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

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

[1] Laboratory for Foundations of Computer Science, University of Edinburgh, EH8 9AB, UK
fciocche@inf.ed.ac.uk
[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 Bio-PEPA events.
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 in the literature and to verify some properties, such as the impact of the stimulus duration and of the NF-κB initial amount on the behaviour of some species. Finally, sensitivity analysis is considered to investigate the most influential parameters of the model.

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 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]. Most of them are defined in terms of Ordinary Differential Equations (ODEs) and the validation and analysis are based on numerical integration of these ODE systems.

Recently there have been various applications of process algebras for the modelling and analysis of biochemical networks [7,8,9,10,11]. These formalisms, originally defined in computer science for the specification and study of complex systems, are also

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

useful in the context of biochemical systems. They offer a formal representation of the model on which various kinds of analysis can be performed, from stochastic simulation to ODE-based analyses, to verification of properties by means of model checking. Furthermore, they support a compositional approach for the construction of the model.

In this work we consider the process algebra Bio-PEPA [12,13]. This language has some features that are useful in describing biochemical networks: it permits the representation of generic kinetic laws by means of functional rates, it supports the explicit definition of stoichiometry, the role of the species in each reaction and locations, and it is enriched with events, constructs able to represent some changes in the system due to trigger conditions.

We focus on a detailed model describing the NF-κB pathway [5] and we represent it in Bio-PEPA. The aim of this work is twofold. First, we would like to demonstrate the power of Bio-PEPA as a modelling language for biochemical networks. In particular, we show how to abstract in Bio-PEPA some features of these networks, such as compartments and the presence of external stimuli which cause the activation of some reactions. These are described by locations and temporal events, respectively. Second, we would like to use some of the analysis techniques supported by Bio-PEPA in order to extend the existing analysis of the model, essentially based on ODE numerical integration, and therefore possibly understand more properties of the pathway. Specifically, we consider stochastic simulation to verify the impact of fluctuations on the behaviour of some species of interest and we use sensitivity analysis to isolate the most influential parameters of the model. Sensitivity analysis is applied to the stochastic version of our model using a novel algorithm, based on the definition of histogram distance over the simulation runs [14], implemented in the version of the Dizzy simulator developed at the University of Edinburgh [15].

For these purposes, we follow the following approach. First we translate the original model in [5] into a Bio-PEPA model. Our model is then validated against the original model, which was proved to be in agreement with the available experimental data [5]. After that we define a stochastic version of the model; it is compared with the ODE model and is used to verify some properties, such as the impact of stimulus duration and the initial amount of NF-κB on the behaviour of some species. Sensitivity analysis is finally applied to the stochastic model.

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. 4, respectively. Bio-PEPA is introduced in Sect. 3. In Sect. 5 the validation of the our model and some analysis results are presented. Finally, in Sect. 6 we report some concluding remarks.

2 The NF-κB pathway

The study of NF-κB is of intense interest as it regulates numerous genes important for pathogen or cytokine inflammation, immune response, cell proliferation and survival [1,2].

There is a vast literature of models for the NF-κB signalling pathway [3,4,5,6]. Here we focus on the model presented in [5]. A schema of the pathway is reported in Fig. 1. The species involved are the IκB kinase (IKK), NF-κB, the protein A20, the protein

IκBα, their complexes and mRNA transcripts of A20, IκBα and an hypothetical control gene (cgen). Under normal conditions, NF-κB is kept in the cytoplasm by the inhibitor protein IκBα. When an upstream stimulus (SIGNAL), such as the Tumor Necrosis Factor (TNF), is activated the protein IKK in the neutral form (IKKn) is transformed into its active phosphorylated form (IKKa). In this form it is capable of triggering the degradation of IκBα and, as a result, free cytoplasmic NF-κB is released. This enters the nucleus and upregulates the transcription of the two inhibitors, A20 and IκBα, and a large number of other genes. The newly synthesized IκBα again inhibits NF-κB while A20 inhibits IKKa by catalysing its transformation into another inactive form (IKKi), which is not able to trigger the degradation of IκBα anymore.

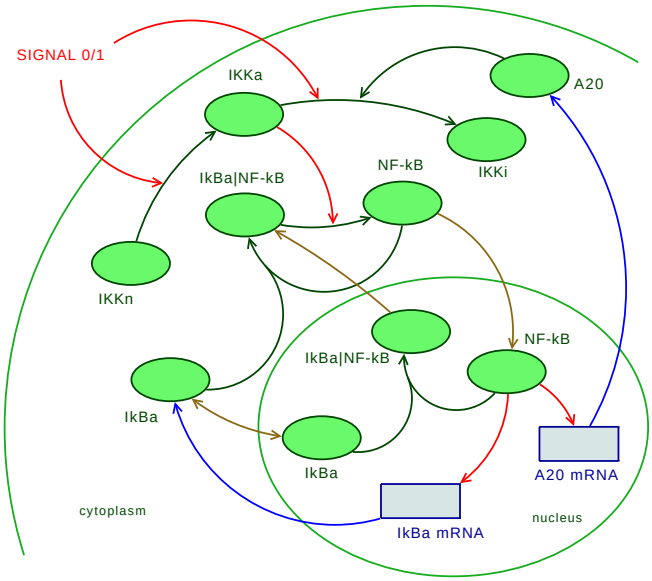

Fig. 1. Schematic depiction of the NF-κB signalling pathway considered in the paper. The rectangles represent mRNAs whereas ovals represent proteins. The red arrows are the interactions triggered by the signal, the brown arrows the transport reactions between compartments, the blue arrows the translation of mRNAs into proteins, black arrows represent all the other kinds of interaction (association and dissociation reactions). Compartments are delimited by green lines.

The pathway is characterized by the following main features:

1. There are *two compartments*, the nucleus and the cytoplasm. Some biochemical species can move from one compartment to the other. In other models compartments are not considered explicitly or they are based on the (non realistic) assumption that they have the same size [3]. Here compartments have different sizes and this information is taken into account in the derivation of rates and concentrations.

2. The *external stimulus* is represented by a *signal*, active from time $T_1 = 1$ to $T_2 = 7$ (hours). When the signal is present, some reactions are enabled (the rate is different from zero): the activation of IKKn and the transformation of IKKa into IKKi.
3. 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.

The model in [5] is in terms of ODEs where variables describe species concentrations (in μM). The model was validated against the available experimental data and some analyses were carried out. A hybrid model for the same pathway is described 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.

3 Bio-PEPA

In this section we give a short description of Bio-PEPA [12,13], a language that has recently been developed for the modelling and analysis of biological systems. The main components of a Bio-PEPA system are the *species 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:

$$S ::= (\alpha, \kappa)\ \mathsf{op}\ S \mid S + S \mid C \quad \text{with } \mathsf{op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot \qquad P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(x)$$

where S is the *species component* and P is the *model component*. In the prefix term $(\alpha, \kappa)\ \mathsf{op}\ S$, κ 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*, $\uparrow$ a *product*, $\oplus$ an *activator*, $\ominus$ an *inhibitor* and $\odot$ a generic *modifier*. We can use "$\alpha\ \mathsf{op}$ " and "$(\alpha, \kappa)\ \mathsf{op}$ " as abbreviations for "$(\alpha, \kappa)\ \mathsf{op}\ S$" and "$(\alpha, 1)\ \mathsf{op}\ S$", respectively. The operator "+" expresses the choice between possible actions, and the constant C is defined by an equation $C \stackrel{def}{=} S$. The process $P \underset{\mathcal{L}}{\bowtie} Q$ denotes synchronisation between components P and Q, the set $\mathcal{L}$ determines those activities on which the operands are forced to synchronise, with $\underset{*}{\bowtie}$ denoting a synchronisation on all common action types. In the model component $S(x)$, the parameter $x \in \mathbb{R}$ represents the initial concentration (or the number of molecules in a discrete-stochastic setting). The reader is referred to [13] for further details on the language and its semantics.

Recently Bio-PEPA has been extended to incorporate events [16] and to support biological locations [17].

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

id is the event name, *trigger* is a mathematical expression involving the components of the Bio-PEPA model and/or time, *event_assignment* is a list of assignments causing some changes to elements in the system, and *delay* is either 0 (*immediate events*) or a positive real value (*delayed events*).

Locations represent both biological compartments and membranes. They are static (i.e. they have a fixed structure) but they can change size with time. In the definition of locations it is possible to express their position with respect to the other locations of the system and their kind (i.e. compartment or membrane). The notation $C@L$ indicates that the species represented by the component C is in the location L. The structure of the biological system is modelled as a *static hierarchy*, represented as a tree whose nodes represent locations (compartments and membranes); each node has one child for each of their sub-locations. The locations are defined as follows.

Definition 1. *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.*

A Bio-PEPA system representing a biochemical network consists of a set of sequential components, a model component, and a context (defining information such as kinetics rates, parameters, locations, and events). Its formal definition is the following.

Definition 2. *A Bio-PEPA system $\mathcal{P}$ is a 8-tuple $\langle t, \mathcal{L}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P, Events\rangle$, where: t is time, $\mathcal{L}$ is the set of locations, $\mathcal{N}$ is the set of (auxiliary) information for the species, $\mathcal{K}$ is the set of parameters, $\mathcal{F}_R$ is the set of functional rates, $Comp$ is the set of species components, P is the model component and $Events$ is the set of events.*

Bio-PEPA offers a formal intermediate compositional representation of biochemical systems, on which different kinds of analysis can be carried out, through the defined mappings into continuous-deterministic and discrete-stochastic modelling languages. The Bio-PEPA language is supported by software tools (for instance the Bio-PEPA Workbench [18]) which automatically process Bio-PEPA models and generate other representations in forms suitable for different kinds of analysis [13,19]. In particular, the generated simulation model can be executed using MATLAB [20] and the Dizzy simulation tool [21], 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 [15], which extends the original tool with sensitivity analysis techniques and additional simulation methods.

4 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.

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:

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

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 constant rate r.

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:

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

This species is in the cytoplasm and it is involved in three interactions: its synthesis, its degradation and it is an activator of the transformation of IKKa into the inactive form IKKi.
The species and their possible interactions are represented by the *model component*:

$$IKKn@cyt[0] \underset{*}{\bowtie} IKKa@cyt[0] IKKi@cyt[0] \underset{*}{\bowtie} \underset{*}{\bowtie} \cdots A20@cyt[0]$$

where the number between square brackets represents the initial concentration.

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

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

In our case $T_1 = 1$ and $T_2 = 7$ (time expressed in *hours*) and both events are immediate.

Note that the rates and the initial amounts in the Bio-PEPA model are in terms of concentration. In order to derive a stochastic model or, generally, a model in terms of molecule numbers, the continuous concentration values must be translated into discrete numbers of molecules and the rates must be modified in order to take this transformation into account. In general, assuming concentrations are expressed in *molars* (M), the

initial amounts must be multiplied by the scaling factor $N_A \cdot V$ (where N_A is the Avogadro number and V is the volume of the compartment where the reactions take place), and the kinetic parameters must be rescaled accordingly (see [22] for details).

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

$$nscale = V_n \cdot N_A \cdot 10^6 = 2 \cdot 10^5 \qquad cscale = V_c \cdot N_A \cdot 10^6 = 10^6$$

where *nscale* and *cscale* are the scaling factors and V_n and V_c for the volume sizes of the nucleus and cytoplasm, respectively. As the concentrations are in terms of μM instead of M we have to multiply by the scaling factor 10^6. Furthermore, we have to define the proportion factor between the two compartments $k = V_c/V_n = 5$. This parameter is used to rescale in the appropriate way the rates of the reactions involving reactants and products in two different compartments, such as the transport of a species from the nucleus to the cytoplasm.

5 Validation and analysis

The Bio-PEPA model for the NF-κB pathway was implemented in the Bio-PEPA Workbench [18]; from it various kinds of analysis can be performed. Here we consider the ODE MATLAB model for the validation and the Dizzy model for the stochastic simulation and sensitivity analysis. Gillespie's direct method [22] is the stochastic algorithm chosen for our analysis. The same results are obtained if other stochastic algorithms implemented in Dizzy (including the ones supporting events) are used. The graphs have time expressed in seconds.

5.1 Validation

In order to validate our model against the original ODE model and the available experimental data in [5], we consider the ODE MATLAB model obtained from the Bio-PEPA Workbench [18]. The species are defined in terms of concentration.

In [5] the following approach is used for the simulation. At time 0 just the complex cytoplasmic IκBα |NF-κB is present (0.06 μM) and all other species are zero. First, the simulation is run for 100 hours in order to reach the resting cell equilibrium state. The simulation is then run for a further 7 hours, with the external signal enabled after 1 hour. The ODE solver used is the MATLAB *ode23tb*, an implementation of *TR-BDF2* [23]. This is an implicit Runge-Kutta formula with a first stage that is a trapezoidal rule step and a second stage that is a backward differentiation formula of order two. This solver is useful for stiff systems.

In Figure 2 we report the results obtained by running our MATLAB ODE model following the approach described above. The results are in full agreement with the results shown in the paper.

The same results (but rescaled by the appropriate scaling factors) are obtained if we consider the ODE model in terms of number of molecules both in MATLAB and Dizzy.

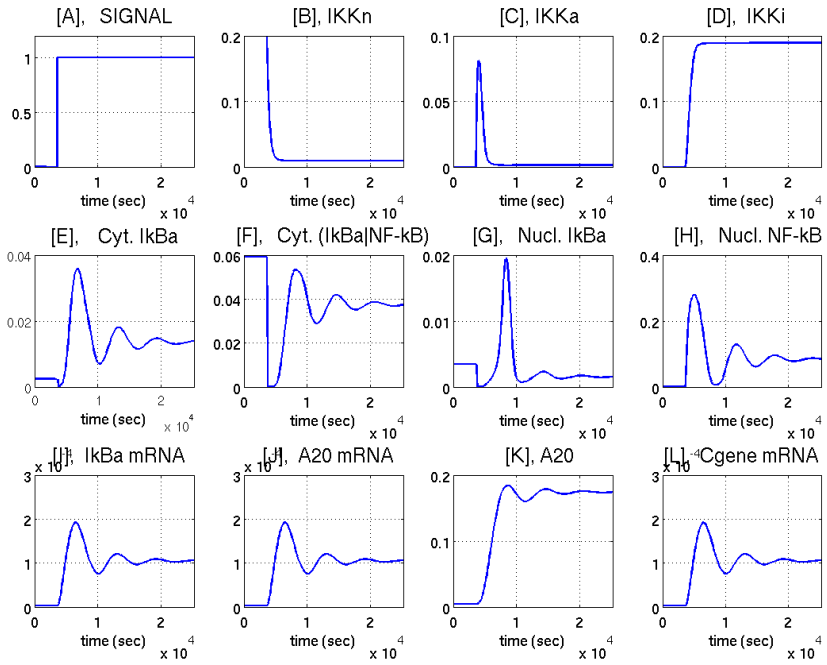

Fig. 2. Validation of the ODE MATLAB model obtained from Bio-PEPA (concentrations). The time is in seconds.

In Figure 3 we report the results in terms of molecules for the nuclear NF-κB and IKKa, as these two species are the ones which our further analyses focus on and the other two species cytoplasmic IκBα and IκBα |NF-κB. In Dizzy we use the *ODEtoJava-dopr54-adaptive* solver [24]. This solver uses a variable time-step size that is controlled by an adaptive method involving a formula for estimating the error. If the error gets too large, the time-step size is decreased until the error is acceptable.

The agreement between the two ODE models in terms of concentration and number of molecules confirms that our scaling is appropriate.

5.2 Stochastic simulation

From Bio-PEPA it is possible to define a stochastic version of the model, in terms of number of molecules. 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 [25]. This is particularly true when the activation of genes is involved, as generally there ate few copies of each gene in the cell.

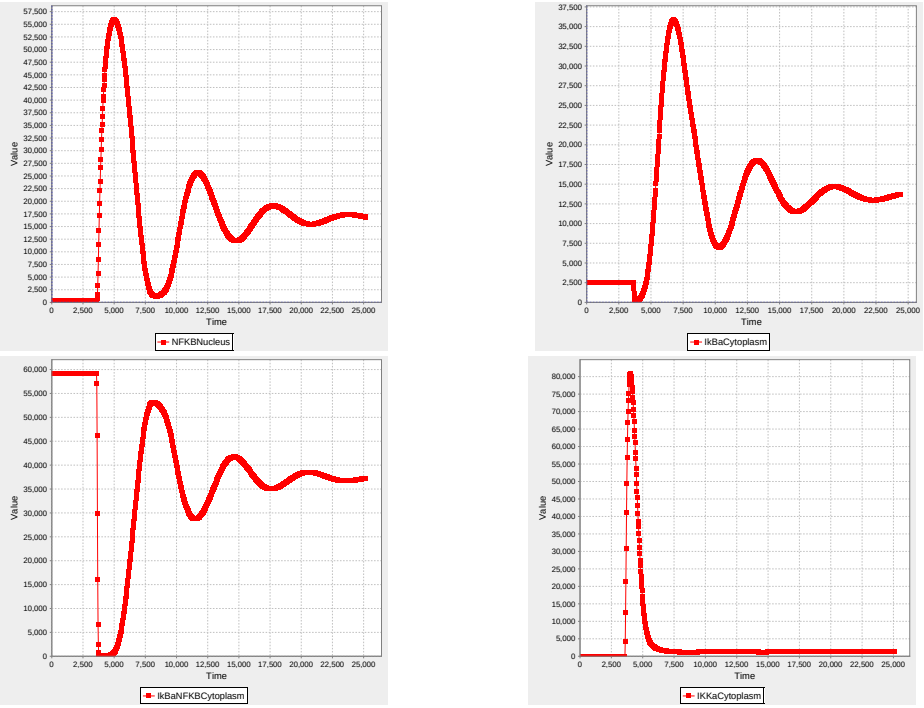

Fig. 3. ODE simulation results in terms of number of molecules for some species in the pathway: nuclear NF-κB (top-left), cytoplasmic IκBα (top-right), cytoplasmic IκBα |NF-κB (bottom-left) and IKKa (bottom-right). These species correspond to the species in the subgraphs H, E, F and C of Figure 2, respectively.

Our aim was to verify the effect of stochasticity on the system behaviour. Figure 4 reports the average and the standard deviation of some species. 100 simulation runs are considered. 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.

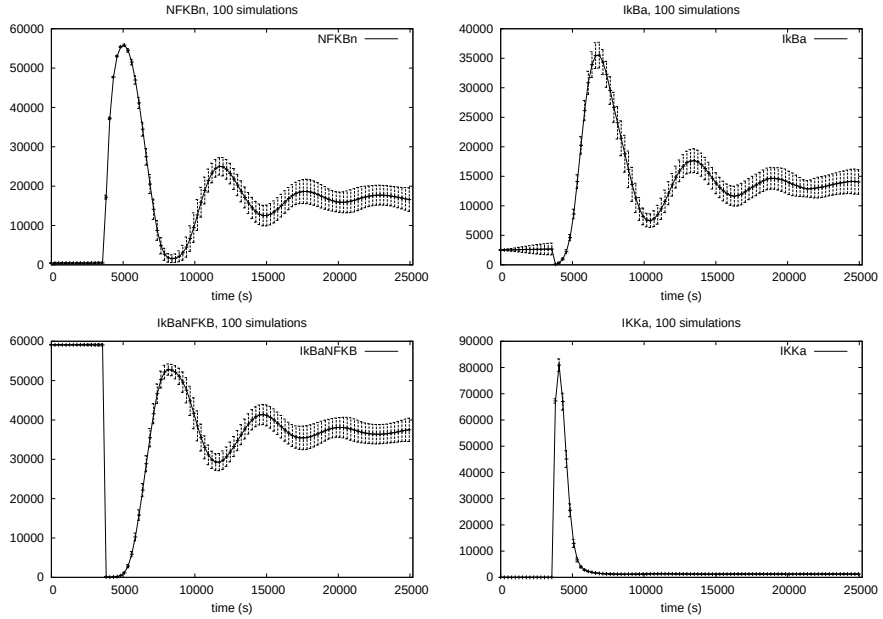

Fig. 4. Stochastic simulation for some species in the pathway: nuclear NF-κB (top-left), cytoplasmic IκBα (top-right), cytoplasmic IκBα |NF-κB (bottom-left) and IKKa (bottom-right). For each of them the average value and the standard deviation are shown. 100 simulation runs are considered. Gillespie's direct method is the algorithm used. These species correspond to the species in the subgraphs H, E, F and C of Figure 2, respectively.

5.3 Some experiments

Our model can be then used to predict the behaviour of the system under different assumptions. In particular, it is interesting to study what happens when we change the duration of the external stimulus and what the effect of the initial amount of total NF-κB over the nuclear NF-κB is.

Figure 5 shows the average over 100 stochastic runs for IKKa and nuclear NF-κB when the stimulus lasts for 15 minutes, 60 minutes and 6 hours (original value). 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 one hour and a half and lasts about one hour. This similarity between the three situations

described above may be due to the fact the pulse of NF-κB is strictly influenced by IKKa and the behaviour of IKKa is the same for all the three cases. Indeed, the IKKa activation seems, at least in part, independent from the duration of the TNF stimulus. However, the duration of the stimulus has an impact on the behaviour of NF-κB after the 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). Moreover with the longer stimulus we see a pronounced oscillation.

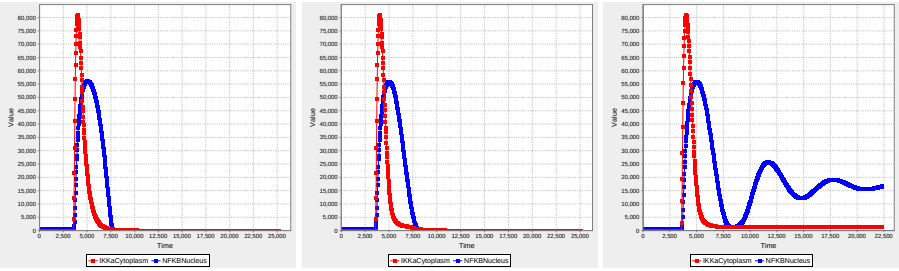

Fig. 5. IKKa (red line) and nuclear NF-κB (blue line) at and after 15 minute-long (left), 60 minute-long TNF stimulation (centre) and at and after 6 hours (right). The graphs show the average amount over 100 simulation runs.

In Figure 6 we report the average behaviour over 100 runs of the nuclear NF-κB for different assumptions about the initial amount of the total NF-κB. An increase in the total amount of total NF-κB makes the oscillation more pronounced whereas a decrease smooths out the oscillations.

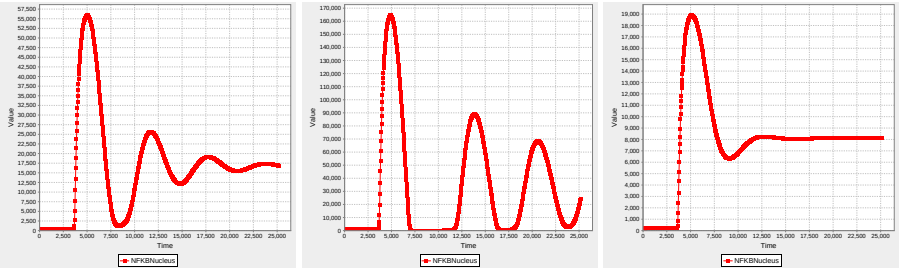

Fig. 6. Nuclear NF-κB when the initial total NF-κB is $6 \cdot 10^4$ (original value, left), is $1.8 \cdot 10^5$ (three times the original value, centre), and $2 \cdot 10^4$ (one third of the original value, right). The graphs show the average amount over 100 simulation runs.

5.4 Sensitivity analysis

Sensitivity Analysis (SA) aims to identify the relationships between the inputs and outputs of mathematical models of biochemical networks [26]. A key goal is the production of Sensitivity Indices (SI) that quantify these relationships, revealing which factors are the most influential with respect to model outcome. The most widespread SA method is "one-at-a-time" (*OAT*). Given a mathematical model with parameters set to those considered the most likely (also called *nominal parameters*), each parameter is perturbed individually by a fixed value or by a percentage of its nominal value, and the change in the output(s) of interest measured. OAT has seen widespread use in ODE models of biochemical interactions.

In [14] this method has been extended to stochastic models. In this case, the output at a given time is not just a value representing the amount of a species as in ODEs; it is, instead, a set of possible values, obtained from independent stochastic simulations. The SA extension has been obtained by substituting the difference between perturbed and nominal output values employed in the traditional approach with a difference measure based on the density distribution surface of the output, estimated with a suitable number of simulations. An estimation of this density distance based on stochastic simulations can be obtained using *histogram distance*, as originally presented in [27]. This stochastic version of OAT therefore 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 was implemented in the version of the Dizzy simulator developed at the University of Edinburgh [15].

Here we present a preliminary investigation concerning the isolation of the three parameters that have most impact on the various forms of NF-κB (nuclear or cytoplasmic, free or in complex with other species). We consider an OAT approach with histogram distance as described above and we apply it to our stochastic model. As a result of our study, we obtain that the most influential parameters are the following:

- c_{1a}, the rate of transcription of IκBα mRNA induced by NF-κB;
- c_{4a}, the rate of translation of IκBα;
- c_{3a}, the rate of degradation of IκBα mRNA.

In the following we focus on the the parameter c_{1a} and we show what happens to the behaviour of nuclear NF-κB when we change the value of c_{1a}. Figure 7 reports the results of stochastic simulation (Gillespie direct, average of 100 runs) when $c_{1a} = 0$ and c_{1a} is three times the original value. The first case corresponds to the case of A20 deficient cells. With respect to the wild-type cells (represented by our original model) NF-κB accumulates in the nucleus and remains there throughout the TNF stimulation. This is probably due to the fact that IKKa presents a constant level of activity different from zero in the tail following the peak. In the second case, nuclear NF-κB shows stronger oscillations and the final amount is lower than in the original wild-type cells. These results are in agreement with the reported experimental data [5].

6 Conclusions

In this work we presented a Bio-PEPA model for the NF-κB signalling pathway. With our model we were able to describe some features of the system, such as compartments

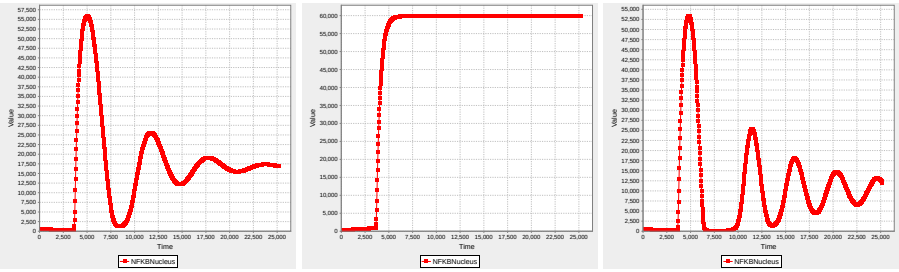

Fig. 7. NF-κB when $c_{1a} = 2.5 \cdot 10^{-6}$ (original value, left) $c_{1a} = 0$ (A20 deficient cells, centre), $c_{1a} = 3 \cdot 2.5 \cdot 10^{-6}$ (right). Gillespie direct algorithm with 100 runs is considered.

and the activation of NF-κB by an external stimulus. The former are represented by locations whereas the latter by means of temporal events. We validated our model against the experimental data present in the literature [5]. Furthermore we defined a stochastic version of the model in order to investigate the effect of fluctuations over the behaviour of the system. Finally, sensitivity analysis was considered to isolate the most relevant parameters of the model.

There is a vast literature of models describing NF-κB signalling pathway [3,4,5,6]; most of previous work involves ODE models. There are just a few applications of process algebras for the modelling and the study of this pathway [28,29]. In [29] the pathway previously reported in [3] is represented in *BetaWB*, a language based on Beta binders [10]. The analysis is based on stochastic simulation. Note that in [3] the two compartments are assumed to have the same size and they are not considered explicitly in the BetaBW model but their size is just reflected in the derivation of the number of molecules and stochastic rates. In [28] the authors defined a PEPA [30] model for the pathway presented in [4], describing in detail the first part of the signalling cascade up to the activation of NF-κB. The map from the PEPA model into the associated ODE model is presented and ODE numerical integration results are shown.

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. Here we focused just on deterministic simulation, stochastic simulation and sensitivity analysis.

Concerning sensitivity analysis, we considered a local approach, i.e. it focuses around a specific point in the parameter space. This can still be informative, giving an idea of the impact of parameter changes on the behaviour of the system. In the future, we plan to apply some global methods in order to explore the full parameter space (or a meaningful subset of it) and to quantify the relationships between different parameters.

Among the different mappings from Bio-PEPA to models for analysis, there is the mapping to continuous time Markov chain (CTMC). In particular, it is possible to derive a PRISM [31,32] model in order to verify some properties by model checking. Two main problems for the use of model checking with our model are the dimension of the state space (it is large) and the presence of temporal events. For the former problem, one possibility is to apply some abstract-view for the CTMC in terms of concentration levels

[33]. This approach is helpful for the reduction of the state space. The latter problem is still open and we are looking for some approaches to handle it.

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. 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
8. Curti, M., Degano, P., Priami, C., Baldari, C.: Modelling biochemical pathways through enhanced π-calculus. Theoretical Computer Science **325** (2004) 111–140
9. Regev, A., Panina, E., Silverman, W., Cardelli, L., Shapiro, E.: BioAmbients: an Abstraction for Biological Compartments. Theoretical Computer Science **325** (2004) 141–167
10. 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
11. 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
12. 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)
13. Ciocchetta, F., Hillston, J.: Bio-PEPA: a Framework for the Modelling and Analysis of Biochemical Networks. Theoretical Computer Science (2009, to appear)
14. 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
15. Dizzy Edinburgh version: (http://homepages.inf.ed.ac.uk/stg/software/Dizzy/)
16. Ciocchetta, F.: Bio-PEPA with SBML-like events. In: Proc. of Computational Models for Cell Processes. (2008)
17. Ciocchetta, F., Guerriero, M.: Modelling Biological Compartments in Bio-PEPA. In: Proc. of MeCBIC 2008. Volume 227 of ENTCS. (2009) 77–95
18. The Bio-PEPA Workbench: (http://www.dcs.ed.ac.uk/home/stg/software/biopepa/about.html)
19. Bio-PEPA Home Page: (http://www.biopepa.org/)
20. MATLAB Home Page: (http://www.mathworks.com/products/matlab/)
21. Dizzy Home Page: (http://magnet.systemsbiology.net/software/Dizzy)
22. Gillespie, D.: Exact stochastic simulation of coupled chemical reactions. J Phys Chem **81** (1977) 2340–2361

23. Bank, R., et al.: Transient simulation of silicon devices and circuits. IEEE Transactions on Electron Devices **32** (1985) 1992–2007
24. Stoer, J., Bulirsch, R.: Introduction to Numerical Analysis. New York: Springer-Verlag (1993)
25. McAdams, H., Arkin, A.: Stochastic mechanisms in gene expression. Proc Natl Acad Sci USA **94** (1997) 814–9
26. Heinrich, R., Schuster, S.: The Regulation of Cellular Systems. Chapman & Hall (1996)
27. 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
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. Kwiatkowska, M., Norman, G., Parker, D.: Prism: Probabilistic model checking for performance and reliability analysis. ACM SIGMETRICS Performance Evaluation Review (2009)
32. PRISM Home Page: (http://www.prismmodelchecker.org)
33. 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

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

In this Appendix we report the full Bio-PEPA model of 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 entrainment are defined. 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 concentration.

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

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

$k_{prod} = 0.000025$; $k_{deg} = 0.000125$; $k_1 = 0.0025$; $k_3 = 0.0015$; $k_2 = 0.1$;
$c_1 = 5 \cdot 10^{-7}$; $c_2 = 0$; $c_3 = 0.0004$; $c_4 = 0.5$; $c_5 = 0.0003$; $t_1 = 0.1$; $t_2 = 0.1$;
$c_{1a} = 5 \cdot 10^{-7}$; $c_{2a} = 0$; $c_{3a} = 0.0004$; $c_{4a} = 0.5$; $c_{5a} = 0.0001$; $c_{6a} = 0.00002$;
$a_1 = 0.5$; $a_2 = 0.2$; $a_3 = 1$; $a_{1n} = 0.5$; $c_{1c} = 5 \cdot 10^{-7}$; $c_{2c} = 0$; $c_{3c} = 0.0004$;
$i_{1a} = 0.001$; $e_{1a} = 0.0005$; $e_{2a} = 0.01$; $i_1 = 0.0025$

$$
\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 + (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\_synthesis, 1)\uparrow + (a20\_degradation, 1)\downarrow \\
I\kappa B\alpha@cyt &\stackrel{def}{=} (IkBa\_synthesis, 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
\end{aligned}
$$

$$
\begin{aligned}
A20t@cyt &\stackrel{def}{=} (a20\_synthesis, 1) \odot + (a20t\_transcription, 1)\uparrow + \\
&\quad (a20t\_transcription\_by\_NFkBn, 1)\uparrow + (a20t\_degradation, 1)\downarrow \\
I\kappa B\alpha t@cyt &\stackrel{def}{=} (ikBat\_transcription, 1)\uparrow + (ikBat\_transcription\_by\_NFkBn, 1)\uparrow + \\
&\quad (ikBat\_degradation, 1)\downarrow + (ikBa\_synthesis, 1)\odot \\
Cgent@cyt &\stackrel{def}{=} (gent\_transcription, 1)\uparrow + (cgent\_transcription\_by\_NFkBn, 1)\downarrow + \\
&\quad (cgent\_degradation, 1)\downarrow
\end{aligned}
$$

where the names *complex*1, *complex*2, *complex*3 and *complex*4 stand for $I\kappa B\alpha \mid NF\text{-}\kappa B@cyt$, $IKKa \mid I\kappa B\alpha@cyt$, $IKKa \mid I\kappa B\alpha \mid NF\text{-}\kappa B@cyt$ and $I\kappa B\alpha \mid NF\text{-}\kappa B@nuc$, 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.

$$
\begin{aligned}
&IKKn@cyt[0] \underset{*}{\bowtie} IKKa@cyt[0] \underset{*}{\bowtie} IKKi@cyt[0] \underset{*}{\bowtie} \\
&\qquad A20@cyt[0] \underset{*}{\bowtie} I\kappa B\alpha@cyt[0] \underset{*}{\bowtie} NF\text{-}\kappa B@cyt[0] \underset{*}{\bowtie} \\
&\qquad I\kappa B\alpha \mid NF\text{-}\kappa B@cyt[0.06] \underset{*}{\bowtie} IKKa \mid I\kappa B\alpha@cyt[0] \underset{*}{\bowtie} \\
&\qquad\qquad IKKa \mid I\kappa B\alpha \mid NF\text{-}\kappa B@cyt[0] \underset{*}{\bowtie} \\
&\quad I\kappa B\alpha@nuc[0] \underset{*}{\bowtie} NF\text{-}\kappa B@nuc[0] \underset{*}{\bowtie} I\kappa B\alpha \mid NF\text{-}\kappa B@nuc[0] \underset{*}{\bowtie} \\
&\qquad A20t@cyt[0] \underset{*}{\bowtie} I\kappa B\alpha t@cyt[0] \underset{*}{\bowtie} Cgent@cyt[0]
\end{aligned}
$$

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

In our case $T_1 = 1$ and $T_2 = 7$ (time expressed in *hours*).

Parameters for the stochastic model

The stochastic version of the model is obtained rescaling all the initial species concentrations by the factors *nscale* (species in the nucleus) and *cscale* (species in the cytoplasm) and modifying the parameters in order to take the number of molecules into account. 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.

Below we report how the parameters are rescaled; just the modified ones are shown.

$$
\begin{aligned}
&k_{prod} = 0.000025 \cdot cscale = 25; && k_2 = 0.1/cscale = 1 \cdot 10^{-7}; \\
&c_1 = 5 \cdot 10^{-7} \cdot k = 2.5 \cdot 10^{-6}; && c_2 = 0 \cdot k = 0; \\
&c_{1a} = 5 \cdot 10^{-7} \cdot k = 2.5 \cdot 10^{-6}; && c_{2a} = 0 \cdot k = 0; \\
&a_1 = 0.5/cscale = 5 \cdot 10^{-7}; && a_2 = 0.2/cscale = 2 \cdot 10^{-7}; \\
&a_3 = 1/cscale = 1 \cdot 10^{-7}; && a_{1n} = 0.5/nscale = 2.5 \cdot 10^{-6}; \\
&c_{1c} = 5 \cdot 10^{-7} \cdot k = 2.5 \cdot 10^{-6}; && c_{2c} = 0 \cdot k = 0; \\
&e_{1a} = 0.0005 \cdot k = 0.0025; && e_{2a} = 0.01 \cdot k = 0.05
\end{aligned}
$$