Universiteit
Leiden
The Netherlands

Computational modelling of mycobacterium infection and innate immune response in zebrafish

Viana de Carvalho, R.

Citation
Viana de Carvalho, R. (2015, October 15). *Computational modelling of mycobacterium infection and innate immune response in zebrafish*. Retrieved from https://hdl.handle.net/1887/35896

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/35896

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/35896 holds various files of this Leiden University dissertation.

Author: Viana de Carvalho, Rafael
Title: Computational modeling of mycobacterium infection and innate immune reponse in zebrafish
Issue Date: 2015-10-15

5 Quantitative Petri net model approach for the *Mycobacterium* infection process

Based on:

Carvalho, Rafael V., van der Sar, Astrid M., Kleijn, Jetty, Verbeek, Fons J.
In preparation for submission 2015 (2015).

Our days as a polar winter .

Its hours of joy are glancing as lightning.
And adversities are as a night, never ends.
For happiness, Times run fast.
And for darkness, the seasons reside.

And in days that are heavy as lead
Flashes as a quick lightening, a love
this is illogical.
Then poetry erupts as fire
in straw hit by a thunderbolt.
In all these spreading ashes
How did these burning words survived?

Jabra Ibrahim Jabra, Palestina (1920 – 1994)

أيامُنا كالشتاءِ القطبيِّ:

ساعاتُ الفرح فيها, كالضياء, خاطفة.
والفواجعُ, كالليل, لا تنتهي.
للإشراقات أوقات ما أسرع ركضها
وللظلماتِ المواسمُ المقيمة.

وفي نهارات, أثقالُها كالرصاص
يومض كخطفِ البرق حبُّ
لا يُفهم منطقه,
ويندلعُ الشعرُ كاللهيب
في هشيم ضربتُه الصاعقة:
في هذا الرماد العنّيِّ المنتشر
كيف بقيت هذه الكلمات الحارقة ؟

جبرا إبراهيم جبرا، فلسطين، (1920 - 1994)

Abstract

In the introduction of this thesis we have stated that a model has to concisely summarize and replicate the relevant information about the biological process. In the previous chapters, we have presented models, which outline the components and dynamics of the *Mycobacterium* infection process. These models have been used as building blocks to create a qualitative model, with, according to the results, a validstructure . However, a reliable model has to reproduce the structure and behavior of the biological process. In this chapter we present a model that fits these requirements and can simulate scenarios that are in accordance with the behavior of the infection process. We used data from biological experiments to validate the model leading to a quantitative analysis of the simulations.

5.1 Introduction

The availability of data representing various biological states, processes and their time dependencies enables the study of biological systems at various levels of organization, from molecules to organism and even up to population level. It is also essential to build models that can simulate scenarios and predict behavior of the system. Moreover, we can use the biological data to validate the model leading to a quantitative analysis of the experiments. Therefore it can contribute to identify important parameters that can help to unravel mechanisms related to the biological process. Walpole *et al.* [123] review the importance of integrating data across spatial, temporal and functional scales and how multi-scale models, that are developed to work across these scales, may provide insights into biological systems using quantitative methods.

The QPN^C model presented in Chapter 4 provides a visualization of the dynamic processes involved in Mycobacterium infection in the zebrafish host as well as the interconnected pathways that interact once bacteria are inside the host immune cell (phagocytosis). Although we have used quantified boundaries to express the dynamics of the model, our earlier focus was to create a pure qualitatively model using data information only as boundary condition for the model simulation. Analyzing the deterministic aspect of the model, these boundaries add an oscillation in the final result according to the simulations. The reason was the accumulation of tokens in the sub-nets that model the pathways related to the mycobacterial proliferation (intracellular, intercellular and molecular interaction). This accumulation occurs along the iterations of the simulation process in order to provide a non-deterministic result. The accumulation of tokens did not interfere in the model simulation. In this chapter, we analyze data from biological experiments on *Mycobacterium marinum* infection in zebrafish embryos. We use this information to redesign our previous qualitative model from Chapter 4 into a quantitative model using SPN^C. For validation, we execute two scenarios, i.e. simulations of certain conditions, and compare the results with the data from *in vivo* experiments. The process of validation and refinement may lead to the discovery of new knowledge about the infection process.

The remainder of this chapter is structured as follows. In Section 5.2 we focus on the material and methods describing the data analysis process and the description of the extension of the previous model emphasizing the quantitative aspects. We discuss about the changes in the model to fit to the problem and how the data is implemented in the model. In Section 5.3 we show our results from the simulations. We compare with the data from *in vivo* experiments and validate the model. Finally, in Section 5.4 we present our conclusions and discuss the results pointing to possible directions for future work.

5.2 Material and methods

The original motivation was to refine the previous models, removing the unnecessary accumulation of tokens used to provide a non-deterministic qualitative output of the simulation process. To tackle the quantitative aspects of the infection process, we have converted the QPN^C model into a Stochastic Colored Petri net (SPN^C). This process assigns a rate function to the transitions that will address a probability to "fire" during the simulation. Therefore, the firing rule adds an extra condition for a transition to consume/produce tokens. When a particular transition becomes enabled, then a local timer is set to an initial value, which is computed at this time point by means of the corresponding probability distribution. The local timer is then decremented at constant speed, and the transition will fire when the time reaches zero [53].

The hierarchical structure of our model is not affected (cf. Section 4.3), though we restructure the dynamic level in order to add quantitative aspects related to the distribution of the infection (granuloma formation) along the positions. These data are derived from a statistical analysis from *in vivo* experiments with *Mycobacterium marinum* in zebrafish embryos [113]. The stochastic model is implemented in the Snoopy tool. This software environment also allows to execute simulations and animations with SPN^C [50]. In the following sections, we will present the data analysis and the stochastic model implemented.

5.2.1 Data collection and analysis

As the basis for our model, we have analyzed data from *in vivo* experiments on zebrafish embryos. The empirical data consist of microscope images of zebrafish, that are 6 day post fertilization (6 *dpf*), infected with wild-type *Mycobacterium marinum* (E11 strain) inoculated into the caudal vein immediately after the onset of blood circulation, i.e. at 28 hours post fertilization (28 *hpf*). The inoculated amount of bacteria ranges from 50 to 200 colony forming units (CFUs). The images were acquired by Leica stereo fluorescence microscope. The images contain infected zebrafish at 5 days post infection (5 *dpi*).The images were part of experiments performed at the department of Medical Microbiology and Infection control (MMI) from the VU Medical Center Amsterdam, under the procedures presented in [113]. For screening, two channel images are used; one channel with a bright field image of the zebrafish and a second channel containing a fluorescent image visualizing the bacteria colonies.

For the analysis of the images, we have used a data analysis environment for high throughput screening (HTS) called *eLaborant* [90]. This computational platform offers a complete solution for image analysis and pattern recognition for zebrafish screening and extracts information about size, shape factors, intensity, texture and location of specific anatomical parts of the zebrafish in the images. The *eLaborant* platform analyzes the zebrafish images [89], by dividing the specimen in 12 regions and quantifying the fluorescence in the regions. It identifies the infection spread within the object and calculates the surface area of the bacterial signal by performing relative measurements about the proportional surface area of infection, cluster count and cluster infection intensity. In **Figure 5.1** one microscope image is depicted containing two zebrafish as analyzed by *eLaborant*. Here, we can notice the centers of the 12 regions (represented by dots and connected by a red line) as well as the granuloma distribution as obtained from the fluorescent channel and overlaid as magenta blobs [89, 113].

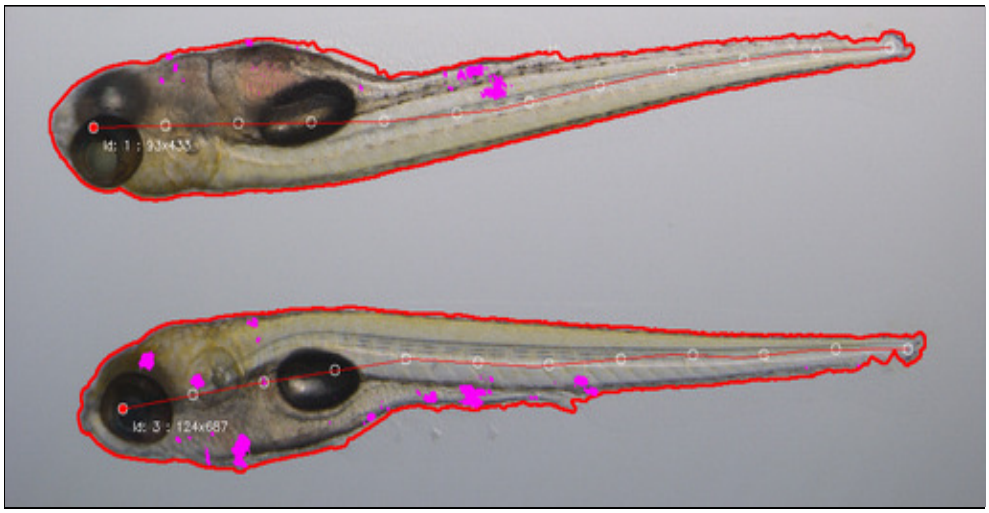


Figure 5.1: The image analysis result for the spread of the wild-type E11 strain on zebrafish with 5 *dpi*. Circles represent the centers of the 12 regions and the granulomas are displayed as magenta blobs.

For the scope of this work, we have considered the total area of infection, which represents the distribution of the infection along the fish according to the 12 pre-defined regions. To that end, we have analyzed 138 zebrafish infected larva by the E11 strain at 6 *dpf*. We calculate the average distribution of infection as proportional surface area of the total infection per segment (the pre-defined regions). The *eLaborant* measures the surface area of the infection by evaluating the fluorescence in the fluorescence channel. This sums to the total area of infection (A_t), i.e. the total surface are of granulomas in the zebrafish. Moreover, per segment (i) the surface area of the

granuloma is known (A_i). Consequently, the contribution per segment to the total infection, i.e. total granuloma size, is computed as:

$$PA_i = \frac{A_i}{A_t} \tag{5.1}$$

The contribution per segment to the total infection is expressed as a proportion of the sum of all regions (A_t). This is computed for our set of 138 samples and from the results we calculate the average proportion per region. This result is depicted in **Figure 5.2** where one can notice that the infection has a large concentration of granulomas localized near the head (referred to as regions 2 and 3), and also at the center of the zebrafish (referred to as regions 6 and 7). The caudal area (referred to as regions 11 and 12) is the site of bacterial inoculation, thus, the most infected regions have developed from dissemination of the bacteria over the body. We can also analyze the behavior of the infection from the perspective of the circulatory system; an infected macrophage travels while the bacteria proliferates inside its cytosol (cf. Section 4.2.1) we then can reason that the infected macrophage travels at least about 4 regions before leaving the circulatory system and attaches into the tissue to form the granuloma. One of the objectives of this work is to replicate this behavior as well as predict this dynamics over time using a quantitative model with stochastic behavior and taking the PN from chapter 4 as a starting point. In the next section we present the model and explain its characteristics.

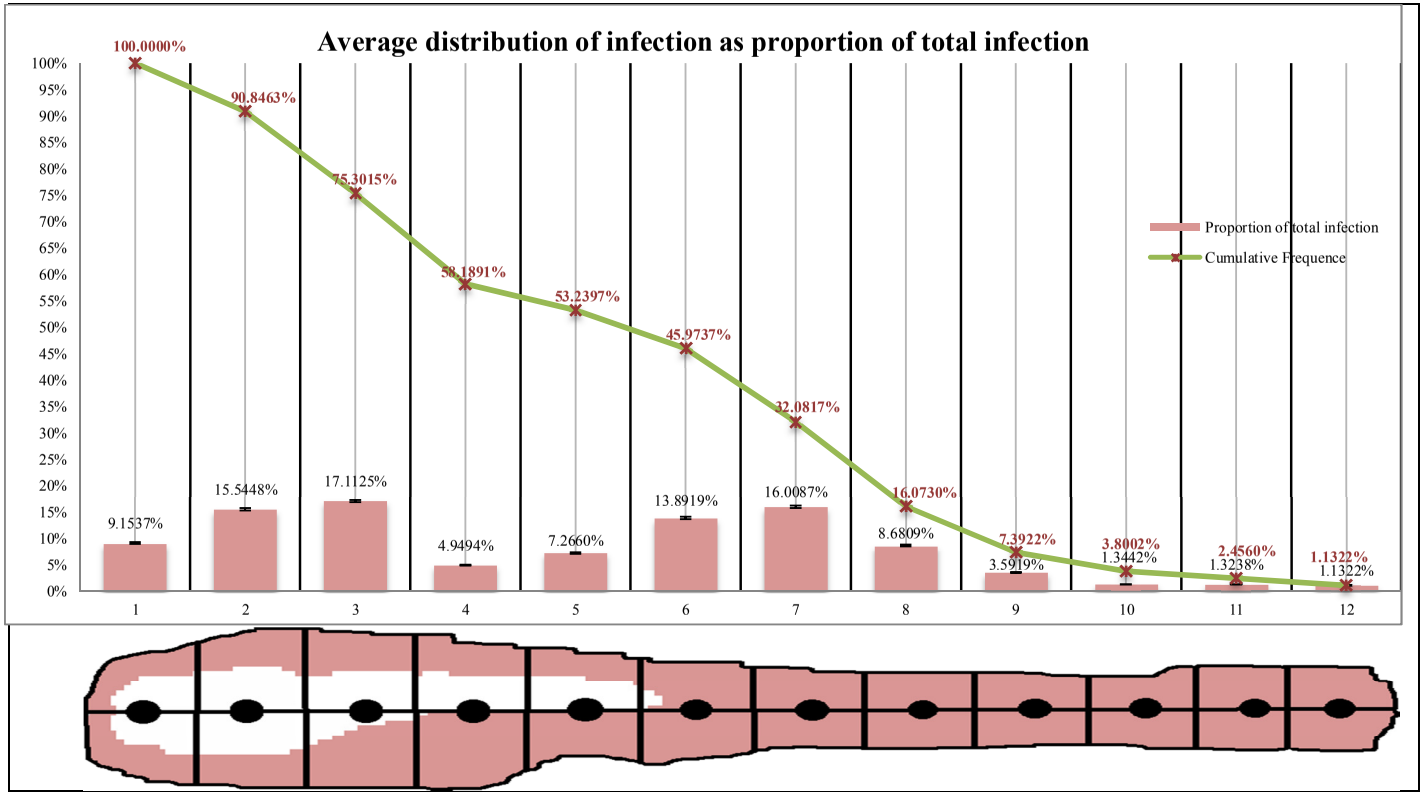
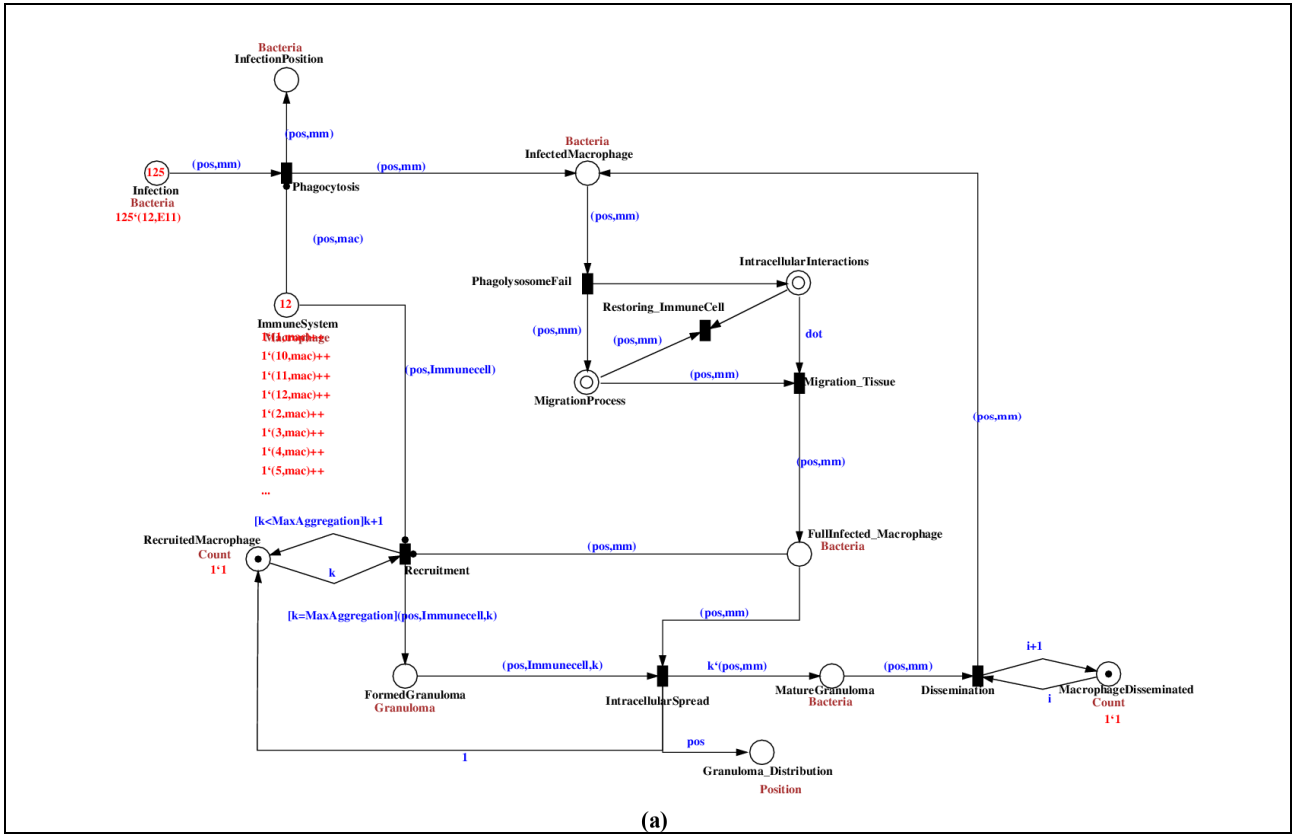


Figure 5.2: Empirical data analysis result: percentage of the distribution of the infection per region from 138 zebrafish at 6 *dpf* infected with wild-type (E11 strain) after 5 *dpi*. The result considers the total area of infection in the concentration of granuloma distributed over 12 regions.

5.2.2 Modeling decisions

The first step in the modeling process was to export our earlier model based on a QPN^C to a stochastic colored Petri net (SPN^C). The Snoopy tool provides portability processing between the methods [50]. By exporting the net, the tool preserves the discrete state space, but assigns exponentially distributed waiting times (probability) to transitions, which are specified by firing rate functions. The firing rates are typically state dependent and the predefined function is the kinetic mass action [74] assigned by the rate equal to 1 (cf. Chapter 1,**Appendix B** and **Appendix C** for a formal definition). The exporting process does not change components and structures, i.e. color-set, places, arcs, variables and guards. Therefore, the sub-nets remain the same.

In the next step, we refine the model by adapting its structure in order to make it suitable to simulate the quantitative aspects of the infection process. Therefore, we use the data obtained through experimental analysis in the stochastic transitions to regulate the migration process of the infected macrophage. There are 12 transitions with their own mass action function. The proportional granuloma distribution (PA_i) is used as rate ω_i in the kinetic mass action function for the 12 transitions. These data assign a probability of an infected macrophage migrating to a specific region of the zebrafish to form a granuloma. Besides the transitions related to the migration process, we replace all other stochastic transitions by immediate transitions. The immediate transitions are time independent (no rate/probability) and have higher priority than stochastic transitions to occur. Therefore, they will always fire, without any delay, if they are enabled. We have decided for this adaptation due to the limitations in our data-set, the quantitative information that we extracted from biological experiments relates to the infection distribution. There are no data (rates or time dependencies) about other interactions involved in the infection process, i.e. protein concentrations, time delays for molecular interactions, probability of a pathway, etc. Moreover, we have adapted the color-set and markings previously defined (cf. Chapter 4) to ensure that the model will be able to replicate the dynamics of the infection process in a stochastic simulation. In **Figure 5.3** the new model and its hierarchical structure are depicted.



(a)

Figure 5.3 *Cont.*

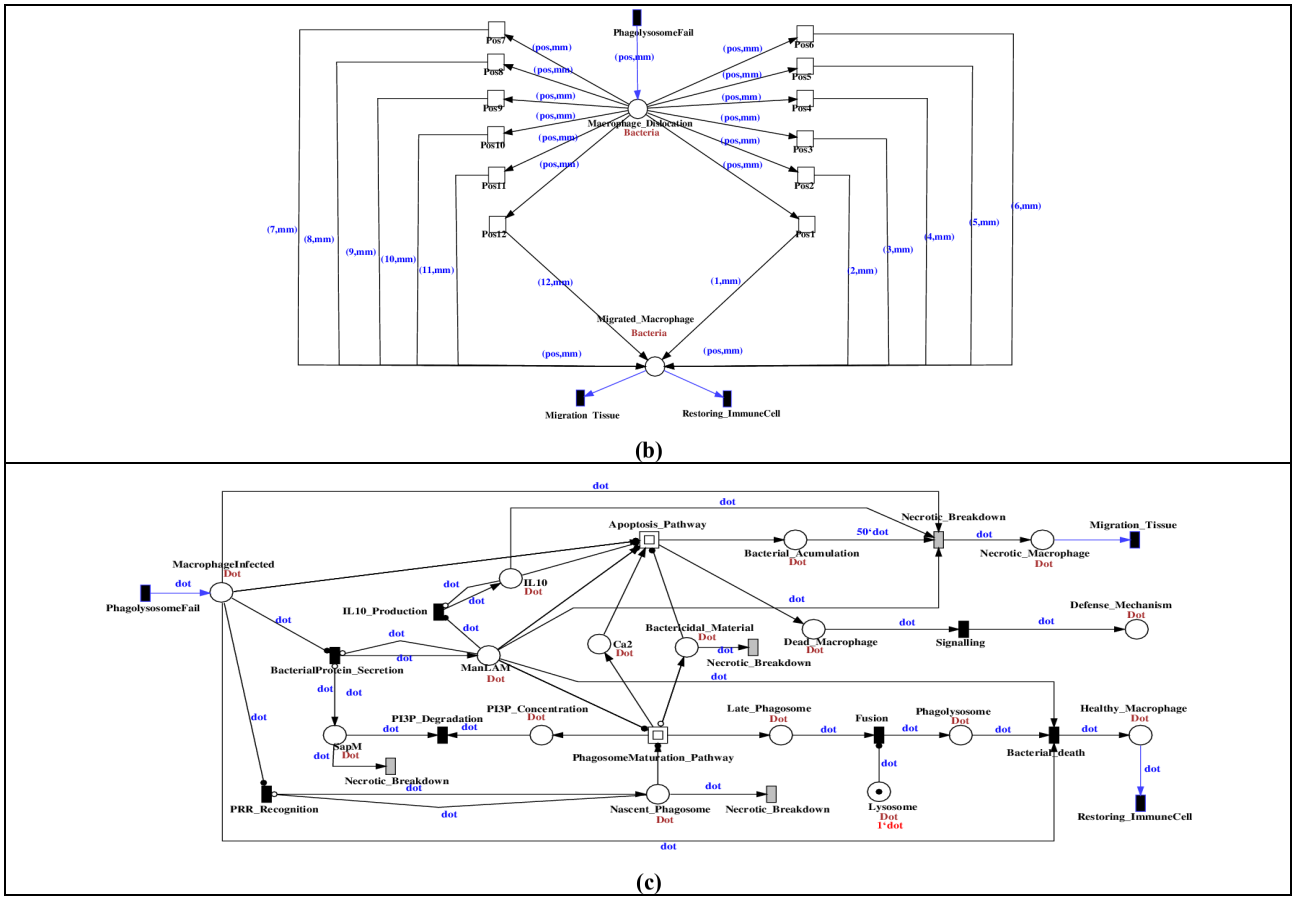


Figure 5.3 *Cont.*

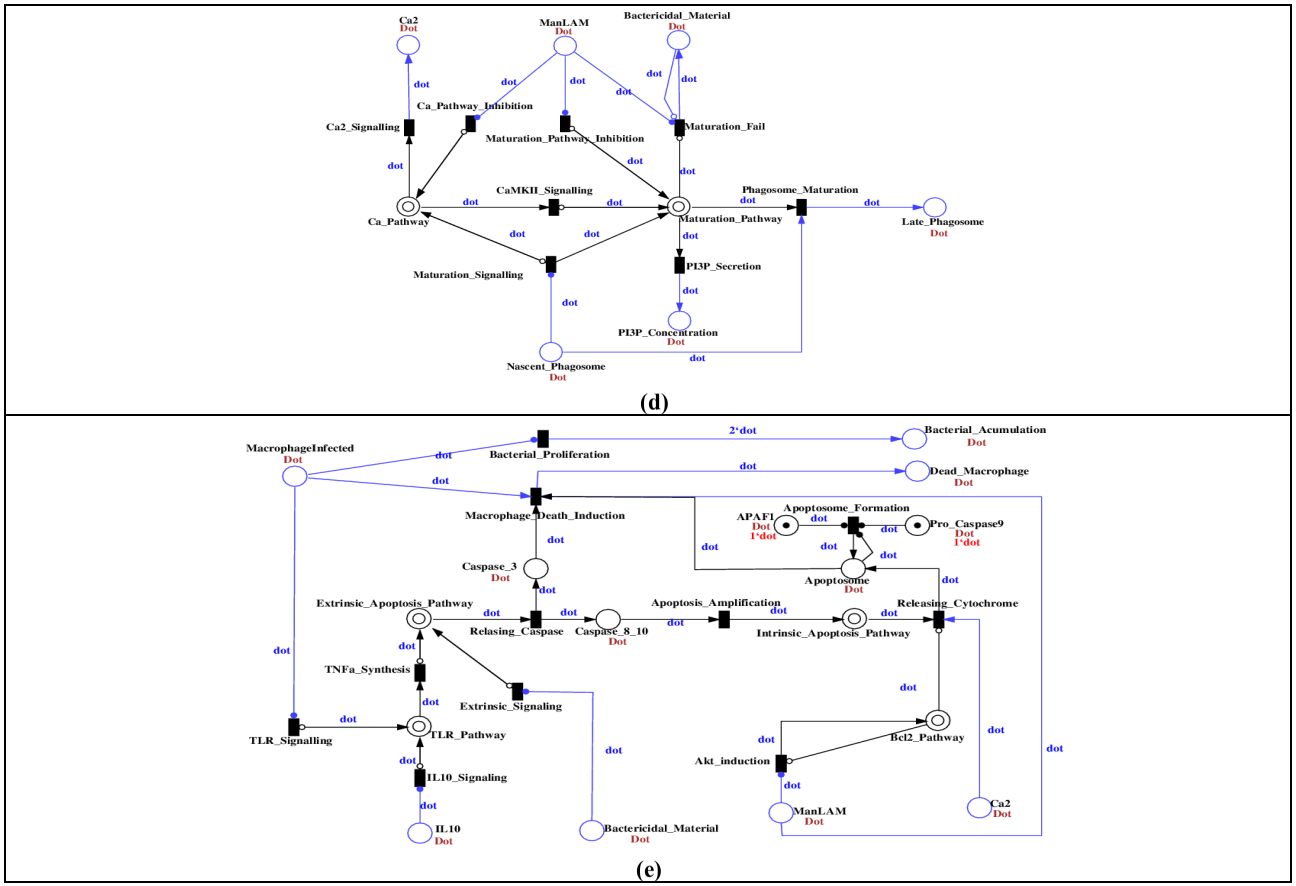


Figure 5.3 *Cont.*

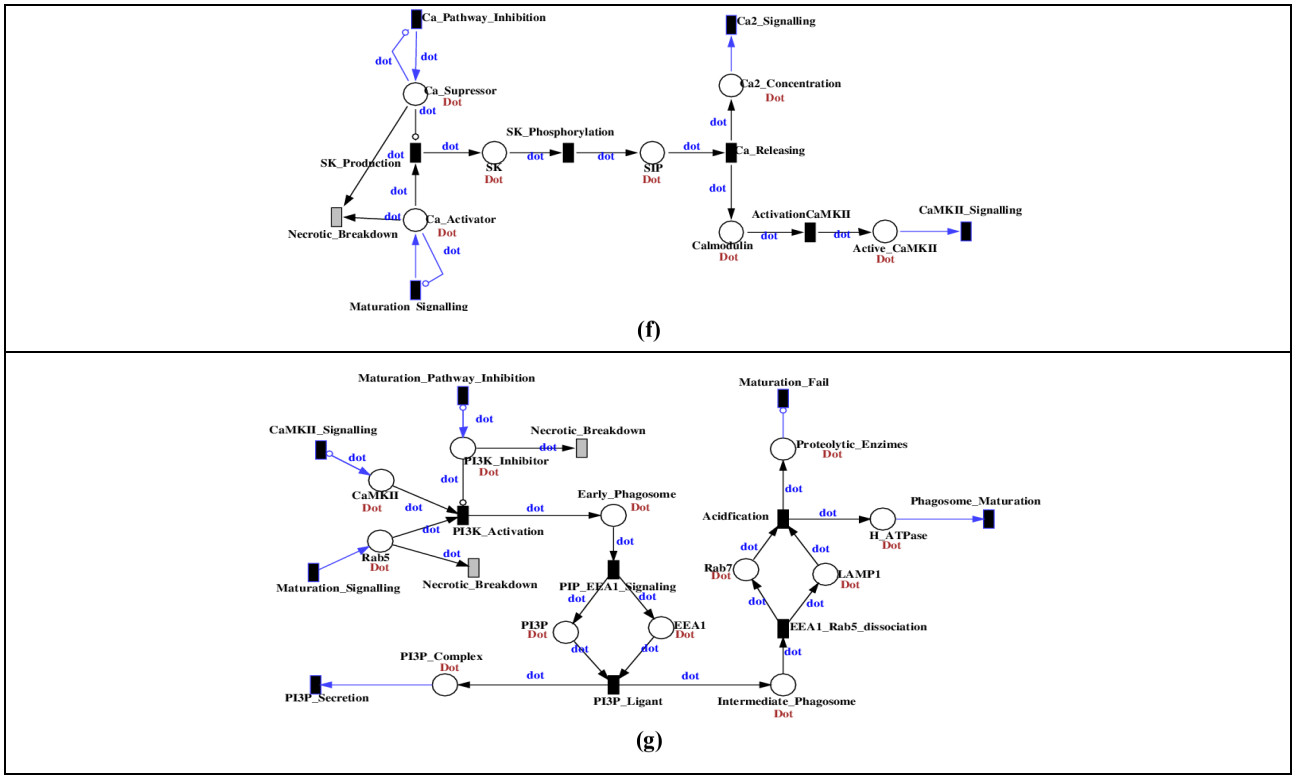


Figure 5.3 *Cont.*

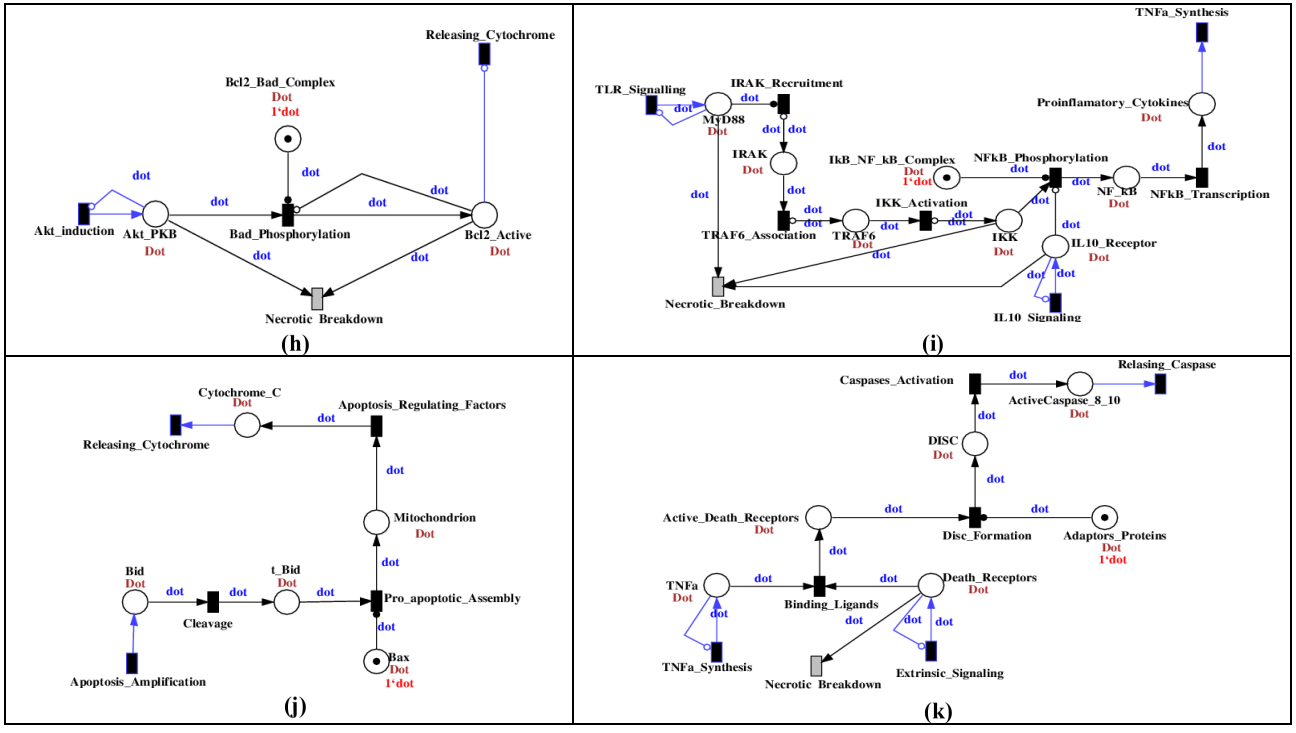


Figure 5.3 Model structure of the SPN^{C}. The black rectangles represent the immediate transitions and grey rectangles represent the logic nodes (fusion transitions). Stochastic transitions are represented by white squares. **(a)** Model the dynamic of the infection process. **(b)** Model the stochastic migration process **(c)** Model the bacteria-macrophage interaction at intercellular level. **(d)** Phagosome maturation pathway and **(e)** Apoptosis pathway at intracellular level. **(f)** Ca pathway; **(g)** Maturation pathway; **(h)** Bcl2 pathway; **(i)** TLR pathway; **(j)** Intrinsic apoptosis pathway and **(k)** extrinsic apoptosis pathway (**f**) to (**k**) are sub-models at molecular level.

5.2.3 Structure of the stochastic model

The hierarchical structure presented in Chapter 4 remains intact; the first layer net (top level) represents the dynamics of the infection process, while the sub-nets (levels 2, 3 and 4) represent the bacteria-macrophage interactions at intracellular, intercellular and molecular level. In the refinement process, we have changed the structure of the apoptosis pathway and its sub-models to better fit to the biological process (cf. **Figure 5.3 (e)**): The intrinsic apoptosis pathway (cf. **Figure 5.3 (j)**) is an amplification process triggered by the release of caspase from the extrinsic apoptosis pathway and cellular stress (cf. **Figure 5.3 (k)**). Moreover, the apoptosome is formed at the intracellular level (apoptosis pathway) which leads to the induction of macrophage death [68] (cf. **Figure 5.3 (e)**). We also removed the unnecessary accumulation of tokens in the sub-nets by adding logical nodes to consume those extra tokens in the simulation process. The logical nodes, a.k.a. fusion nodes, are copies of a single node used for the sake of readability of the net. They are features for the design, working as graphical copies with identical function. We use the transition **Necrotic_Breakdown** as a logical node on the sub-models in level 2 (intercellular model, cf. **Figure 5.3 (c)**) and in level 4 (molecular models, cf. **Figure 5.3 (f)** to **(k)**). Therefore, whenever this transition is enabled, it will consume all tokens on each level at the same time. With this solution, we have guaranteed that for each infected macrophage there is an individual host-pathogen interaction at the intracellular, intercellular and molecular levels. **Table 5.1** describes the places that previously accumulate tokens on each level and now are connected with the logical transition.

On the dynamic level (level 1), we have replaced the migration process, represented by the place **ActiveMacrophage** and the transition **Migration**, for a sub-net at level 2 implemented in the coarse place **MigrationProcess** (cf. **Figure 5.3 (b)**). This sub-net represents the migration of the infected macrophage towards the 12 regions according the data analysis presented in the Section 5.2.1. Therefore, the sub-net is composed of 12 stochastic transitions that specify the region where the infected macrophage will move to form the granuloma when it fires. The firing rule uses the kinetic mass action function and the firing rates are distributed based on the average distribution of infection previously analyzed (cf. **Figure 5.2**). The mass action function is used because it depends on the rate and not on the concentration of tokens. It follows the *mass action law* in which at a fixed concentration, the quantities (mass) have no influence upon the state of equilibrium [74]. It is not how much or how little of a substance is present that will be decisive, but how dense or sparse it is distributed in the space. Therefore, we correlate the distribution of the infection concentration along the 12 regions from the empirical data with the probability of the 12 transitions of the model to fire.

Table 5.1: Description of places connected through the logic node: **Necrotic_Breackdown**. Whenever this transition occurs, all the remains tokens are consumed.

Place	Sub-model (level)	Description
Nascent_Phagosome and **Bactericidal_Material**	**IntracellularInteraction (Level 2)**	Accumulates tokens that represent the signals triggering the phagosome maturation and the apoptosis process respectively (cf. **Figure 5.3 (c)**)
SapM	**IntracellularInteraction (Level 2)**	Accumulates tokens that represent the SapM protein released by the bacteria inside the macrophage (cf. **Figure 5.3 (c)**)
Ca_Activator and **Ca_Supressor**	**Ca_Pathway (Level 4)**	Accumulates signals (tokens) to the production/inhibition of calcium (cf. **Figure 5.3 (f)**)
Rab5 and **PI3K_Inhibitor**	**Maturation_Pathway (Level 4)**	Rab5 place accumulates maturation signals while PI3K inhibitor accumulates the ManLAM signals. (cf. **Figure 5.3 (g)**)
MyD88; IKK and **IL10_Repressor**	**TLR_Pathway (Level 4)**	MyD88 and IKK are proteins accumulated by the TLR signals and IL10_Repressor is accumulate signals to repress NFkB production (cf. **Figure 5.3 (i)**)
Akt_PKB and **Bcl2_Active**	**Bcl2_Pathway (Level 4)**	Akt protein is accumulated induced by presence of ManLAM while Bcl2 is accumulated by phosphorylation (cf. **Figure 5.3 (h)**)
Death_Receptors	**Extrinsic_Apoptosis_Pathway (Level 4)**	Accumulates the signals (tokens) to activate the receptors that bind with TNFa (cf. **Figure 5.3 (k)**)

5.2.4 Stochastic model extension

For a proper implementation, the set of color-sets Σ predefined in Chapter 4 also requires some modification. This involves the restructuring of the semantics of the color-set in order to make them suitable for the stochastic model. The SPN^C uses two sets of color-sets, the simple color-set described in **Table 5.2**, and the compound color-set described in **Table 5.3**.

Table 5.2: Simple color-sets modified for the colored stochastic Petri net model

Color-set	*Data type*	*Description*
Dot	dot	Default black color. Defined for all the places at the **IntracellularInteractions** sub-models
Position	Integer	Represents a location (of a macrophage, bacteria and/or granuloma) within the zebrafish embryo. It connects to the compound color-set.
count	Integer	Represents a threshold for the simulation in the recruitment of macrophage and calculate the amount of the dissemination.
BacterialType	String	Distinguishes the type of bacteria. Here represented by wild-type (**E11**) c.f. Section 5.2.1
ImmuneCellType	String	Distinguishes the type of immune cells. Here represented by **macrophages**.

Table 5.3: Compound color-sets modified for the colored stochastic Petri net model

Color-set	*Product type*	*Description*
Macrophage	**Position, ImmuneCellType**	Represents host macrophage immune cells
Bacteria	**position, BacterialType**	Represents the bacteria that are modeled
Granuloma	**position, ImmuneCellType, count**	Represents granuloma with a number of infected immune cells

There are no significant changes with respect to the set of places P and set of transitions T as previously defined in Chapter 4. Since we add a new sub-model implemented in the coarse place **MigrationProcess** (cf. **Figure 5.3 (b)**), we also add two places and 12 stochastic transitions to model the stochastic distribution of the infection. In order to extend our simulation experiments, we added the place **Granuloma_Distribution** to the dynamic model (cf. **Figure 5.3 (a)**). In this manner we are able to collect the information about the amount and position of the granulomas that were formed. Moreover, at this level we also add the transition **Restoring_Immunecell** to consume tokens from the place **Healthy_Macrophage** at **IntracellularInteractions** sub-model, cf. **Figure 5.3 (c);** and from the place **Moving_Macrophage** at **MigrationProcess** sub-model, cf. **Figure 5.3 (b)**. This transition represents the return of an infected macrophage to its functionality in case it has been able to digest the bacteria (phagolysosome process and bacterial death). In this scenario, the macrophage becomes healthy and no granuloma formation will take place. In **Table 5.4** and **Table 5.5** the set of places and transitions added to our model are described respectively.

Table 5.4: Set of places *P* added to the stochastic model

Place	*Color-set*	*Description*
Moving_Macrophage	**Bacteria**	Represents the macrophage in the moving process. It contains the initial infected macrophage that will move within the bacteria.
Migrated_Macrophage	**Bacteria**	Represents the infected macrophage repositioned into a region where it will leave the circulatory system and form the granuloma.
Granuloma_Distribution	**Position**	This place is used to keep track on the spatial distribution of the infected macrophage along the 12 regions.

Table 5.5: Set of transitions *T* added to the stochastic model

Transition	*Type*	*Rate function*
Pos1	Stochastic	MassAction(0.106042565)
Pos2	Stochastic	MassAction(0.165508512)
Pos3	Stochastic	MassAction(0.196383367)
Pos4	Stochastic	MassAction(0.056442759)
Pos5	Stochastic	MassAction(0.071513451)
Pos6	Stochastic	MassAction(0.148477928)
Pos7	Stochastic	MassAction(0.121384737)
Pos8	Stochastic	MassAction(0.057348882)
Pos9	Stochastic	MassAction(0.030160435)
Pos10	Stochastic	MassAction(0.015993484)
Pos11	Stochastic	MassAction(0.021050388)
Pos12	Stochastic	MassAction(0.00969349)
Restoring_Immunecell	Immediate	-

The initial marking *I* also does not significantly change compared to the qualitative model from Chapter 4. The condition markings remain the same, whereas the example marking regarding the amount of bacteria changed so as to represent the concentration of bacteria inoculated in the *in vivo* experiments. We defined that **125** tokens represent the amount of wild-type bacteria (**E11**) injected at position **12**. This value is derived from the average concentration of bacteria inoculated at the caudal region according to our empirical data (c.f. Section 5.2.1). If we change the number of tokens that represent the number of bacteria, it will influence in the total of granuloma that will be formed, i.e. the more bacteria, the more macrophage will phagocytosis and become infected; the relation is one to one (one macrophage to one bacterium). This effect does not occur with the initial marking related to the total number of macrophage. The place **ImmuneSystem** works as an infinite resource, it always contains macrophages in all positions. All the transitions connected to this place do not consume its token, only check if it has a macrophage in a specific position by the read arcs. Therefore, adding more macrophage in a specific position does not influence in the amount of granuloma, although, if there is no macrophage in a specific position, the bacteria that eventually

have settled in this position, will not be phagocytosis by the immune cell. The same process occurs for an infected macrophage that eventually migrates to a position without a macrophage; it will not form a granuloma. In **Table 5.6** the example markings initially defined for the stochastic model are described.

Table 5.6: Example markings initially defined in the colored stochastic Petri net model

Place	*Marking*	*Description*
Infection	**125`(12,E11)**	Defines the initial concentration of mycobacteria that will intrude the host.
ImmuneSystem	**1`(1,mac)++** **1`(2,mac)++** **1`(3,mac)++** **…** **1`(12,mac)**	Defines the initial position of non-infected macrophages in the host. We considered that macrophage can be present in all the 12 regions and the first macrophage to arrive at the infection site is and the nearest macrophage positioned

5.2.5 Simulation environment

Now that our SPN^C model has been set up, we simulate the stochastic process in order to reproduce the behavior of the mycobacterial infection process and innate immune response compared to the *in vivo* experiments. Snoopy software [50] provides a simulation environment for stochastic models. In this environment, it is possible to set-up the initial markings for the places; the function rates defined for the stochastic transitions; apply weights for immediate transitions. Moreover, it is also possible to specify the simulation time interval and the number of time points in the time interval that the information is recorded. In addition, the number of runs of the simulation process can be defined. The result is determined from calculation of the average values over all runs. The output of the simulation can be visualized in the simulation environment in different graphical formats (e.g. scatter plot, histogram, tabular) and also can be exported to a file (csv format).

5.3 Simulation and results

We perform stochastic simulations to test how well the model captures the dynamics of the system and how changes might affect the behavior of the infection process. Two simulation scenarios were performed reproducing the dissemination of the infection. We use the results of the data analysis from Section 5.2.1 as input parameter for our simulations as well as a means to compare the end result. The output of the simulations is the number of tokens accumulated at the place **Granuloma_Distribution** along the simulation time. The number of tokens represents the average granuloma formation on each position. The time is simulated in days, from the inoculation of the

bacteria (day 0) till the 5th day post infection (5 *dpi*), according to our data analysis from the zebrafish images. Hereafter we present the simulation scenarios and their outcome.

5.3.1 Stochastic simulation: ideal scenario for bacterial burden

In our first simulation, we consider that each macrophage that phagocytosis bacteria is not able to digest the bacteria (phagolysosome process) and neither suppress the infection (apoptosis process). Therefore, the infected macrophage migrates to form granuloma in one of the 12 regions, according to the probability distribution based on the rates given in **Table 5.5**. The granuloma disseminates the infection by releasing infected macrophages that will form more granulomas along the regions.

In this scenario, the infection behavior is simulated from the time point the bacteria are introduced in zebrafish embryo (time 0), till the time the zebrafish larva were screened by taking images at 5 *dpi*. The place **Granuloma_Distribution** contains the information about the total of granulomas that are formed per region. We plot the average of the number of tokens accumulated during the simulation time per region. **Figure 5.4** and **Figure 5.5** depict the results. In **Figure 5.4 (a)** we depicted the accumulation of the infection, as average of the number of tokens per region (**Macrophage_Distribution_1** till **Macrophage_Distribution_12**) in 5 days, for one random specimen of zebrafish (1 simulation run). We performed 138 simulation runs within one experiment in order to compare with the number of samples (138 zebrafish) from our empirical data. In **Figure 5.4 (b)** we depicted the average value of number of granuloma accumulated over 138 runs. It represents the accumulation of the infection per region during the 5 days from 138 zebrafish. **Figure 5.5** presents the percentage of distribution per region (**Pos1** till **Pos12**) for both simulations (1 run and 138 runs). It is possible to verify the similarity of the distribution and the behavior of the system by comparing the simulation results with the empirical data. Most of the infection regions in the simulations are concentrated in the head (position 2 and 3) and also at the center of the zebrafish (position 6 and 7), which is similar to the *in vivo* experiments. For one simulation run, we have a total of 224 granulomas formed, while the 138 simulation runs totalized an average of 221.88 granulomas formed. The infection distribution grows exponentially, reaching a steady state during the run time. This means that, for this model, there will be hardly any difference witch simulations for 6, 7 *dpi*. However we cannot predict the same behavior for the *in vivo* experiments since there is no data yet that report such behavior. The tendency of the simulations to reach the steady state in the number of tokens (infection distribution) is reflected in the single parameter set for the stochastic properties of the model. For this simulation we only have considered the probability for the distribution of the infection, whereas the other factors are considered to occur at as instantaneous events (immediate transition) and have no stochastic influence on the system. If we add more stochastic influence in the system, a different behavior might occur. Next, we simulate a second

scenario exploring a different simulation set in order to approximate the model to the *in vivo* experiments results observed in our empirical data.

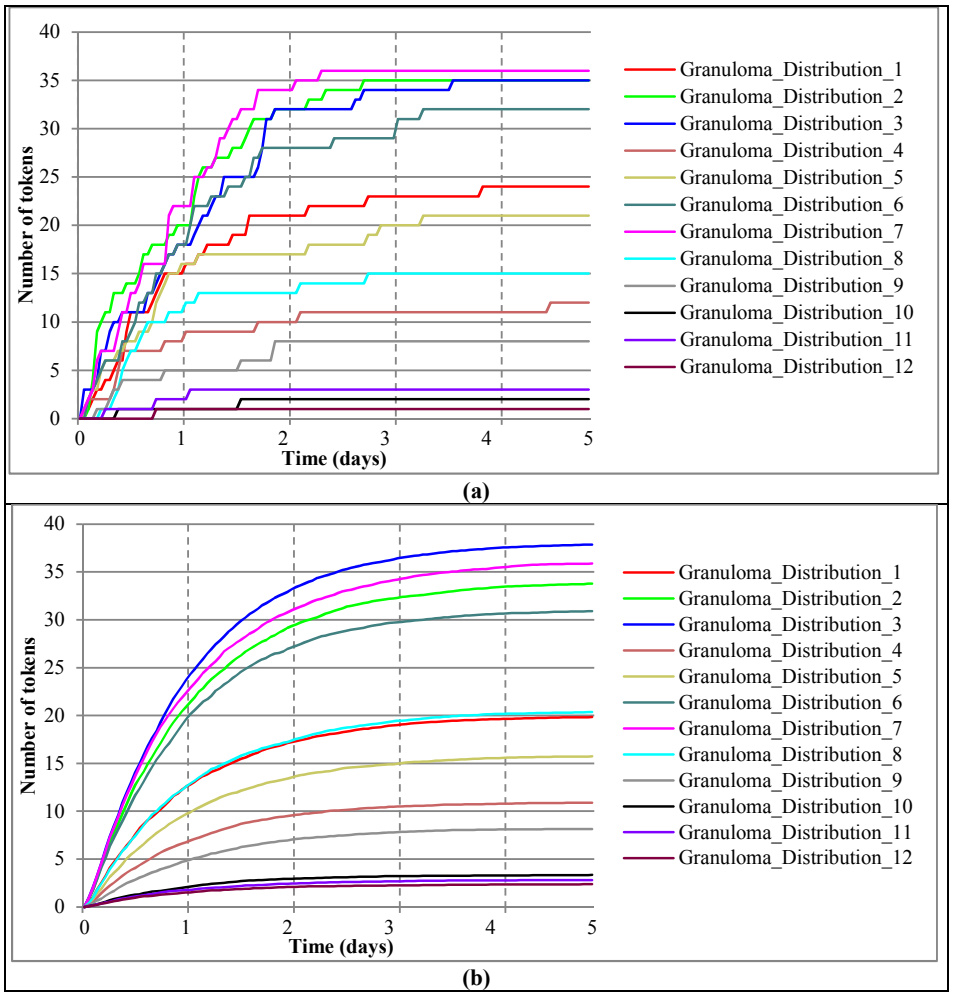
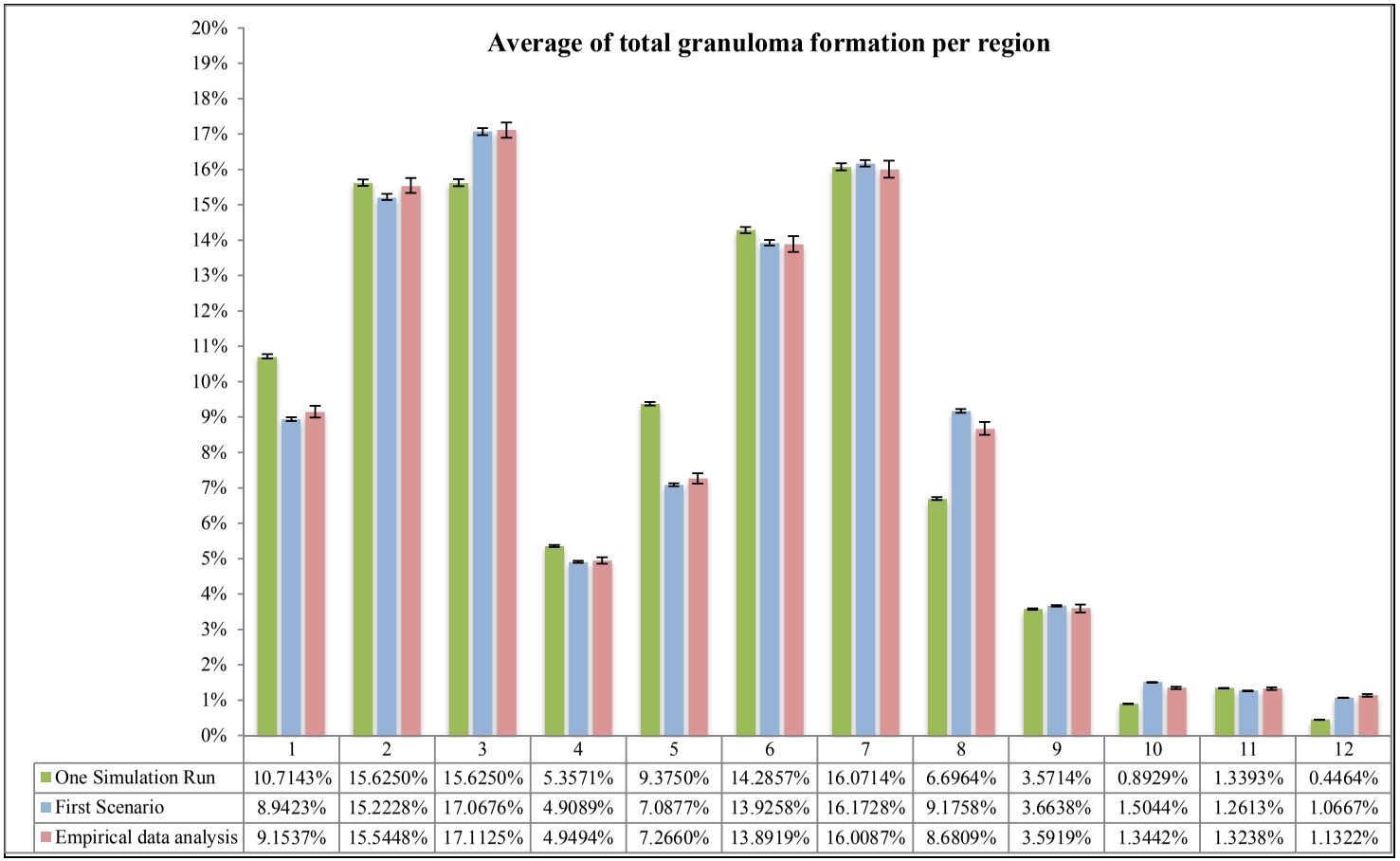


Figure 5.4 Simulation results of the stochastic Petri net model: ideal scenario for bacterial burden. **(a)** Infection distribution over time (5dpi) for one simulation run and **(b)** for an average of 138 simulation runs.



	1	2	3	4	5	6	7	8	9	10	11	12
One Simulation Run	10.7143%	15.6250%	15.6250%	5.3571%	9.3750%	14.2857%	16.0714%	6.6964%	3.5714%	0.8929%	1.3393%	0.4464%
First Scenario	8.9423%	15.2228%	17.0676%	4.9089%	7.0877%	13.9258%	16.1728%	9.1758%	3.6638%	1.5044%	1.2613%	1.0667%
Empirical data analysis	9.1537%	15.5448%	17.1125%	4.9494%	7.2660%	13.8919%	16.0087%	8.6809%	3.5919%	1.3442%	1.3238%	1.1322%

Figure 5.5 Percentage of distribution per region (Pos1 till Pos12) for both simulations (1 run and 138 runs) compared to our empirical data analysis.

5.3.2 Stochastic simulation: macrophage response to infection

In the studies on mycobacterial infection, Armstrong and Hart [5] were the first to identify the bacterial influence in the phagolysosome. They performed *in vitro* experiments with cultured macrophages infected with *Mycobacterium tuberculosis (Mtb)*. They showed that approximately 70% of the infected macrophages display failure in the process of fusion of the phagosome with the lysosome. Although their experiments were with *Mtb*, in our simulations, we assume that *Mycobacterium marinum* behaves similar and used this particular fact in our second scenario. We add a probabilistic factor of the macrophage to perform the phagolysosome process, degrading the bacteria, becoming healthy, complying with its immune properties, and thereby reducing the virulence spread. In this manner, we consider to reproduce the infection process in a more realistic manner.

To perform this simulation, we added a stochastic transition at **IntracellularInteractions** sub-model (cf. **Figure 5.3 (c)**) which will fire a token to the place *P*={**Late_Phagosome**} indicating the maturation process has occurred. This stochastic transition will fire according to the **MassAction** function that will indicate the probability of the phagosome maturation take place in an infected macrophage. For this scenario we performed 3 simulations with different probabilities (rates) for a macrophage to perform the phagolysosome process. In simulation one, we consider the probability of 70% of chance for a macrophage failure in the phagosome maturation process, become infected and disseminate the infection, forming granuloma as observed in [5]. Therefore, a macrophage has 30% of chance to kill the bacteria by performing the phagolysosome process and becoming healthy. For the simulation two and three we arbitrary extrapolate the system, by considering 40% of chance for an infected macrophage form a granuloma (60% of chance to kill the bacteria) and 10% of chance for an infected macrophage form a granuloma (90% of chance to kill the bacteria) respectively. The reason for these rates is to analyze the behavior of the system in different scenarios, where the immune cell is able to respond to the infection process, as observed by Armstrong and Hart [5]. In **Table 5.7** we describe the rates (probability) of phagolysosome manifestation assigned for the transition for each simulation.

Table 5.7: Rates assigned to the **MassAction** function indicating the probability of a macrophage degrade the bacteria

Simulation	*Function*	*Description*
Simulation 1	**MassAction(0.3)**	Indicate there is a 30% of chance to a macrophage perform the phagolysosome and degrade the bacteria (become healthy)
Simulation 2	**MassAction(0.6)**	Indicate there is a 60% of chance to a macrophage perform the phagolysosome and degrade the bacteria (become healthy)
Simulation 3	**MassAction(0.9)**	Indicate there is a 90% of chance to a macrophage perform the phagolysosome and degrade the bacteria (become healthy)

We performed the three simulations under the same initial condition of the First Scenario, using the average of granuloma distribution per region from 138 simulation runs (cf. Section 5.3.2). In **Figure 5.6** we depict the average of granuloma formed for the three simulations. Moreover, in **Figure 5.7** we plot the average of concentration of granuloma per region formed during the simulation, compared with the First Scenario and the empirical data.

Analyzing the results from the simulations and the empirical data we can conclude that the model behaves similar to the biological experiments from the data analysis. Despite the bacteria CFUs are injected in the tail, a high concentration of granuloma is formed near the head and the center of the fish (regions 2, 3 and region 6, 7 respectively). We also noticed that the average of granuloma formation diminishes once we increase the probability of a macrophage to perform the phagolysosome and kill the bacteria. The Simulation 1 presented an average of 30.49 granulomas formed, simulation 2 resulted in 4.79 and simulation 3 presented an average of 1.60 granulomas considering 138 simulation runs in 5 days (cf. **Figure 5.6**).

Despite the fact that the amount of infection (number of tokens) has been less than the previous simulation, we can notice that the distribution of the granulomas slightly differentiates from the First Scenario and the empirical data. The higher the probability of a macrophage to be able to degrade the bacteria, the more the percentage of granuloma distribution per region from the empirical data differs. The percentage of distribution for the simulation without probability of phagolysosome (First Scenario) and with 30% of chance for the phagolysosome (simulation 1) are more similar to the empirical data, compared with the simulations of 60% and 90% of chance of a macrophage to degrade the bacteria. This is logical because we this to be part of the empirical data. In the next section, we perform a statistical analysis to assess the confidence interval and hypothesis on the similarity of the percentage of the granuloma distribution from our simulation experiments and the empirical data.

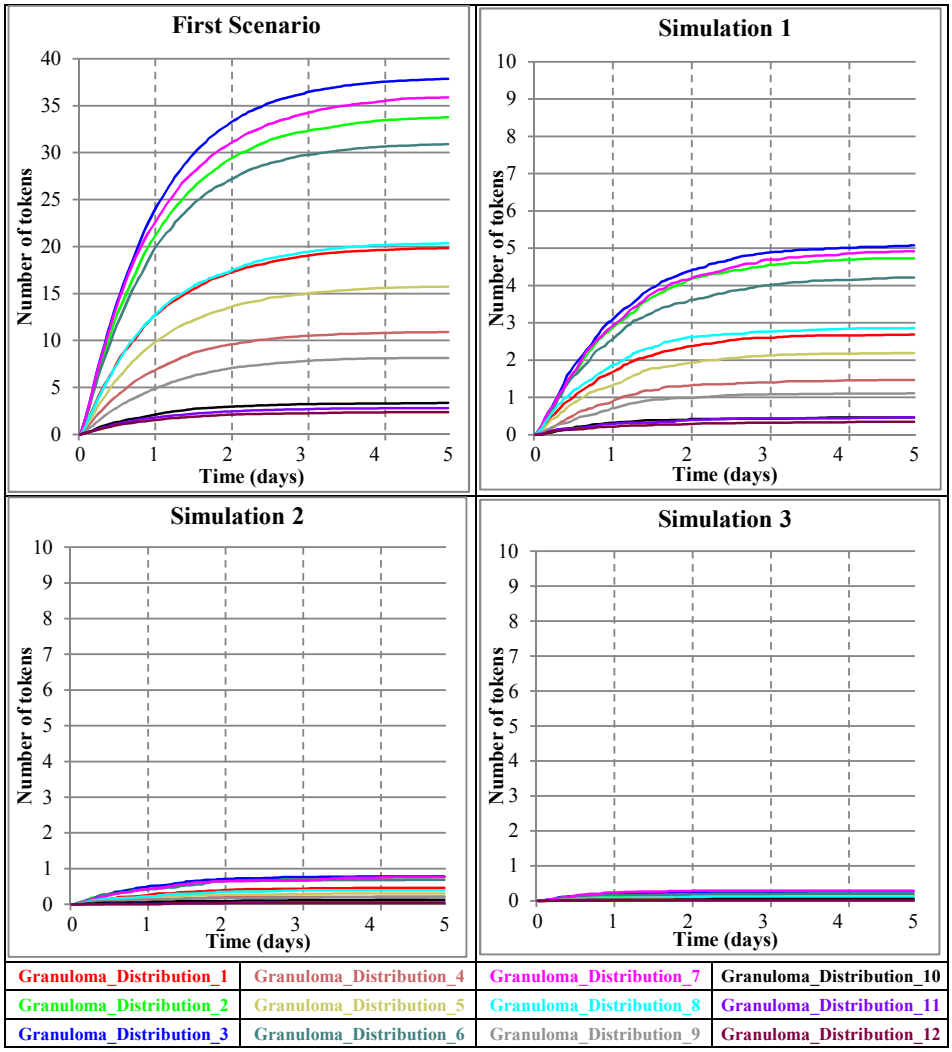
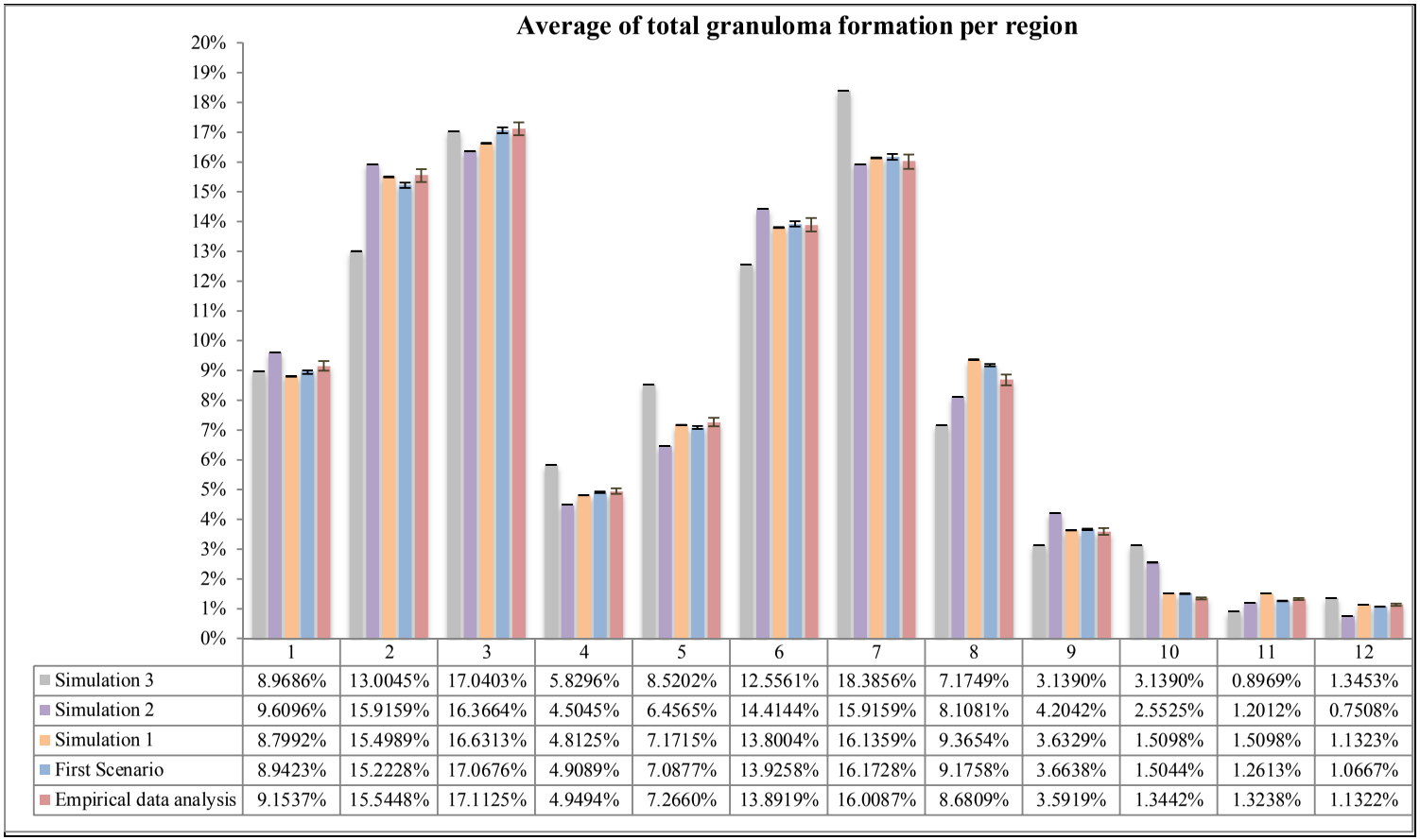


Figure 5.6 Simulation results of the stochastic Petri net model: **First Scenario** average of 221.88 granulomas formed; **Simulation 1** average of 30.49 granulomas formed; **Simulation 2** average of 4.79 granulomas formed; **Simulation 3** average of 1.60 granulomas formed.



	1	2	3	4	5	6	7	8	9	10	11	12
Simulation 3	8.9686%	13.0045%	17.0403%	5.8296%	8.5202%	12.5561%	18.3856%	7.1749%	3.1390%	3.1390%	0.8969%	1.3453%
Simulation 2	9.6096%	15.9159%	16.3664%	4.5045%	6.4565%	14.4144%	15.9159%	8.1081%	4.2042%	2.5525%	1.2012%	0.7508%
Simulation 1	8.7992%	15.4989%	16.6313%	4.8125%	7.1715%	13.8004%	16.1359%	9.3654%	3.6329%	1.5098%	1.5098%	1.1323%
First Scenario	8.9423%	15.2228%	17.0676%	4.9089%	7.0877%	13.9258%	16.1728%	9.1758%	3.6638%	1.5044%	1.2613%	1.0667%
Empirical data analysis	9.1537%	15.5448%	17.1125%	4.9494%	7.2660%	13.8919%	16.0087%	8.6809%	3.5919%	1.3442%	1.3238%	1.1322%

Figure 5.7 Simulation results of the stochastic Petri net model using Snoopy: 70% of infected macrophage will form granuloma. Percentage of distribution per region compared with previous simulation (blue bar) and the data analysis (red bar)

5.3.3 *Analysis of the results*

A statistical analysis on the average of granuloma formation per region was performed, verifying the similarity of the percentage of distribution of each simulation with the analysis of our empirical data (cf. Section 5.2.1). The idea is to check if the result is statistically similar to what the plot in **Figure 5.7** shows, or if what we see is a coincidence of the stochastic process. Therefore, we use a one sample hypothesis test (t-test) since we are dealing with the average of the infection distribution per region. The t-test is the most commonly used confidence interval and hypothesis test of population means [13]. It is used to test whether the mean of a population takes a particular value, for repeated measurements (which is our case). It requires the sample to be independent, normally distributed and equal in sample size and variance.

We applied the t-test in the simulation of the First Scenario, in which an infected macrophage does not perform the phagolysosome process (0% chance to kill the bacteria); also in the simulations of the second scenario, in which an infected macrophage has a chance to perform the phagolysosome process (simulation 1 with 30% of chance to kill the bacteria, simulation 2 with 60%, and simulation 3 with 90%). In **Table 5.8** the t-test results for each simulation are given. It presents the percentage of similarity between the averages of concentration of granuloma per position similar to the analysis of our empirical data.

Table 5.8: One sample t-test result: probability of each position been similar to the empirical data

Simulation	*Pos1*	*Pos2*	*Pos3*	*Pos4*	*Pos5*	*Pos6*	*Pos7*	*Pos8*	*Pos9*	*Pos10*	*Pos11*	*Pos12*
First Scenario	88.07%	86.00%	98.07%	95.82%	88.62%	98.59%	93.71%	75.28%	94.01%	54.69%	84.76%	83.19%
Simulation 1	80.12%	97.99%	79.51%	85.91%	93.95%	96.18%	95.12%	66.31%	96.58%	53.34%	56.81%	99.96%
Simulation 2	74.61%	83.90%	68.72%	56.43%	51.64%	78.48%	96.44%	71.55%	52.20%	0.00%	70.64%	21.75%
Simulation 3	89.54%	16.57%	96.89%	25.48%	31.53%	48.54%	25.40%	33.86%	63.56%	0.00%	19.11%	48.99%

In **Figure 5.8** the confidence levels from the t-test for each simulation are depicted according to their position in the zebrafish. A high confidence level, i.e. similarity > 95%, demonstrates that in that particular region there is no statistically significant difference between the simulation and the empirical data (indicated by color blue). A mid confidence level, i.e similarity in >80%, <95%, demonstrates that in that region there is an acceptable significant difference between the simulation and the empirical data (indicated by color gray). A lower confidence level (similarity < 80%) demonstrates that in that region there is a statistically significant difference between the simulation and the empirical data (indicated by color red). Analyzing the results, we can notice that Simulation 1, i.e. a Macrophage with 30% chance to kill the bacteria, has more similarities in the distribution of the granuloma per position than the other simulations. We can correlate these results with the *in*

vitro experiments performed by Armstrong and Hart [5], indicating that indeed not all the macrophages that phagocytosis bacteria will generate a granuloma.

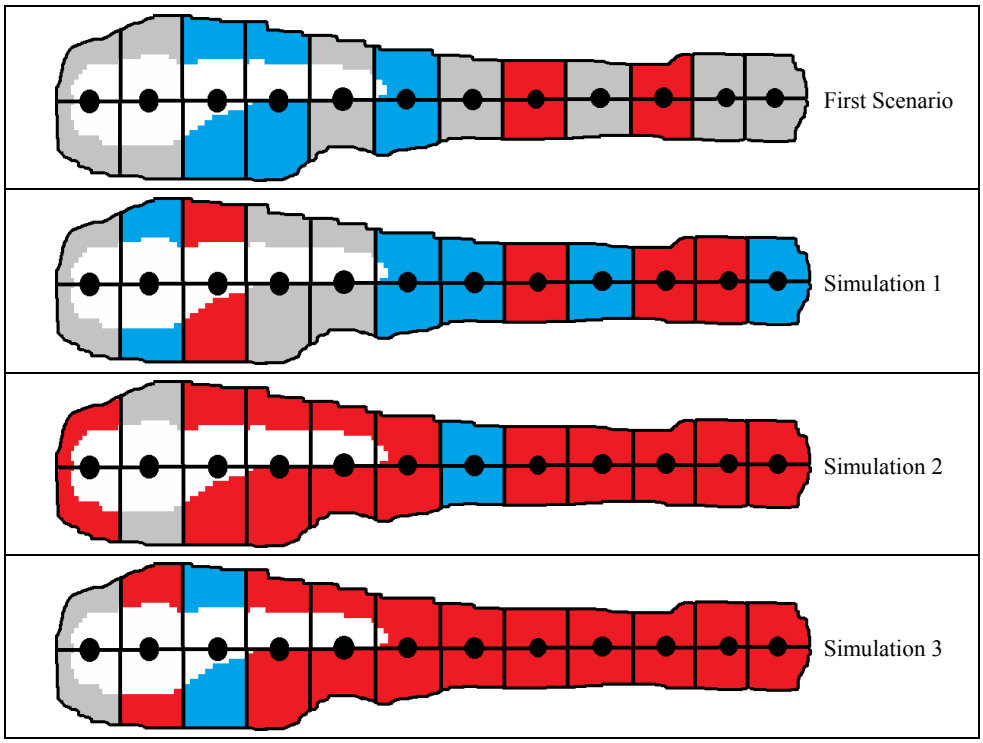


Figure 5.8 Confidence level per region of the t-test result for each simulation: in blue high confidence level (bigger than 95%); in gray mid confidence level (between 95% and 80%); in red the low confidence level (lower than 80%)

5.4 Conclusion and discussion

In this chapter, we have used a colored stochastic Petri net to model and simulate the mycobacterial infection process. To that end, we had to redesign our previous QPN^C model presented in Chapter 4 making it a quantitative model, in order to simulate scenarios that can reflect the behavior of the infection process. The objective was to convert our QPN^C model into a colored stochastic Petri net model using information we have obtained from *in vivo* experiments as the basis for the simulation and model validation.

We have analyzed data from biological experiments on *Mycobacterium marinum* (*Mm*) infection in zebrafish embryos. We used a dedicated image analysis software package [89, 90, 113] to measure the "spread" of the *Mm* bacteria after 5 *dpi*. The image analysis software has shown to be an indispensable tool in the modeling process. It provides information about the granuloma concentration along a pre-defined set of 12 regions within the zebrafish. We organized the data as a

proportion of infection per region (cf. **Table 5.2**). This information was used in our model as a probability rate for an infected macrophage to form granuloma on a specific position.

We redesigned the colored Petri net model into a SPN^C model taking in account the specified position from where the infected macrophage migrates to form the granuloma. We implemented a sub-model with stochastic transitions, addressing the position of each region granuloma are formed, as described in the empirical data (cf. **Figure 5.3 (b)**). We decided to use stochastic transitions only to model the process related to the infected macrophages migration, as that corresponds to the information extracted from *in vivo* experiments. The other transitions of the model are implemented as immediate transitions to represent processes present in the model but not explicitly measured in the empirical data. The use of immediate transitions implies no time consumption (delays) for a specific event (transition) to happen, and they have higher priority over the stochastic transitions. We could have used the stochastic transition with a rate function of 1.0 (100% probability to occur) but the addition of a time delay for each transition would skew the result. The motivation for using immediate transitions was to keep control over the transitions of no stochastic information available, providing more precise simulation results. In this manner, we are able to gradually increase the model by replacing the immediate transitions for stochastic transitions according the availability of data or to perform controlled simulations experiments.

We have performed two simulation scenarios with the new model, and validated them by comparing the results with the data from *in vivo* experiments.

In the first simulation scenario, we considered that all infected macrophages could form a granuloma in a specific region according to the rate function defined from the data analysis. This could be the ideal scenario for the bacteria since all inoculated *Mm* bacteria would survive and spread the infection over the time. We simulated the behavior of the infection for one zebrafish and for 138 zebrafishes, comparing the results with the distribution in our empirical data. The analysis of the results showed that the infection tends to reach a steady state over the time (cf. **Figure 5.4**). We assume that this behavior is related to the deterministic aspects of the immediate transitions. Despite of this, the granuloma distribution remains similar to the analysis of the empirical data, which corresponds to the biological behavior of the *in vivo* experiments. Therefore, this result validates our model.

For the second scenario, we have added another stochastic process to represent a natural factor that influences the infection process. This process consists of a stochastic transition that adds a probability of an infected macrophage to perform the phagosome maturation. For this scenario we performed three more simulations, considering that an infected macrophage has 30%, 60% or 90% of chance to successfully perform the phagolysosome process, hence kill the bacteria and avoid the formation of granuloma. This strategy could be accomplished by adding quantitative information on the calcium regulation and maturation process, by changing their respective pathways in the model.

However, such measured information was not available from biological experiments or through literature. Therefore, for convenience, we connected a stochastic transition to the late phagosome indicating that it has a specific probability to occur according to each simulation. The result was a reduction of the infection (amount of granuloma), expressed by the number of tokens accumulated in the place **Granuloma_Distribution**. This result should be expected since, for each simulation, we have increased probability of one infected macrophage to kill the bacteria and consequently, not forming a granuloma. By analyzing the concentration of the granulomas per region for each simulation, we noticed that the behavior of the distribution of the infection remained similar compared to the analysis of the empirical data. Although, we found that the percentage of granuloma concentration per region (position) for the simulation, where a macrophage has 30% chance to kill the bacteria, was more accurate than the first scenario and the other simulation - 60% and 90% chances of phagolysosome - (cf. **Figure 5.7**). We corroborate this result by performing a t-test to statistically verify the similarity between the simulations and the empirical data cf. **Table 5.8** and **Figure 5.8**). We can correlate these findings with the results found by Armstrong and Hart on their *in vitro* experiments [5]. Although their experiments where in tuberculosis and we do not have information of such experiment replicated *in vivo* for *Mm* with zebrafish, we can, to a certain extent, validate our model.

There are many natural factors that influence the infection in the biological process for which at the moment no measurements are available. Having demonstrated the robustness of our model in the simulation scenarios, we can extrapolate the model in order to approximate the results with its biological behavior. It is possible to add more complexity by replacing the immediate transitions for stochastic transitions and using results from *in vivo/in vitro* experiments as data input. An example of such possible extension is modeling the influence of calcium in the phagolysosome maturation. As we demonstrated before, 30% of infected macrophages are able to degrade the bacteria through this process. Future simulations could focus on the level of calcium that can influence in this percentage and that can suggest new *in vivo/in vitro* experiments to validate the results. In addition, we can perform the animation of the model in the 3D environment using the 3D visualization tool introduced in Chapter 4. The 3D tool was designed to read qualitative models, modeled using colored Petri nets. Therefore, a necessary adaptation to read the output of the stochastic simulation is required in order to express the quantitative aspects from the model.

In summary, in this chapter, we analyzed data from biological experiments on *Mycobacterium marinum* infection in zebrafish embryos. This information was used to redesign our previous model to obtain a quantitative model. We performed simulations on the new model and validated it by comparing the empirical data from *in vivo* experiments. We demonstrated the robustness of our model by performing different simulation scenarios, comparing them with the behavior observed from the experiments and suggested further model refinement. The process of validation and

refinement contribute to identify important parameters that can help to unravel mechanisms related to the biological process.