Chapter 2

INDISIM-RBC: model of *P. falciparum* infected RBCs in *in vitro* cultures

In vitro cultivation of *P. falciparum* is a custom practice carried out with mastered techniques that were set thirty years ago. Yet, many limitations that make malaria harvesting costly and tenuous are accepted without examining whether they could be avoided with alternative cultivation methods. *Plasmodium falciparum in vitro* cultures (see Section 1.3) have been modelled using INDISIM-RBC, an Individual-based model (see Section 1.4) that has been framed up with the information on the RBC and the parasite found in literature (see Section 1.2).

This chapter describes the working methodology (Section 2.1) and the specific model (Section 2.2) used to tackle malaria spread *in vitro*. It also includes the results obtained with the 2D version of the model (Section 2.3), which is especially concerned with the processes affecting individual cells. These results refer to specific problems detected in real systems and reproduce observed phenomena and trends. The chapter ends with a brief discussion on the obtained results and the loose ends that should be tackled in further chapters (Section 2.4).

2.1 Background and general outline of INDISIM

2.1.1 Basis of INDISIM

The acronym INDISIM stands for INDividual DIScrete SIMulation. It refers to a modelling methodology and also to the software developed in Fortran to examine the outcome of particular models through computer simulations. INDISIM is specifically designed to study microbial communities and their environment through a mechanistic approach.

INDISIM has features typical of the Monte Carlo Methods (MC). This term describes a large class of computational algorithms that use repeated random sampling and reckoning to obtain their results. MC simulations are especially useful to address systems with a large number of interacting particles and a good tool to study stochastic phenomena (see Section 1.4).

Both INDISIM and MC require the use of large sets of random numbers. As truly random numbers are only obtained through measurement of probabilistic processes, computers use pseudorandom sequences. They are long strings of integers that exhibit the statistical properties of randomness, at least to the extent required by the simulations. They can be generated with an algorithm out from an initial value called the random seed. Two simulations with the same initial configuration and random seed will produce exactly the same results. Modifying the random seed can produce different results because the stochastic contributions to the simulations are altered (Metroppolis and Ulam., 1949).

The model is also related to the ones used in Molecular Dynamics (MD). MD comprises those computational models in which interacting atoms (or molecules) are allowed to move in discrete events, following the model rules, to give rise to the long-term evolution of systems composed by a large number of particles (Alder and Wainrwright, 1959). Both methods simulate the evolution of dynamical systems based on statistical mechanics.

However, the methodology of INDISIM differs from the MD approach in several main aspects. Firstly, the main entities in MD are molecules while INDISIM deals with cells. This means that the processes represented by each approach cover different spatial and temporal scales. Secondly, MD typically uses deterministic rules while INDISIM depicts stochastic processes and uses probability distributions. Thirdly, MD typically deals with systems in the thermodynamic equilibrium (or in the steady state) while INDISIM simulates open systems that are far from the thermodynamic equilibrium. This list of divergences could be furthelry extended. Although INDISIM originally arose from MC and MD models, it has its own specific features.

For this reason, the model is better framed in the ambit of System Dynamics, the

branch of physics that tackcles generic complex systems. In this context, INDISIM is classified as an Individual-based Model (IbM, a term mainly used in theoretical ecology) or as an Agent-based Model (AbM, a term mainly used in social sciences). A thorough discussion on the application of the IbMs approach to microbial populations is detailed in Chapter 4.

The root of the simulator is the program Barcelonagram (Valls, 1986; Giró et al., 1986). It was initially employed to better understand the normalized biomass distribution function among populations growing under certain constraints: the weight of individuals in a population undergoing unrestricted growth is distributed according to a lognormal distribution function (Koch, 1966). Such shape observed in real populations can also be theoretically obtained from the maximisation of biological diversity among the bacterial population (Wagensberg et al., 1988b). Finally, simulations can be used to show how the theoretical assumptions give rise to the observed behaviour (Wagensberg et al., 1988a). The methodology proved an appropriate tool to address the theoretical study of complexity when Solé et al. (1992) used a version of Barcelonagram to study self-organized criticality in ecosystems.

Meanwhile, the approach to particular systems was tackled when Bermúdez et al. (1989) built the direct precursor of INDISIM and succeeded in simulating the growth of *Serratia marcesens* and *Escherichia coli* in different situations, in accordance with experimental data. In the year 2002, the current term INDISIM was coined for first time and its general methodology was explained in detail (Ginovart et al., 2002a). From there, INDISIM evolved by studying specific cases of interest, such as bacterial growth in agar plates (Ginovart et el., 2002c), the study of the influence of bacteria size and shape in yoghurt processing with *S. thermophilus* and *L. bulgaricus* (Ginovart et al., 2002b), bacterial ecosystems in soil dynamics (Ginovart et al., 2005; Gras, 2004) and during the composting process (Gras et al., 2006; Prats et al., 2006b), flocculation in brewing yeasts (Ginovart et al., 2006; Ginovart et al., 2007) and the study of the lag phase in bacterial growth (Prats et al., 2006; Prats et al., 2007; Prats et al., 2008).
Arguments on the use of Individual-based methodology in the context of microbial systems, in general, and an overview of the range of applications that can be tackled with INDISIM, in particular, will be presented in Chapter 4.

2.1.2 General outline of the INDISIM methodology

INDISIM is an spatially explicit IbM. The modelled space is split into a set of regular divisions called spatial cells. The modelled events occur at finite and regular spans, called time steps. Two main entities are taken into account by the model: the microbial cell and

the spatial cell. The population is the ensemble of microbial cells and the environment is the set of spatial cells.

The state of a population composed by N microbial cells at a given time step t is described with the matrix $P_N(t)$.

$$P_N(t) = \{\overrightarrow{B}_i[v_1(t), v_2(t), ...v_s(t)]\}_{i=1,N} \tag{2.1}$$

The state of each microbial cell is defined with a vector $\overrightarrow{B_i}(t)$. The s components of this vector stand for: the individual label (integer number i), the spatial position of the cell and the microbial characteristic variables (e.g. age and mass, among others), which may change throughout the simulation according to the set of rules governing individuals (e.g. motion, uptake and metabolism, among others).

Microbial cells are located in a regular grid. The modelled space is divided into Q spatial cells, each representing a site in the grid. The state of the environment at a given time step t is described with the matrix $G_Q(t)$.

$$G_Q(t) = \{\overrightarrow{E}_j[w_1(t), w_2(t), ...w_r(t)]\}_{i=1,Q} \tag{2.2}$$

Each spatial cell represents the local environment of microbial cells, and is described with the vector $\overrightarrow{E_j}(t)$. The r components of this vector stand for: the cell label (integer number j), the coordinates in the spatial grid and the local characteristic variables (e.g. concentration of substrate, among others). The processes that affect each spatial cell are described by the set of rules governing the environment (e.g. substrate diffusion, among others).

The operating procedure of INDISIM follows the general scheme of Monte Carlo Methods. Firstly (a), the initial state of the system is set. Secondly (b), the rules governing the system are applied recurrently, and finally (c), once the simulation is finished, the output data is analysed.

The matrixs $P_N(t)$ and $G_Q(t)$ are usually very big ($N \sim Q \sim 10^4 - 10^6$) and not easily handled. For this reason, rules governing the model are applied through explicit first-order methods (the state of the system at next time step is computed just from the state of the system at the current time step). This allows the identification of individual cells and facilitates the implementation of rules describing local interactions. It also avoids performing operations that require a lot of computation (such as the inversion of $P_N(t)$ and $G_Q(t)$). Explicit solving of the proposed rules entails the reduction of the computational time. However, a burdensome limitation comes from the fact that explicit methods are numerically unstable for certain values of the parameters used by

the model. This hindrance was found during the execution of this thesis. The formulation of the problem (model of substrate diffusion through the hematocrit layer) and how it was handled is described in Section 3.5.

Rules governing the individuals and their local environment are implented as independent submodels that can be switched on and off at will. This means that models can be built with increasing complexity and that the outcome of the different versions of the same model are easily compared one with each other. Such procedure provides the staggered study of complex systems and allows determining the appropriate degree of complexness of the models.

INDISIM can track the evolution of single individuals. It can also store data at any time during the time step, thus following in detail the course of an event within the time step. This feature allows the comprehensive monitoring of any process ocurring in the model.

The general operation of the simulator is depicted in the flowchart in Figure 2.1. An exhaustive description of the general procedure of the simulator may be found in (Ginovart, 1996).

INDISIM-RBC is the specific application of the presented methodology to the study of the spread of the malaria parasite in *in vitro* red blood cell cultures. It was developed in close collaboration with the EMG-GSK, in response to their specific needs and interests. The next two chapters provide a detailed description of the procedure and a discussion of the obtained results with different versions of INDISIM-RBC: a preliminary 2D model with a time step of 1 hour, and the 2D and 3D models with a time step of 6 minutes.

2.2 ODD description of INDISIM-RBC in 2D

The *ODD* protocol is the standardized protocol used to present Individual-based Models in ecology in order to facilitate their communication and sharing (Grimm et al., 2006). It proposes a formal layout that consists of three blocks (Overview, Design concepts, and Details). In the first block, the generalities of the model are outlined. This includes the *Purpose* of the model, its operating *State variables* and *scales*, and a schematic *Process overview*, as well as a *Scheduling* details (the order in which the simulator applies the outlined rules). The second block should deal with the key concepts for designing and understanding IBMs in ecology, such as the *emergence* of behaviours at a system level, the use of *stochasticity* by the model, the nature and range of the *interactions* among individuals, and the extraction of information from the model to be compared with real world *observations*. Finally, the third block offers a detailed description of the particu-

larities of the model, such as the *Initialization* of the values for the variables, the use of *Input* data from external databases, and a thorough description of the *Submodels*: the rules and functions implemented.

A description of the 2D version of INDISIM-RBC is presented below in line with the ODD standard protocol.

Overview

Purpose:

The model aims to decipher essential underlying mechanisms that control *in vitro* cultivation of *Plasmodium falciparum* IRBCs. It specially focuses on how system-level behaviours may emerge from individual characteristics and population structure. Two versions of the program are presented here: *2Dv.1* and *2Dv.2*.

Characteristic scales:

The model is spatially explicit and stands for a small fraction of the hematocrit. In the first version (*2Dv.1*), the space is modelled as a regular 2D grid formed by $L \times L$ spatial cells (*sc*), with $L = 100$-300 spatial cells. Each spatial cell is a square of $l_{sc} = 10\ \mu m$ side (l). Therefore, the grid represents a surface of $1\ mm^2$. Processes are modelled discretely and events take place at finite time steps. The time step (*ts)* for the first version of the model (*2Dv.1)* is set to 1 hour. The value of the simulation unit that represent substrate particles (*su*) is set to 10^6 molecules.
The later version of the simulator (*2Dv.2*) operates at smaller temporal scales, in particular the time step is reduced one order of magnitude and set to $ts = 6$ minutes. Such modification comes together with the reduction of the simulation unit to $su = 10^5$ molecules, and entails modifying some parameters defined in the model. These modifications are specified *in situ*.
The characteristic scales used by this version are outlined in Table 2.1.

Boundary conditions:

The model represents a small part of a real culture that exchanges biota and substrate with the rest of the system. Or what is the same, most of the surface of the culture system consist on thousands of replicas of the model tyling it (exceptionally, the regions next to the walls of the culturing flasks adon't limit with other parts of the culture). Periodic boundary conditions (PBC) are used to effectively simulate this kind of infinitely tiled

systems. Basically, they make any RBC or substrate particle going beyond the boundaries of the model get in again through a point symmetric to its exit. Such boundary condition reduces the effect that the finite size of the simulation grid has on the dynamics of the model (Hoffmann and Chiang, 2004).

State variables:

Two types of low level entities are defined, the Red blood cell (RBC) and the spatial cell (SC). Each entity has aset of variables that determines its state. Average values of the characteristic X computed throughout the population of RBCs (or spatial Cells) are represented with the notation $\overline{X}$. Values temporally averaged throughout the simulation course are represented with the notation $\langle X \rangle$.

1. Variables of the RBC (the Red Blood Cell susceptible to infection). We distinguish three kinds of individual variables:

   a) Common characteristics for every RBC. They include constant values, such as the cellular volume, or the types of infection states: healthy RBC (type 0), and IRBC in the ring (R), trophozoite (T), schiszont (S) and fragmenter (F) stages of the infection cycle. And also properties that deppend on the cellular state, such as the infection susceptibility (P_{inf}), which deppends on the age of the RBC, or the accidental death rate during the infection cycle (P_{death}), which deppends on the stage of the infection cycle.

   b) Individual characteristics that are set when the RBC is introduced into the model. Each RBC has its own specific traits that are represented as characteristics that do not vary through the simulation. The values representing them (X) are set with normal distributions ($N\left(\overline{X},\sigma_X\right)$) around mean values ($\overline{X}$) and with deviations ($\sigma_X$) that are extracted from measurements of real systems. Some individual characteristics are: the uptake requisites for healthy RBCs (U), the maximum number of time steps that viable RBCs last (t_{MRBC}), the time for optimal infection susceptibility (t_{MSUS}) or number of time steps that the RBC is regarded as an optimal target of invasion by the malaria parasite, the duration of the infection cycle (t_{MINF}) and the duration of each of the stages of the infection cycle ($t_{MCL}\left(\{R, S, T\, and\, F\}\right)$).

   c) Individual variables that vary through the simulation. They define the instant state of the cell. Some variables are: location in the spatial

grid (x, y), RBC age (t_{RBC}) and time in culture (t_{cult}), infection stage ($ICL = \{0, R, S, T\, or\, F\}$), and post-invasion time (t_{INF}), and metabolic stress index (IM), among others.

2. Variables of the spatial cell. They represent characteristics of the local environment. They include: total number of RBCs ($n_{RBC}(x, y)$), and of each of the infection stages ($n_{ICL}(x, y)$, $ICL = 0 : F$), number of merozoites ($n_{mero}(x, y)$), and total amount of glucose ($C_{gluc}(x, y)$) and lactate ($C_{lact}(x, y)$).

A scheme presenting the characteristic scales, entities and spatial structure of the model is depicted in Figure 2.2. The list of variables, their values (in simulation units and in measured amounts) and their reference sources to set them are presented in Table 2.1.

Process overview and scheduling:

The general scheduling of the simulator follows the scheme depicted in Figure 2.1.

At each time step, the processes affecting the low level entities of the model are split into: processes affecting the RBCs, processes affecting the spatial cells and processes affecting the system as a whole.

First, the actions on every RBC take place. RBCs act sequentially, and each one carries out its actions consecutively. Individual actions affect solely the spatial cell in which they are and modifications in the spatial cell due to RBC actions are updated as each individual action goes along. Once a RBC has finished all his actions, another RBC starts to act. Output data regarding ony of these processes is labeled as *7i* in Figure 2.1.

The processes affecting the RBCs alter their local environment (characteristic variables of the spatial cells). This causes inhomogeneities within the system. Transport phenomena occur to relax the appearing gradients, these are the processes affecting the spatial cells. Output data reflecting the gradients created by the biological activity is labeled as *7ii* in Figure 2.1.

Once every RBC has acted, the actions on each spatial cell take place. Modifications on the characteristic variables of each spatial cell are stored as the actions over the spatial cells go along, and updated only after every spatial cell has acted.

Finally, process that affect the system as a whole affect every RBC and spatial cell. They occur once in a while and represent global modifications of the population and environment. Processes on the whole system are implemented after actions on the individuals and spatials cells occurred. They represent a modification of the scenario for the next time step. Output data showing the configuration of the system after all actions occurred is labeled as *7iii* in Figure 2.1.

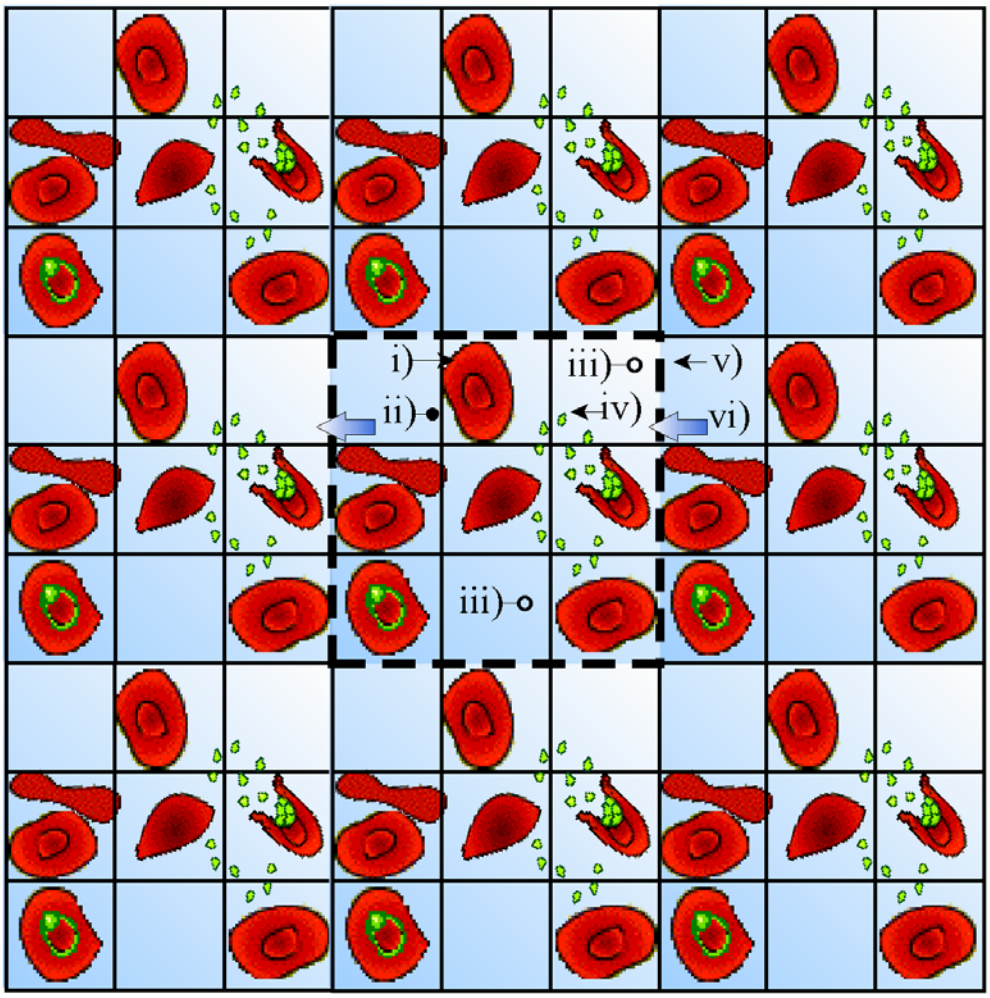

Figure 2.2: *Depiction of the entities, variables and spatial structure of the 2D versions of INDISIM-RBC. Figure shows a 3x3 square grid containing i) RBCs and ii) Spatial cells. Up to 2 RBCs can be contained in a spatial cell; iii) concentration of substrate glucose is indicated as the gradation in a blue-white scale; iv) extracellular merozoites are depicted as green dots. Figure shows nine replica of the simulated system to ilustrate the implementation of periodic boundary conditions. Whenever a magnitude is transported through the beyond of the system, it reenters it from a simetrical spot.*

Parameter description and notation		Value		Reference source
Time step	*ts*	1 h (*0.1 h*)		
Grid size	L	100		
Cell size	l_{sc}	10 μm		
Substrate unit	*su*	10^6molecules		
Maximum RBC age	$\bar{t}_{MRBC}$	2800 (*28000*) ts		(Capps and Jensen, 1983)
Maximum time for infection	$\bar{t}_{MSUS}$	170 (*1700*) ts		(Pasvol et al., 1980)
Average glucose uptake rate	$\overline{U}$	20 su / ts		(Guyton, 1987)
Maximum metabolic stress index	IM_{MAX}	3		
Maximum and minimum infection susceptibility for RBCs	P_{max}, P_{min}	0.7 - 1, 0 - 0.5		
Duration of infection cycle	$\bar{t}_{MINF}$	48 (*480*) ts		
Maximum duration ($\bar{t}_{MCL}(-)$) and death probability (P_{death}) of the infected stages:		$\bar{t}_{MCL}(-)$	P_{death}	(Sherman, 1998a)
- Healthy RBC	*0*	-	0	
- Ring	*R*	18 (*180*) ts	0	
- Trophozoite	*T*	17 (*170*) ts	10^{-3}*(*10^{-4}*)*	
- Schizont	*S*	7 (*70*) ts	10^{-2}*(*10^{-4}*)*	
- Fragmenter	*F*	6 (*60*) ts	10^{-2}(10^{-3})	
Initial parasitaemia	$\%I_0$	0.1 - 5 %		(MR4, 2008; Trager and Jensen, 1976)
Initial distribution of post-invasion times	$\bar{t}_{INF}(0)$	0 - 48 (*480*) ts		
Initial glucose per spatial cell	$\overline{C}_{gluc}(ts=0)$	1200 su		
Initial lactate per spatial cell	$\overline{C}_{gluc}(ts=0)$	0 su		
Subcultivation period	-	48 (*480*) - 96 (*960*) ts		(MR4, 2008)
Subcultivation ratio	-	66% -80% of RBCs		

Table 2.1: *Caracteristic parameters of INDISIM-RBC that describe: the scale and spatial structure of the model, the RBC population, the parasite strain, the suitability of the culturing medium (captured as an accidental death probability), the initial conditions and the experimental protocols. Enclosed italic values correspond to the version 2Dv.2 and v3D of the simulator that operate at reduced time steps.*

The modelled processes are listed outlined below following this scheduling:

1. Processes affecting the RBC (including IRBCs).
   (a) Motion: RBCs randomly shift positions in agitated cultures and dont move in static culures (see process 3c).
   (b) Uptake: each RBC uptakes an amount of glucose that depends on its requisites and on the concentration of substrate at the same spatial cell.
   (c) Metabolism: the uptake requisites for the next time step are determined and an amount of lactate is excreted as a function of the uptake.
   (d) Infection:
      i. IRBCs undergo the parasite life cycle. After a complete infection cycle, the IRBC dies and may infect neighbor cells, so
      ii. any healthy RBC neighboring an IRBC at its lysis may become infected with a fixed probability.
   (e) Death:
      i. accidental rupture, with probability that increases with the infection stage;
      ii. exceeding the maximum age;
      iii. exceeding maximum metabolic stress index.
      iv. lysis at the end of thew infection cycle.

      Dead RBCs are removed from the model.
2. Processes affecting the spatial cells (SC).
   (a) Diffusion of substrate: Substrate propagates from cell to cell following the Fick's law to compensate concentration gradients. The total surface between cells is not taken into account.
   (b) Motion/death of RBCs: RBCs may abandon a spatial cell (see processes 1a and 1e).
   (c) Reseting of spatial cell characteristics due to a subculture (process 3a) or medium renewal (process 3b).
3. Process affecting the system as a whole (WS).
   (a) Subcultivation: A variable fraction of the RBC population is replaced by healthy RBCs and the medium is renewed. Glucose and lactate concentrations are set to initial values.

(b) Medium renewal: A fraction of the culture medium is renewed.

(c) Agitation: RBCs shift their position in the spatial grid at random.

Design Concepts

Emergence:

Population structure (e.g. distribution of post-invasion times among the IRBC population, among others) and spatia patterns emerge from the individual defined infection cycles, from the local interactions among RBCs and merozoites, as well as from the inclusion of stochasticity in the processes. The average infection proliferation within the culture (growth ratio) deppends on these features, thus also emerges from the individually defined rules.

Adaptation/Fitness/Prediction/Sensing:

RBCs percieve just their inmediate surroundings ($\sim 10\ \mu m$). Rules governing individuals take into account the number of glucose and lactate particles, extracellular merozoites and RBCs at the same spatial cell. Rules governing each spatial cell $(i,\ j)$ take into account its nearest neigbourhood (nn) which consists on eight cells, four side-to-side spatial cells -$(i,\ j \pm 1)$ and $(i \pm 1,\ j)$- and the four diagonal neighbours $(i \pm 1,\ j \pm 1)$.
The principle of parsimony suggests that, since cellular behaviour can be explained as a mechanism, it is better not to resort to more complex explanations, such as individual preferences or free will. For this reason, RBCs don't make decisions to increase their fitness and no adaptation rules are defined. RBCs have neither predictive capability and only sense their environment in the sense that they uptake as much nutrient as available.

Stochasticity:

Randomness is introduced in the model in different ways. Firstly, population is heterogeneous. Some individual characteristics (e.g. U, t_{MRBC}, among others) are randomly distributed among the population with normal distributions around mean values. Secondly, uncertainty and variability are introduced in individual processes as a gaussian noise on the expected values (e.g. instant uptake needs u are set following $N(U, \sigma_U)$), and also as probabilistic rules (e.g. infection of healthy RBCs -submodel *1d.ii*- or accidental death -submodel *1e*-).
It must be stressed that the inclusion of stochasticity is essential in order to obtain the appropriate emergent behaviours. Finally, random numbers are also used to create the

tables that determine the order of action when RBCs act sequentially.

Interactions:

The main interaction between individuals is the spreading of the infection. RBCs also interact with each other by competing for nutrient, and accumulating harmful substances. Both interactions are local and range to the first nearest neighbor cells.

Observation:

The graphical interface of the model shows data colected at the end of each time step (label 7iii in Figure 2.1). It is comprised by three (six, in version *2Dv.2*) graphical windows and a numeric display showing system-level variables that are updated on the fly. The windows represent (1) the spatial distributions of RBCs and IRBCs, and (2) merozoites (in version *2Dv.2*, concentrations of (3) glucose and (4) lactate are also depicted), (5) the temporal evolution of the number of RBCs ($N_{RBCs}(ts)$) and parasitaemia ($\%I\ (ts)$, percentage of IRBCs), and (6) the IRBC age structure (histogram of the distribution of t_{INF} among IRBCs: $q(t_{INF})$). The numeric display (7) shows the number of time steps and equivalent time in real cultures, and the number of RBCs, IRBCs and extracellular merozoites. All this information and additional data (for instance, track of the life cycle of a single cell, measurments of the infection growth ratio (GR) and external modifications on the whole system that represent manipulation of the cultures, among others) may be recorded and stored in files for further analysis. System-level variables (parasitaemia, growth ratio, age structure, abundance of multiply infected IRBCs, and average concentration of glucose and lactate) are compared to real world observations. Measurments of local characteristics of the real systems have not been performed during the ellaboration of this thesis, so spatial distributions of local variables are not compared to experimental results (see Table 1.2 in Section 1.3). Figure 2.3 shows a screenshot of the *2Dv.2* version of the model.

Details

Initialization:

Cultures are set with random distributions of RBCs, infected RBCs and glucose. No lactate and no extracellolar merozoites are initially introduced into the model. The individual characteristics (X) for each RBC are set following normal distributions around the average values ($\overline{X}$) with variance $\sigma_X^2 = 10\%$ and trunkated ($\overline{X} - 1.96\sigma_X^2 < X < \overline{X} + 1.96\sigma_X^2$) to avoid unrealistic values (Prats, 2008). The distribution of post-invasion times among IRBCs is set in correspondence to the specific experimental input measurement.

Parameters that describe the general characteristics of the RBC and *Plasmodium* (see Section 1.2) and those describing the general features of the culture system (see Section 1.3) are taken from literature and presented in Table 2.1. Dimensions of the spatial grid, initial parasitaemia, total number of RBCs, and age structure of the IRBC inoculum are specifically set to reproduce experimental behaviours. A detailed list of the particular input parameters are presented for each simulation in Section 2.3, together with the inputs of the experimental observations, and particular modifications of the values shown in Table 2.1.

Inputs:

The experimental culturing protocols are external manipulations of the system represented by imposed dynamics of the variables of the system. They comprise subcultivation, medium renewal and agitation.

Subcultivation of the parasite is modelled in accordance to the protocols proposed by Trager and Jensen (1976) and the MR4 recommendations (MR4, 2008). It consist on the removal of most of the population and substitution by healthy RBCs. Depending on the protocol, it occurs each 48 ts- 96 ts, and the fraction of removed population varies from 66% to 80%.

Medium renewal is modelled as the (partial or total) reseting of the glucose and lactate concentrations. It takes place each 24 hours and together with the subcultivation, but not the day after subcultivation.

Agitation is modeled as the shift of positions of every RBC and the homogenisation of glucose and lactate concentrations. It occurs together with subcultivations or at fixed spans. In particular, suspended cultures are modeled as being agitated every time step.

Other external inputs that can control the dynamics of the culture are the storage time of RBCs, the strain of *P. falciparum*, the composition of the culturing medium or serum, the treatment with drugs or the exposition to other harmful situations (temperature, chemicals). They are modeled as specific modifications of the parameters in the model and are explained in detail when required, in Section 2.3.

Submodels:

Two kinds of submodels are distinguished: 1) those refered to the erythrocytes (RBC), and 2) those refered to the spatial cells (SC).

Shuffling. RBCs act sequentially. In order to reduce the artificial bias on the model outcome caused by the process of sequentation, we impose random orders of action

for the RBCs at each time step, to avoid having an RBC always favored with he first turn of action. Randomization of the sequences of action i using disordering tables that can reshuffle any given list. These tables are created at the begining of the simulation (step 3 in Figure 2.1).

1a) RBC Motion. At the begining of the simulation and after an agitation, RBCs are randomly distributed through the spatial grid. X and Y coordinates for each RBC are integers set randomly between 1 and L. The spatial cells have limited size and allow only 2 RBCs occupancy. If there are already two RBCs in the selected spatial grid, the coordinates are drawn again.

1b) RBC Uptake. At each time step, each RBC attempts to uptake as much nutrient as it is required by the ingestion needs. They are represented by the number of particles of glucose to be ingested ($U_{in} = u \cdot IM$). The value for the instant uptake needs (u) is drawn from a normal distribution centered on the average uptake rate for $\overline{U} = 20\ su/ts$. The real uptake (U_{eff}) is limited by the availability of glucose at the same spatial cell (x,y): $U_{eff} = max(U_{in}, C_{gluc}(x, y))$. The excess of lactate hinders the uptake of glucose. If the amount of lactate in the same spatial cell (x,y) exceeds a maximum threshold ($C_{lact}(x, y) > 10000\ su$), U_{eff} decreases to zero. If the U_{in} requisites are not fulfilled, the RBC will suffer metabolic stress (increase of the metabolic stress index IM), enhancing further ingestion needs.

Infected RBCs have more energetic demmands than healthy RBCs and must consume more glucose (see Section 1.2.2). Their ingestion needs (U_{in})are fixed in the same way as for healthy RBCs, but they are multiplied by a factor that varies with the post-invasion time $f(t_{INF})$ (Kirk, 2001; Lew et al., 2003).

$$f(\tau) = \begin{cases} 1\ su & ; \quad \tau < 18\ ts \\ max\left(20\Delta(\tau - 18), 100\right)\ su & ; \quad 18\ ts < \tau < 35\ ts \\ 100\ su & ; \quad 35\ ts < \tau < 40\ ts \\ max\left(100 - 20\Delta(\tau - 40), 1\right)\ su & ; \quad 40\ ts < \tau < t_{MINF} \end{cases} \tag{2.3}$$

The shape of $f(\tau)$is depicted in Figure 2.4. For each ingested glucose unit two lactate units are dumped into the extracellular medium.

1c) RBC Metabolism. For each ingested glucose unit two lactate units are dumped into the extracellular medium. The metabolic stress index (IM) increases if the RBC starves and is reduced to its minimum whenever the RBC fullfils its ingestion

needs. If the effective uptake is smaller than the RBC requisites ($U_{eff} < U_{in}$), then $IM = IM+1$. Otherwise, $IM = 1$. Excessive metabolic stress may cause the death of the RBC ($IM > IM_{MAX}$).

$t_{MINF} <$ *44 ts*	$t_{MCL}(F) = 2\ ts$	$n_r = 2^2 = 4$
44 ts $< t_{MINF} <$ *46 ts*	$t_{MCL}(F) = 4\ ts$	$n_r = 2^3 = 8$
46 ts $< t_{MINF} <$ *50 ts*	$t_{MCL}(F) = 6\ ts$	$n_r = 2^4 = 16$
50 ts $< t_{MINF}$	$t_{MCL}(F) = 8\ ts$	$n_r = 2^5 = 32$

Table 2.2: *Duration of the fragmenter stage ($t_{MCL}(F)$) and number of merozoites released to the extracellular medium (n_r) as a function of the duration of the infection cycle (t_{MINF}).*

1d) RBC Infection. It comprises: i) the evolution of the infection cycle for the IRBCs, and ii) the process of invasion of healthy RBCs by extracellular merozoites.

i) At the beginning of the simulation, the duration of the infection cycle t_{MINF} is set for each RBC. If the RBC becomes infected, the span t_{MINF} constrains the duration of the fragmenter stage and determines the number of merozoites released to the medium. This model assumes that the duration of every mitotic partition is 2 hours and that the average number of nuclei divisons is 4 (see Section 1.2.2). The corresponding values are outlined in Table 2.2.

At each time step, the post-invasion time is actualised ($t_{INF} = t_{INF} + 1$) for every IRBC. At fixed spans the IRBC switch their infection stage, from ring to trophozoite, schizont and fragmenter (see Table 2.1). When $t_{INF} > t_{MINF}$, the IRBC lyses and n_r merozoites are released into the medium. Merozoites are randomly distributed among the first neighbor spatial cells, including the own spatial cell (see Figure 2.9 in Section 2.3.3).

ii) The process of invasion is specific for each version:

2Dv.1 For each healthy RBC, the model counts how many merozoites are in its same spatial cell. Each merozoite can invade the RBC. Invasion is a probabilistic event that occurs with a probability P_{inf}. P_{inf} linearly decreases with the age of the RBC (t_{RBC}) and abruptly decreases with time in culture (t_C, see Submodel 3a) after the threshold of optimal susceptibility (t_{MSUS}).

$$P_{inf}(t_{RBC}, t_C) = \left[P_{max} - \frac{t_{RBC}}{t_{MRBC}}(P_{max} - P_{min})\right] \Delta min(1,\ e^{-\frac{t_C}{t_{MSUS}}}) \quad (2.4)$$

For a given culture trial, the storage time of RBCs (τ_{store}) is a fixed value, $IE = ITCULT + \tau_{store}$ and P_{inf} depends on a single variable. The shape of this function is depicted in Figure 2.4, b. The values P_{max} and P_{min} depend on the parasite strain and on the procedence of RBCs. Multiple invasion of IRBCs is allowed until reaching the maximum value of 3 parasites per IRBC. Beyond this value, the IRBC is no longer invadible. The merozoites that don't invade a healthy RBC are removed from the model.

2Dv.2 In the second version of the 2D model, invasion occurs like in *2Dv.1*. The sole exception is that now merozoites are let in the culture for up to 5 time steps before being removed. While they stay in the extracellular medium, they can eventually move to a nearest neighbour cell after each unsuccessful invasion attempt.

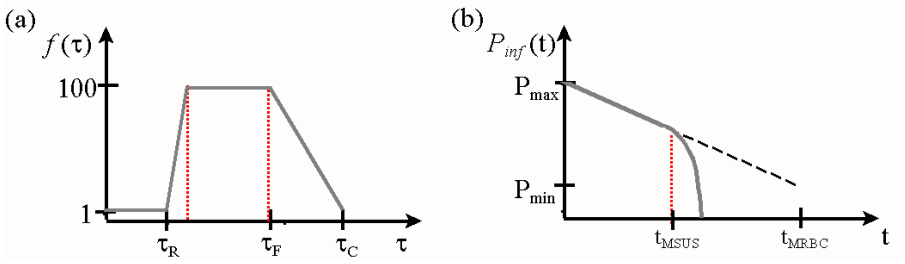

Figure 2.4: *a) Factor of individual metabolic requirements for infected RBCs ($f(\tau)$ as a function of the post-invasion time (t_{INF}). $\tau_R = t_{MCL}(R)$: end of the ring stage of the infection cycle. $\tau_F = t_{MCL}(R) + t_{MCL}(T)$: beginning of the fragmenter stage of the infection cycle. $\tau_C = t_{MCL}(R) + t_{MCL}(T) + t_{MCL}(S) + t_{MCL}(F)$: end of the infection cycle. b) Evolution of the individual infection susceptibility. t: time in culture. P_{max}: maximum invasion probability, for young RBCs. P_{max}: minimum invasion probability, for old RBCs. t_{MSUS}: maximum time of infection susceptibility. t_{MRBC}: maximum individual age.*

1e) RBC Death. Each RBC may die due to i) accidental death probability P_{death}, ii) because it exceeded its maximum age ($t_{RBC} > t_{MRBC}$), or iii) could not fulfill its metabolic demands ($IM > IM_{MAX}$). IRBCs, in particular, die at the end of their infection cycle (iv, when $t_{INF} > t_{MINF}$). Whenever a RBC dies, it is completely removed from the model. The accidental death probabilities per time step depend on the culturing medium and are fixed to best fit experimental results.

2a) SC Diffusion. Thue number of glucose and lactate units at each spatial cell is updated at each time step due to biological activity. This entails a local heterogeneity furtherly compensated through diffusion, according to the Fick's law. For each spatial cell and at each time step, the model evaluates the difference in concentration with the neighbor cells and allows a partial balance of the concentration of substrate. Diffusion is carried out with the FTCS (Forward-Time Central-Space) explicit Euler method (Hoffmann and Chiang, 2004). Let C_{ij}^t be the concentration of substrate (C_{ij}^t refers to $C_{gluc}(i,j)$ or $C_{lact}(i,j)$, indistinctly, at time step t). The amount of substrate at the same spatial cell the next time step is given by:

$$C_{i,j}^{t+1} = C_{i,j}^t + \widetilde{D} \sum_{k,l}^{nn(i,j)} w_{k,l} C_{k,l}^t \tag{2.5}$$

Where k and l are the spatial coordinates of the 9 cells that constitute the inmediate environment of cell *(i,j)*. Equation 2.5 adds up the contributions of the each cell to the diffusive transport of substrate. The set of parameters $w_{k,l}$ weight the contribution each spatial cell, inversely proprtional to the distances bewteen centers of the spatial cells, and normalized so that $|w_{k,l}| = 1$. This results in:

$$w_{k,l} = \begin{cases} \frac{1}{4(1+\sqrt{2})} & ; \ (k \neq i) \cap (l \neq j) \\ -1 & ; \ (k,l = i,j) \\ \frac{\sqrt{2}}{4(1+\sqrt{2})} & ; \ otherwise \end{cases} \tag{2.6}$$

The weights in equation 2.6 are presented and justified in (Ginovart, 1996). $\widetilde{D}$ is the effective diffusion coefficient. Its value is set to $\widetilde{D} = 0.6$, the maximum numerically stable value.

This explicit model of the diffusion process is just a preliminary approach to assessing the local limitations of substances. Subsequent work has shown that version of the diffusion model here presented should be improved. This problem is furtherly revisited in Section 3.5.2.

2b) SC Merozoite propagation. In version *2Dv.1*, merozoites are released into a spatial cell and removed during the next time step, so no propagation of the parasite is modelled beyond merozoite eggress at the end of an infection cycle. In contrast, in version *2Dv.2* mrozoites can move to nearest neighbour spatial cells during the next five time steps after eggress. At the end of each time step, merozoites remaining

in the medium randomly shift their position to a nearest neighbour cell with a probability set to P_{prop}. The value of P_{prop} is adjusted to best fit experimental observations and typically takes values ranging from 0 to 0.1.

3a) WS Subcultivation. Subcultivation may occur at prefixed periodes or be triggered whenever a prefixed threshold parasitaemia is surpassed. The fraction of population to be replaced may also remain fixed or depend on the parasitaemia prior to subcultivation. The mechanism to set the rate and extent of subcultivations reproduces the experimental setup to be simulated.

Subcultivation entails the extraction of a fraction of the RBC population, the inclusion of healthy RBCs to reach the initial number of RBCs ($N_{input} = N_{RBC}(t = 0) - N_{RBC}(t = t_{SC})$), the complete renewal of the culture medium (see Submodel 3b), and the agitation of the new population (Submodel 3c).

Whenever a new RBC is introduced into the culture, the values of all its characteristics are set as in the initialization and its time in culture is set to zero ($t_C = 0$). Therefore, the time in culture of a RBC is the difference between the age of the RBC at the present time step and the age it had when it was introduced into the culture.

3b) WS Medium renewal. Medium renewal can occur at discrete events or continuously, depending on the experimental system to be simulated. Discrete medium renewal occurs daily, except the day after a subcultivation and it entails the complete renovation of medium. Contiuous medium renewal entails the extraction of a fraction of C_{gluc}, C_{lact} and n_{mero} from each of the spatial cells and the replacement of such amounts for the corresponding values under the initial conditions. After each subcultivation the medium is always completely renewed.

3c) WS Agitation. Agitation of the culture system is simulated as the random shift of position of RBCs, together with the uniformization of the local values of C_{gluc}, C_{lact} and n_{mero}.