclear NF-κB and cytoplasmic IκBα, using 200 simulation runs. The analysis is performed every 15 minutes, as the sensitivity indices might be time dependent. Results have been grouped in four heat diagrams, shown in Figure 16, with average distance at the top and histogram distance below. Each diagram can be read in two ways: horizontally, for the sensitivity of a parameter through time and vertically, to compare sensitivity of the parameters at a specific time. A scale of colours is used to represent sensitivity indices, from dark red (low SI), passing through yellow, to white (high SI).

The sensitivity analysis gives rise to the following observations:

- In contrast with [9], four of the nine originally identified parameters, a_2, t_1, t_2 and k_{deg}, present a low sensitivity and are clearly the least influential as they show very little variation through time with respect to both nuclear NF-κB and cytoplasmic IκBα. This is probably due to the introduction of the A20 feedback loop. High sensitivity of the three parameters connected with A20 activity sustains this hypothesis, suggesting a key role for this molecule.
- The stochastic sensitivities reveal that with a 10% perturbation of the considered rates there is still significant overlap in the distributions of values obtained from the simulations. Moreover, the trend of the stochastic sensitivities seems to agree with the average sensitivities, suggesting a distribution of the runs around the mean.
- c_{1a} and c_{4a} (bottom two rows of the heat diagrams) exhibit a similar trend, with high sensitivities when nuclear NF-κB peaks (compare with Fig. 4), suggesting that

the rates of transcription and translation of IκBα play a major role in the determination of peak intensity. In particular, the sensitivities of c_{1a} and c_{4a} oscillate in time, suggesting agreement with the nominal model (low sensitivity) alternated to disagreement (high sensitivity) at the times when NF-κB peaks. Moreover, this trend is complementary to the one of c_{3a}, that presents low sensitivity when nuclear NF-κB peaks;

- Similarly, c_1 and c_4, rates of A20 transcription and translation, share a similar trend, and this appears to be dependent on the target species.
- The sensitivity of i_{1a}, the rate of IκBα nuclear import, shows a very clear oscillatory trend only when targeting cytoplasmic IκBα. In particular, a perturbation of i_{1a} yields major changes in the cytoplasmic IκBα peaks.

7 Conclusions

In this work we have presented a Bio-PEPA model for the NF-κB signalling pathway. This is the first algorithmic model [18] of the pathway, which captures how and why the system changes state: a much more mechanistic account of the pathway than that given by previous representations as systems of ODEs, or even as chemical reactions since the role of compartments and events are fully captured in the model. We have studied it using a selection of the analysis techniques supported by the language and implemented in the Bio-PEPA Workbench [43]. With our model we were able to describe in detail important features of the system, such as compartments and the activation of NF-κB by an external stimulus, using a recent extension of the Bio-PEPA language which incorporates explicit representation of locations [32] and time-dependent events [31].

A main feature of Bio-PEPA is that it is a intermediate formal representation of biochemical systems, on which various kinds of analysis can be performed. The access to a variety of analysis techniques can foster a better understanding of the behaviour of the system, and help to discover errors due to the use of a particular solver/simulator [44]. Furthermore, the modeller can select the approach that is most appropriate for specific model under study. Here we focused on deterministic simulation, stochastic simulation and sensitivity analysis.

The system of ODEs corresponding to the Bio-PEPA model was used to validate the system against the experimental data from the literature [23,3], which is based on a cell population. Subsequently, we performed stochastic simulation and we considered both the average of several runs, in order to validate the model against the well-known population behaviour, and single runs, to analyse the behaviour of a single cell. Even though the average behaviour is in agreement with the experimental data and the results obtained from ODEs, some species are characterized by variability between the various runs. In particular, the oscillations of nuclear NF-κB present in the single cells (runs) under various assumptions disappear when the average is considered. This behaviour is confirmed by recent experiments showing that NF-κB under persistent stimulation has an oscillatory behaviour at the level of the single cell, but, as the oscillations are not completely synchronized across a cell population, the oscillations are damped when the average behaviour is considered. Note that most of the models in the literature are

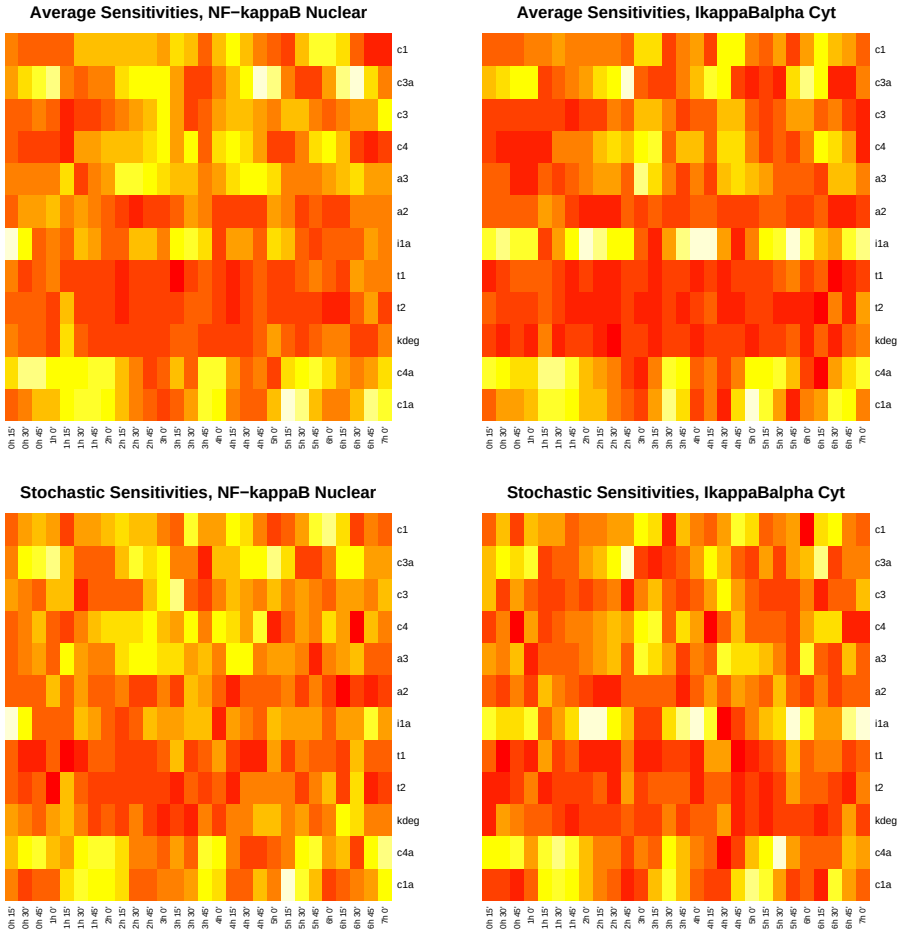

Fig. 16. Heat diagrams of the sensitivity analysis performed on selected parameters of the NF-κB model. Red colours indicate low sensitivity, while yellow or white indicate high sensitivity. The two diagrams above, refer to sensitivity indexes computed using average of simulations, while the two diagrams at the bottom refer to the stochastic approach, that considers a density distance.

described in terms of systems of ODEs [3,5] or as hybrid systems [6] or Gepasi models [9] and therefore the focus is generally on the average behaviour of cells. The single cell approach facilitated by the stochastic simulation emulates the data generated from live cell imaging techniques, as opposed to those obtained from population studies such as western blotting [45,46].

In addition to reproducing and discussing a number of in-silico experiments, sensitivity analysis was applied to investigate the most influential parameters of the model. Our results are complementary to the previous work [9]. Our model has some significant differences, such as the inclusion of a negative feedback loop, and is analysed stochastically. As a result we observe some interesting new phenomena. Moreover, we investigated model properties from a different point of view, using alternative sensitivity measures with respect to [9]. In the presented analysis, we considered a local approach, i.e. one focused around a specific point in the parameter space. This can be informative, giving an idea of the impact of parameter changes on the behaviour of the system. Moreover, this strongly suggested that parameters have specific dependences at particular points in the time evolution of the pathway. This prediction would be amenable to experimental test by single cell imaging methods. In the future, we plan to apply global methods in order to explore the full parameter space (or a meaningful subset of it) and to quantify the relationships between different parameters.

Another route to analysis of a Bio-PEPA model is via the mapping to continuous time Markov chains (CTMC). In particular, it is possible to derive a PRISM [47,48] model in order to verify some properties expressed as a logical formula, by model checking. We have yet to explore the possibilities offered by this route for our NF-κB model. Two main challenges for the use of model checking with this model are the dimension of the state space (it is extremely large) and the presence of temporal events. For the former, one possibility is to apply an abstract-view for the CTMC in terms of concentration levels [49], which can substantially reduce the state space.

A final biological point is that the growing body of evidence from modelling biological pathways is that they exhibit parameter dependences that characterise the structure of the pathway. Most exhibit a small number of vulnerable nodes and are insensitive to perturbation of many values. This is important for considering the development of therapeutic interventions and may well reflect the evolution of the pathway via quantitative changes in key interactions [50].

Acknowledgements

This work was completed while Federica Ciocchetta was a research fellow at the University of Edinburgh, supported by the EPSRC grant EP/c54370x/01. Andrea Degasperi is supported by a University of Glasgow Lord Kelvin/Adam Smith scholarship. Jane Hillston is supported by the EPSRC ARF EP/c543696/01. John K. Heath is supported by Cancer Research UK.

References

1. Cheong, R., Levchenko, A.: Wires in the soup: quantitative models of cell signalling. Trends in Cell Biology **18** (2008) 112–118

2. Hayden, M., Ghosh, S.: Shared principle in NF-κB Signalling. Cell **132** (2008) 344–362
3. Hoffmann, A., Levchenko, A., Scott, M., Baltimore, D.: The IκB–NF-κB Signaling Module: Temporal Control and Selective Gene Activation. Science **298** (2002) 1241–45
4. Cho, K.H., Shin, S.Y., Lee, H.W., et al.: Investigations Into the Analysis and the Modeling of the TNF-mediated NFκB Signaling Pathway. Genome Res. **13** (2003) 2413–2422
5. Lipniacki, T., Paszek, P., Brasier, A., Luxon, B., Kimmel, M.: Mathematical model of NF-κB regulatory module. Journal of Theoretical Biology **228** (2004) 195–215
6. Lipniacki, T., Paszek, P., Brasier, A., Luxon, B., Kimmel, M.: Stochastic Regulation in Early Immune Response. Biophysical Journal **90** (2006) 725–742
7. Nelson, D., Ihekwaba, A., Elliott, M., Johnson, J., Gibney, C., Foreman, B., Nelson, G., See, V., Horton, C., Spiller, D.G., Edwards, S., McDowell, H., Unitt, J.F., Sullivan, E., Grimley, R., Benson, N., Broomhead, D.S., Kell, D., White, M.: Oscillations in NF-κB Signalling Control the Dynamics of Gene Expression. Science **306** (2004) 704–708
8. Ashall, L., Horton, C., Nelson, D., Paszek, P., Harper, C., Sillitoe, K., Ryan, S., Spiller, D.G., J.U., Broomhead, D., Kell, D., Rand, A., Sée, V., White, M.: Pulsatile Stimulation Determines Timing and Specificity of NF-κB–Dependent Transcription. Science **324** (2009) 242–246
9. Ihekwaba, A., Broomhead, D., Grimley, R., Kell, D.: Sensitivity analysis of parameters controlling oscillatory signalling in the NF-κB pathway: the roles of IKK and IκBα. Systems Biology **1** (2004) 93–103
10. Ihekwaba, A., Broomhead, D., Grimley, R., Kell, D.: Synergistic control of oscillations in the NF-κB signalling pathway. IEE Proc.-Syst. biol. **152** (2005) 153–160
11. Mendes, P.: GEPASI: a software package for modelling the dynamics, steady states and control of biochemical and other systems. Comput. Appl. Biosci. **9** (1993) 563–571 gepasi.
12. Priami, C., Regev, A., Silverman, W., Shapiro, E.: Application of a stochastic name-passing calculus to representation and simulation of molecular processes. Information Processing Letters **80** (2001) 25–31
13. Curti, M., Degano, P., Priami, C., Baldari, C.: Modelling biochemical pathways through enhanced π-calculus. Theoretical Computer Science **325** (2004) 111–140
14. Regev, A., Panina, E., Silverman, W., Cardelli, L., Shapiro, E.: BioAmbients: an Abstraction for Biological Compartments. Theoretical Computer Science **325** (2004) 141–167
15. Priami, C., Quaglia, P.: Beta binders for biological interactions. In: Proceedings of Computational Methods in Systems Biology (CMSB'04). Volume 3082 of LNCS. (2005) 20–33
16. Calder, M., Gilmore, S., Hillston, J.: Modelling the Influence of RKIP on the ERK Signalling Pathway Using the Stochastic Process Algebra PEPA. Transactions on Computational Systems Biology **7** (2006) 1–23
17. John, M., Lhoussaine, C., Niehren, J., Uhrmacher, A.: The Attributed Pi-Calculus with Priorities. Transactions on Computational and Systems Biology (to appear)
18. Priami, C.: Algorithmic systems biology. Communications of the ACM **52** (2009)
19. Fisher, J., Henzinger, T.: Executable cell biology. Nature Biotechnology **25** (2007) 1239–1249
20. Ciocchetta, F., Hillston, J.: Bio-PEPA: a framework for the modelling and analysis of biological systems. Technical report, School of Informatics University of Edinburgh Technical Report EDI-INF-RR-1231 (2008)
21. Ciocchetta, F., Hillston, J.: Bio-PEPA: a Framework for the Modelling and Analysis of Biochemical Networks. Theoretical Computer Science **410** (2009) 3065–3084
22. Carlotti, F., Dower, S., Qwarnstrom, E.: Dynamic shuttling of nuclear factor kappa B between the nucleus and cytoplasm as a consequence of inhibitor dissociation. J. Biol. Chem. **275** (2000) 41028–41034

23. Lee, E., Boone, D., Chai, S., Libby, S., Chien, M., Lodolce, J., Ma, A.: Failure to regulate TNF-induced NF-κB and cell death responses in A20-deficient mice. Science **289** (2000) 2350–2354
24. Degasperi, A., Gilmore, S.: Sensitivity Analysis of Stochastic Models of Bistable Biochemical Reactions. In: Formal Methods for Computational Systems Biology. Volume 5016 of LNCS. Springer-Verlag (2008) 1–20
25. Dizzy Edinburgh version: http://homepages.inf.ed.ac.uk/stg/software/Dizzy/ (2009)
26. Ciocchetta, F., Degasperi, A., Heath, J., Hillston, J.: Modelling and analysis of the NF-κB pathway in Bio-PEPA. In Breitling, R., Gilbert, D.R., Heiner, M., Priami, C., eds.: Formal Methods in Molecular Biology. Number 09091 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany (2009)
27. Gerondakis, S., Grossmann, M., Nakamura, Y., Pohl, T., Grumont, R.: Genetic approaches in mice to understand Rel/NF-κB and IκB function: transgenics and knockouts. Oncogene **18** (1999) 6888–6895
28. Hillston, J., Duguid, A.: Deriving Differential Equations from Process Algebra Models in Reagent-Centric Style. In: Algorithmic Bioprocesses. Natural Computing Series, LNCS (2009)
29. Larcher, R., Ihekwaba, A., Priami, C.: A BetaBW model for the NF-κB pathway. Technical Report TR 25/2007, The Microsoft Research-University of Trento Centre for Computational and Systems Biology (2007)
30. Hillston, J.: A Compositional Approach to Performance Modelling. Cambridge University Press (1996)
31. Ciocchetta, F.: Bio-PEPA with events. Transactions on Computational Systems Biology **XI** (2009) 45–68
32. Ciocchetta, F., Guerriero, M.: Modelling Biological Compartments in Bio-PEPA. In: Proc. of MeCBIC 2008. Volume 227 of ENTCS. (2009) 77–95
33. The Bio-PEPA Workbench: http://www.dcs.ed.ac.uk/home/stg/software/biopepa/about.html (2009)
34. Bio-PEPA Home Page: http://www.biopepa.org/ (2008)
35. MATLAB Home Page: http://www.mathworks.com/products/matlab/ (2009)
36. Dizzy Home Page: http://magnet.systemsbiology.net/software/Dizzy (2008)
37. Gillespie, D.: Exact stochastic simulation of coupled chemical reactions. J Phys Chem **81** (1977) 2340–2361
38. Bank, R., et al.: Transient simulation of silicon devices and circuits. IEEE Transactions on Electron Devices **32** (1985) 1992–2007
39. McAdams, H., Arkin, A.: Stochastic mechanisms in gene expression. Proc Natl Acad Sci USA **94** (1997) 814–9
40. Ihekwaba, A., Wilkinson, S., Waithe, D., Broomhead, D., Li, P., Grimley, R., Benson, N.: Bridging the gap between in silico and cell-based analysis of the nuclear factor κB signalling pathway by in vitro studies of IKK2. FEBS Journal **274** (2007) 1678–1690
41. Saltelli, A., Chan, K., Scott, E.: Sensitivity Analysis. Wiley (2000)
42. Cao, Y., Petzold, L.: Accuracy limitations and the measurements of errors in the stochastic simulation of chemically reacting systems. J. Comput. Phys. **212** (2006) 6–24
43. Ciocchetta, F., Duguid, A., Gilmore, S., Guerriero, M., Hillston, J.: The Bio-PEPA Tool Suite. In: Proceedings of the 6th International Conference on Quantitative Evaluation of SysTems (QEST 2009), Budapest, Hungary (2009) 309–310
44. Calder, M., Duguid, A., Gilmore, S., Hillston, J.: Stronger computational modelling of signalling pathways using both continuous and discrete-state methods. In: Proceedins of Computational Methods in Systems Bioogy (CMSB'06). Volume 4210 of Lecture Notes in Computer Science., Springer (2006) 63–77

45. Ankers, J., Spiller, D., White, M., Harper, C.: Spatio-temporal protein dynamics in single living cells. Curr. Opin. Biotechnology **19** (2008) 375–380

46. Sillitoe, K., Horton, C., Spiller, D., White, M.: Single-cell time-lapse imaging of the dynamic control of nfκb signalling. Biochem. Soc. Transactions **35** (2007) 263–266

47. Kwiatkowska, M., Norman, G., Parker, D.: Prism: Probabilistic model checking for performance and reliability analysis. ACM SIGMETRICS Performance Evaluation Review (2009)

48. PRISM Home Page: http://www.prismmodelchecker.org (2009)

49. Ciocchetta, F., Degasperi, A., Hillston, J., Calder, M.: Some Investigations Concerning the CTMC and the ODE Model Derived from Bio-PEPA. In: Proc. of FBTC 2008. Volume 209 of ENTCS. (2009) 145–163

50. Gerhart, J., Kirshner, M.: The theory of facilitated variation. PNAS (2007)

A The Bio-PEPA Model for the NF-κB pathway

In this Appendix we report the full Bio-PEPA model of the NF-κB pathway studied in this paper. First, the set of locations is considered. Then the set of functional rates and the set of parameters are reported. The name of each action type describes the function of the associated reaction. The notation $fMA(r)$ indicates that the kinetic law is mass-action with constant rate r. After that, there is the definition of species components and of the model component. Finally, the events describing the TNF stimulus are defined. The effect of events is to reset the value of the parameter signal, from 0 (signal inactive) to 1 (signal active) and vicerversa. These events enable /unable two reactions, the transformation of IKKn into IKKa and the transformation of IKKa into IKKi with the support of A20, represented by the action types activation_IKKn and transformation_IKKa_into_IKKi_by_A20, respectively (the corresponding kinetic laws depend on the parameter signal and are 0 if the signal is 0). Here we do not report the set $\mathcal{N}$ with auxiliary information for species as this information is not considered in our study. Note that species and parameters are given in terms of number of molecules.

In order to derive a model in terms of number of molecules from the model in terms of concentration [5], all the initial species concentrations are rescaled by the factors $nscale = 2 \cdot 10^5$ (species in the nucleus) and $cscale = 10^6$ (species in the cytoplasm) and the parameters are modified accordingly. The initial concentration of the cytoplasmic complex IκBα−NF-κB is therefore $0.06 \cdot cscale = 60000$ molecules and all the other species are zero.

$$location\ nuc\ :\ kind = \mathbf{C},\ size = 3.33 \cdot 10^{-13}\ l;$$
$$location\ cyt\ :\ kind = \mathbf{C},\ size = 1.65 \cdot 10^{-12}\ l$$

$activation\_IKKn = [fMA(signal \cdot k_1)]$;
$transformation\_IKKa\_into\_IKKi\_by\_A20 = [fMA(signal \cdot k_2)]$;
$production\_IKKn = [k_{prod}]$;
$transformation\_IKKa\_into\_IKKi = [fMA(k_3)]$;
$degradation\_IKKn = [fMA(k_{deg})]$;
$a20t\_transcription\_by\_NFkBn = [fMA(c_1)]$;
$degradation\_IKKa = [fMA(k_{deg})]$;
$cgent\_transcription\_by\_NFkBn = [fMA(c_{1c})]$;
$degradation\_IKKi = [fMA(k_{deg})]$;
$association\_IkBa\_IKKa = [fMA(a_2)]$;
$dissociation\_IkBaIKKa = [fMA(t_1)]$;
$a20\_translation = [fMA(c_4)]$;
$association\_IKKa\_IkBaNFkB = [fMA(a_3)]$;
$a20\_degradation = [fMA(c_5)]$;
$dissociation\_IKKaIkBaNFkB = [fMA(t_2)]$;
$a20t\_transcription = [c_2]$;
$association\_IkBa\_NFkB = [fMA(a_1)]$;
$a20t\_degradation = [fMA(c_3)]$;
$association\_IkBa\_NFkBn = [fMA(a_{1n})]$;
$cgent\_transcription = [c_{2c}]$;
$transport\_IkBaNFkBn\_nucl\_cyt = [fMA(e_{2a})]$;
$cgent\_degradation = [fMA(c_{3c})]$;
$ikBat\_transcription\_by\_NFkBn = [fMA(c_{1a})]$;
$dissociation\_IKKaIkBaNFkB = [fMA(c_{6a})]$;
$transport\_IkBa\_cyt\_nucl = [fMA(i_{1a})]$;
$transport\_IkBa\_nucl\_cyt = [fMA(e_{1a})]$;
$transport\_NFkB\_cyt\_nucl = [fMA(i_1)]$;
$ikBa\_translation = [fMA(c_{4a})]$;
$ikBa\_degradation = [fMA(c_{5a})]$;
$ikBat\_transcription = [c_{2a}]$;
$ikBat\_degradation = [fMA(c_{3a})]$

$$
\begin{aligned}
IKKn@cyt &\stackrel{def}{=} (production\_IKKn,1)\uparrow \; + \; (degradation\_IKKn,1)\downarrow + \\
&\quad (activation\_IKKn,1)\downarrow \\
IKKa@cyt &\stackrel{def}{=} (activation\_IKKn,1)\uparrow \; + \; (transformation\_IKKa\_into\_IKKi,1)\downarrow + \\
&\quad (transformation\_IKKa\_into\_IKKi\_by\_A20,1)\downarrow + \\
&\quad (degradation\_IKKa,1)\downarrow + \\
&\quad (association\_IkBa\_IKKa,1)\downarrow + \\
&\quad (dissociation\_IkBaIKKa,1)\uparrow + \\
&\quad (association\_IKKa\_IkBaNFkB,1)\downarrow + \\
&\quad (dissociation\_IKKaIkBaNFkB,1)\uparrow \\
IKKi@cyt &\stackrel{def}{=} (transformation\_IKKa\_into\_IKKi,1)\uparrow + \\
&\quad (transformation\_IKKa\_into\_IKKi\_byA20,1)\uparrow + \\
&\quad (degradation\_IKKi,1)\downarrow \\
A20@cyt &\stackrel{def}{=} (transformation\_IKKa\_into\_IKKi\_by\_A20,1) \odot + \\
&\quad (a20\_translation,1)\uparrow \; + \; (a20\_degradation,1)\downarrow \\
I\kappa B\alpha@cyt &\stackrel{def}{=} (IkBa\_translation,1)\uparrow \; + \; (IkBa\_degradation,1)\downarrow + \\
&\quad (association\_IkBa\_IKKa,1)\downarrow \; + \; (association\_IkBa\_NFkB,1)\downarrow + \\
&\quad (transport\_IkBa\_cyt\_nucl,1)\downarrow \; + \; (transportIkBa\_nucl\_cyt,1)\uparrow \\
NF\text{-}\kappa B@cyt &\stackrel{def}{=} (dissociation\_IkBaNFkB,1)\uparrow \; + \; (association\_IkBa\_NFkB,1)\downarrow + \\
&\quad (dissociation\_IKKaIkBaNFkB,1)\uparrow + \\
&\quad (transport\_NFKB\_cyt\_nucl,1)\downarrow \\
complex1 &\stackrel{def}{=} (association\_IkBa\_NFkB,1)\uparrow \; + \; (dissociation\_IkBaNFkB,1)\downarrow + \\
&\quad (association\_IKKa\_IkBaNFkB,1)\downarrow + \\
&\quad (transport\_IkBaNFkBn\_nucl\_cyt,1)\uparrow \\
complex2 &\stackrel{def}{=} (association\_IkBa\_IKKa,1)\uparrow \; + \; (dissociation\_IkBaIKKa,1)\downarrow \\
complex3 &\stackrel{def}{=} (association\_IKKa\_IkBaNFkB,1)\uparrow + \\
&\quad (dissociation\_IKKaIkBaNFkB,1)\downarrow \\
I\kappa B\alpha@nuc &\stackrel{def}{=} (association\_IkBan\_NFkBn,1)\downarrow + \\
&\quad (transport\_IkBa\_IkBa\_cyt\_nucl,1)\uparrow \; + \; (transport\_IkBa\_nucl\_cyt,1)\downarrow \\
NF\text{-}\kappa B@nuc &\stackrel{def}{=} (a20\_transcription\_by\_NFkBn,1) \odot + \\
&\quad (cgent\_transcription\_by\_NFkBn,1) \odot + \\
&\quad (ikBat\_transcription\_by\_NFkBn,1) \odot + \\
&\quad (association\_IkBa\_NFkBn,1)\downarrow + \\
&\quad (transport\_NFkB\_cyt\_nucl,1)\uparrow \\
complex4 &\stackrel{def}{=} (association\_IkBa\_NFkBn,1)\uparrow + \\
&\quad (transport\_IkBaNFKBn\_nucl\_cyt,1)\downarrow \\
A20t@cyt &\stackrel{def}{=} (a20\_translation,1) \odot + (a20t\_transcription,1)\uparrow + \\
&\quad (a20t\_transcription\_by\_NFkBn,1)\uparrow \; + \; (a20t\_degradation,1)\downarrow \\
I\kappa B\alpha t@cyt &\stackrel{def}{=} (ikBat\_transcription,1)\uparrow \; + \; (ikBat\_transcription\_by\_NFkBn,1)\uparrow + \\
&\quad (ikBat\_degradation,1)\downarrow \; + \; (ikBa\_translation,1)\odot \\
Cgent@cyt &\stackrel{def}{=} (gent\_transcription,1)\uparrow \; + \; (cgent\_transcription\_by\_NFkBn,1)\downarrow + \\
&\quad (cgent\_degradation,1)\downarrow
\end{aligned}
$$

where the names $complex1$, $complex2$, $complex3$ and $complex4$ stand for $I\kappa B\alpha$–NF-$\kappa B@cyt$, $IKKa$–$I\kappa B\alpha@cyt$, $IKKa$–$I\kappa B\alpha$–NF-$\kappa B@cyt$ and $I\kappa B\alpha$–NF-$\kappa B@nuc$, respectively. The species components $A20t$, $I\kappa B\alpha t$ and $cgent$ are the mRNA transcripts of the proteins $A20$, $I\kappa B\alpha$ and $cgen$ and are assumed in the cytoplasm as in the original model.

$$IKKn@cyt[0] \underset{*}{\bowtie} IKKa@cyt[0] \underset{*}{\bowtie} IKKi@cyt[0] \underset{*}{\bowtie} IKKa\text{–}I\kappa B\alpha@cyt[0] \underset{*}{\bowtie}$$
$$A20@cyt[0] \underset{*}{\bowtie} I\kappa B\alpha@cyt[0] \underset{*}{\bowtie} NF\text{-}\kappa B@cyt[0] \underset{*}{\bowtie}$$
$$I\kappa B\alpha\text{–}NF\text{-}\kappa B@cyt[60000] \underset{*}{\bowtie} IKKa\text{–}I\kappa B\alpha\text{–}NF\text{-}\kappa B@cyt[0] \underset{*}{\bowtie}$$
$$I\kappa B\alpha@nuc[0] \underset{*}{\bowtie} NF\text{-}\kappa B@nuc[0] \underset{*}{\bowtie} I\kappa B\alpha\text{–}NF\text{-}\kappa B@nuc[0] \underset{*}{\bowtie}$$
$$A20t@cyt[0] \underset{*}{\bowtie} I\kappa B\alpha t@cyt[0] \underset{*}{\bowtie} Cgent@cyt[0]$$

$$Events = [(begin\_signal,\ time = T_1,\ signal = 1,\ 0);$$
$$(end\_signal,\ time = T_2,\ signal = 0,\ 0)]$$

In our case $T_1 = 3600$ and $T_2 = 25200$ (time expressed in *seconds*).