Individual-based modeling of *Plasmodium falciparum* erythrocyte infection in *in vitro* cultures

Jordi Ferrer Savall

Individual-based modeling of *Plasmodium falciparum* erythrocyte infection in *in vitro* cultures

by

Jordi Ferrer Savall

PhD THESIS

Submitted to fulfill the requisites for the degree of doctor by the

Universitat Politècnica de Catalunya

Supervisors: **Daniel López Codina** and **Joaquim Valls Ribas**

Departament de Física i Enginyeria Nuclear

Programa de Doctorat de Física Aplicada i Simulació en Ciències

Universitat Politècnica de Catalunya (UPC)

Castelldefels. May, 2010

Contents

1 Introduction **1**

1.1 Background 2

1.1.1 Research team 2

1.1.2 Malaria in the world 4

1.2 Malaria and *Plasmodium falciparum* 8

1.2.1 Life cycle and pathogenesis 8

1.2.2 Specific features of the intraerythrocytic schizogony 12

1.3 Cultivation methods 20

1.3.1 *In vitro* cultivation of erythrocitic stages 20

1.3.2 Cultures performed by EMG-GSK 21

1.3.3 Pending issues 23

1.4 Models of the infection process 25

1.4.1 Levels of description and other classification criteria for mathematical models 25

1.4.2 Strategies in modeling complex systems 26

1.4.3 Models of the spreading of diseases 28

1.5 Aim, objectives and outline of the thesis 31

2 INDISIM-RBC: model of *P. falciparum* infected RBCs in *in vitro* cultures **35**

2.1 Background and general outline of INDISIM 36

2.1.1 Basis of INDISIM 36

2.1.2 General outline of the INDISIM methodology 37

2.2 ODD description of INDISIM-RBC in 2D 41

2.3 Analysis of experimental observations with the 2D model 56

2.3.1 Short-term preservation of IRBCs. Adjustment of INDISIM-RBC. 56

2.3.2 Long-term cultivation of IRBCs. Readjustment of INDISIM-IRBC. 59
2.3.3 Patterns observed in *in vitro* cultures. 64
2.3.4 Synchronization of the infection process 74
2.3.5 Substrate availability 77
2.3.6 Agitation of the culture systems 80
2.4 Discussion and open questions 84

3 *P. falciparum* infected RBCs in static cultures 87
3.1 Geometrical characterization of the static culture systems 88
3.1.1 Preliminary considerations on the application of INDISIM-RBC.*v3D* 88
3.1.2 Experimental trials on static culture systems 89
3.2 ODD Description of INDISIM-RBC in 3D 97
3.2.1 Need for a 3D model 97
3.2.2 ODD description of the 3D model 97
3.3 Analysis of experimental observations with the 3D model 103
3.3.1 Calibration of version *v3D* 103
3.3.2 Settling processes in the hematocrit layer 103
3.3.3 Multiple invasion of RBCs 105
3.4 Local limitations on the infection process 106
3.4.1 Effect of the walls of the culturing device 107
3.4.2 Effect of the hematocrit layer depth 109
3.4.3 Whole-system model that accounts for HLD and L together (WS model) 111
3.4.4 Outcome of INDISIM-RBC 112
3.4.5 Discussion 113
3.5 Local substrate limitations in the hematocrit layer 116
3.5.1 Diffusion limitations in the hematocrit layer 117
3.5.2 Explicit model of the diffusion process 125
3.5.3 Implicit model of the diffusion process 127
3.6 Discussion and open questions 129

4 Individual-based Models of microbial communities 131
4.1 Pattern-oriented modeling 132
4.1.1 Pattern-oriented strategy to build models in microbiology 132
4.1.2 Simplicity and specification in PoM 133
4.1.3 Application of PoM to INDISIM-RBC 134
4.2 IbM and PbM approaches in microbiology 136

4.2.1 Population-based models 137
4.2.2 Individual-based Models 140
4.2.3 Applicability of each approach 144
4.3 IbM, microbiology and basics of thermodynamics 146
4.3.1 Basic concepts of thermodynamics in biology 147
4.3.2 Characterization of bacterial and cell cultures 153
4.3.3 Thermodynamic description of microbial populations 155
4.3.4 Population structure in *Plasmodium falciparum* in *in vitro* cultures 158
4.4 Closing remarks 163

5 Conclusions and perspectives 165
5.1 General conclusions 165
5.2 Synthesis of specific results 168
5.3 Perspectives and further work 180

A Glossary 183

B Experiments carried out by EMG-GSK 191

C Statistical tests and tools 203

D Own publications 207

Chapter 1

Introduction

Systems featuring a large number of interacting components (agents, processes or mechanisms) may show a complex behavior, not easily derivable from the sum of the activity of individual components. Addressing particular problems often requires an understanding of the system beyond the identification of its components and the phenomenological description of the observed behavior. It requires building models that capture the essential features of the problem. Models are simplified representations of the reality used to address specific questions regarding the modeled systems. Computational models define objects and algorithms that simulate the components and mechanisms of the real system, with a view to assessing their combined effect on the system as a whole. Trial and refinement of the models is continually carried out together with real-world observations and experiments, in order to obtain insight and increase predictive capability.

Biology is crowded with complex systems, from genetic networks to ecosystems. Since the last century, biology has increasingly become an exciting domain to many physicists. The contributions of physical models and techniques to biology are found in many areas and at various levels of description: from the detailed analysis of structure and functions of bio-molecules in the nanoscale, at a molecular level, through the microscopic description of membranes and cells as a whole system, and up to the macroscopic level with the study of population dynamics of living organisms. In particular, if the systems under study are microbial communities, these scales are coupled together, providing a wide range of problems to address and situations to explore. The use of computational models of biology is often unavoidable due to the large amount of data to be handled and to the intrinsic complexity of the systems under study.

The present work is a theoretical approach to a particular microbial community: the

ensemble composed of red blood cells and the parasite *Plasmodium falciparum* which infects them, causing malaria. It includes experimental observations and some computer programming, but mainly focuses on the process of building a model of appropriate complexity. It may be of interest to researchers from disciplines that range from medicine to biology, engineering, physics and applied mathematics. Different educations and trainings entail different worldviews and the communication among colleagues is not trivially straightforward when complexity arises. For this reason, I have tried to adapt the language and orientation of this manuscript to a multidisciplinary approach accessible to anyone with scientific education in the field of biomedical and/or experimental sciences.

The work on malaria is compared to the research carried out regarding other microbial communities. Thereby studying general emerging properties of microbial systems in general, with regard to the effect of cell individuality, heterogeneity and diversity; interactions among cells and between cells and their local environment; and biological and spatial complexity.

This first chapter is a preliminary approximation to malaria infected red blood cells. It is basically a compilation of facts and relevant features that are to be considered by the model. It also provides an outline of the fundamentals of modeling and of the Individual-based approach. A detailed description of the general and specific objectives of the thesis is also included at the end of the chapter.

1.1 Background

1.1.1 Research team

MOSIMBIO -*MOdelització i SIMulació discreta de sistemes BIOlògics*- is an interdisciplinary group (physicists, mathematicians, biologists and agricultural engineers) that has been working for nearly thirty years in modeling and simulation (MOSIMBIO, 2009). It has its origins on the fruitful collaboration of investigators that studied the behavior of liquids with techniques typically used by physics in solving N-bodies problems (Monte Carlo and Molecular Dynamics) (Giró et al., 1980) with researches that worked on the analysis of population dynamics in the framework of theoretical ecology (Lurié et al., 1983). The foundational aim of the group was noticing general properties of microbial populations and ecosystems with models based on the individual cells. This approach allowed combining the techniques and experience from both disciplines: the biological rules governing the individuals could be framed with the physical laws that also describe their local environment, and the collective outcome could be drawn from the statistical

treatment of the population ensemble and the (Giró et al., 1985; Giró et al., 1986).

The methodology provided a way to address questions that could not be covered otherwise, such as describing the effect of the statistical distribution of individual characteristics among the population on the system dynamics (Bermúdez et al., 1989), or proposing mechanisms for the emergence of coordinated global behaviors out from local individual interactions and randomness (Solé and Valls, 1992).

This seminal approach was further developed by López D. (1992), and Ginovart M. (1997), gradually increasing the degree of detail and specification of the models. The modeling methodology was revised, standardized and finally presented (Ginovart et al., 2002a) with the acronym INDISIM that stands for INDividual DIScrete SIMulations.

The research is currently carried out through the development of specific models addressed to particular systems. The microorganisms under study range from bacteria to fungi, yeasts and protozoa. The specific problems to be handled usually respond to a determinate industrial demand or social concern, and the developed models pursue the improvement of the technological use of microbial systems in fields such as food safety, soil quality or pharmaceutics. Yet, the key point and ultimate goal of the group is to acquire holistic understanding of microbial ecosystems, this is to unravel the connections between the local microscopic rules and the macroscopic observations.

The present work comes out from the collaboration of MOSIMBIO with the Experimental Microbiology Group (EMG-GSK) from the laboratory unit Diseases from the Developing World (DDW), a research and development center of GlaxoSmithKline. This unit holds around 2000 employees and engages more than 100 dedicated full-time scientific staff, whose skills include chemistry, biology, biochemistry, toxicology, cytotoxicology, assay development, and *in vivo* and *in vitro* screening, to assess disease targets and find drug candidates for further development. The DDW is currently executing 8 research projects, four on malaria supported by the partnership Medicines for Malaria Venture (MMV) and four on tuberculosis, in the framework of the Global alliance for Tuberculosis. Among other topics in the field of malaria, the research in the DDW pursues the improvement of the current *in vitro* clinical trialing protocols, the development of *in vivo* culturing methods through murine models, and the assessment of the population distribution characteristics through flow citometry (Jimenez-Diaz et al., 2009).

The objective of the four-year collaboration project (2005-2008) was to carry out basic research on the factors affecting the propagation of the malaria parasite *Plasmodium falciparum* in red blood cell (erythrocytes) cultures under controlled laboratory conditions. INDISIM-RBC, a specific version of the root model has been developed in order to undertake this study. The strategic reason for this investigation is to gain understanding

and control of the current culturing protocols of the parasite. This would eventually lead to the improvement and increase in efficiency of the pharmaceutical *in vitro* trials, and consequently facilitate drug development.

1.1.2 Malaria in the world

Malaria *(a.k.a.* paludism when the host is human) is the acute or chronic disease caused by the presence of sporozoan parasites of the genus *Plasmodium* in the red blood cells (a.k.a. erythrocyte and here referred as RBC). It is transmitted from an infected to an uninfected individual by the bite of anopheline mosquitoes, and characterized by periodic attacks of chills and fever that coincide with mass destruction of blood cells and the release of toxic substances by the parasite at the end of each reproductive cycle.

The intracellular parasite is an eukaryote microorganism of the phylum *Apicomplexa*, a large class of strictly parasitic protozoans that pass through a complicated life cycle that involves the alternation of a sexual with an asexual generation, and requires two or more dissimilar hosts. Half of the life cycle is spent in the lymphatic and circulatory system of a vertebrate host, and the other half takes place in the digestive system of a mosquito host, often referred as the disease vector. There are over the 200 known species of *Plasmodium*, and at least four species that infect humans: *P. falciparum, P. vivax, P. ovale* and *P. malariae*. Other species infect animals, including monkeys, rodents, birds, and reptiles. The female mosquitoes of the genera *Culex, Anopheles, Culiceta, Mansonia* and *Aedes* may act as vectors of the disease because they complement their basic nectar nourishment with hematophagy (drinking blood from animals). The currently known vectors (more than 100 species) for human malaria all belong to the genus *Anopheles*.

The evolution of the parasite inferred from the DNA sequentation suggests that *Aplicomplexa* were originally autotrophus prey protozoa that evolved the ability to invade the intestinal cells of jellyfish and subsequently lost their photosynthetic ability. The genus *Plasmodium* is closely related to the genus *Leukocytozoon*, which infects leukocytes (white blood cells), liver and spleen cells and is transmitted by 'black flies', and which formerly evolved the ability to infect the liver from the intestine cells. It is currently believed that *Plasmodium* evolved from *Leukocytozoon* about 130 million years ago, in coincidence with the spread of angiosperms (flowering plants). This expansion presumably led to an increase in the number of mosquitoes and their contact with vertebrates. It seems probable that birds and reptiles were the first group infected by *Plasmodium,* and primates and rodents became infected later(Escalante and Ayala, 1995). *Plasmodium* parasitism of humans has independently arisen several times(Leclerc et al., 2004). Nevertheless, the extant spices *Plasmodium falciparum* is believed to derive from one single parasite an-

cestor within the last 50,000 years. This long history of coexistence between the parasite and the host entails an evolutionary arms race that has lead to sophisticated infective and immunity strategies, respectively (Rich and Ayala, 2000).

The term malaria was coined from the Italian words: "mala aria" meaning "ill airs". The first historical references to this disease are texts dated 1500 B.C, from China and the Ancient Egypt. From the very beginning, the disease was associated to swamps and wetlands, the natural habitat of the vector mosquito. In 1880 Charles Laveran observed for the first time the protozoan parasite associated to paludism. In 1891, Romanowsky successfully stained the parasite with Geimsa coloration, thus allowing the study of its morphology. Meanwhile, the disease was eradicated from northern and western countries during the XIXth and XXth centuries through the reduction of swamplands, and specially as a consequence of the generalization of health care and attention, which implied the control and isolation from possible vectors of the infected human hosts. In the late 1960s, it was expected that malaria would be completely wiped out from the world and that it would not entail a serious health concern. This optimistic prospects were based on the successful eradication from areas were the disease had traditionally been endemic and on the existence of effective treatments. However, in the last 20 years there has been a notable rise, rather than a decrease, in the incidence and harm of this pandemics, particularly, in the southern hemisphere. The pointed causes for this drift are the lack of resources from governments and social agents assigned to face the menace, the increase of mobility and connection of the societies, the increase of precipitations in the affected areas and the abolition of DDT fumigation, among others. The appearance of several new strains of the parasite with cross-resistance to the current treatments and the lack of success of the employed strategies draw a much less optimistic scene for the immediate future (Greenwood and Mutabingwa, 2002).

Today, malaria remains the greatest single cause of debilitation and death throughout the world. According to the World Malaria Report 2008 edited by the World Health Organization (WHO, 2008), half of the world's population is at risk of malaria. There were an estimated 247 million malaria cases among 3.3 billion people at risk in 2006, causing nearly a million deaths, mostly of children under 5 years. This means one child decease every 30 seconds. 109 countries were endemic for malaria in 2008, 45 within the African region (see Figure 1.1). Children are particularly vulnerable to this disease, as well as pregnant women. The incidence in this collective is four times the average for adult men, and survival to the disease in this conditions is drastically reduced to 50%. Malaria, in combination with other major health concerns, such as malnutrition and AIDS, literally decimates the population in the sub-Saharan African countries. Assessing

the real burden of malaria accurately is difficult because most deaths occur at home, the clinical features of paludism are very similar to those of many other infectious diseases, and good quality microscopy is available in only a few centers.

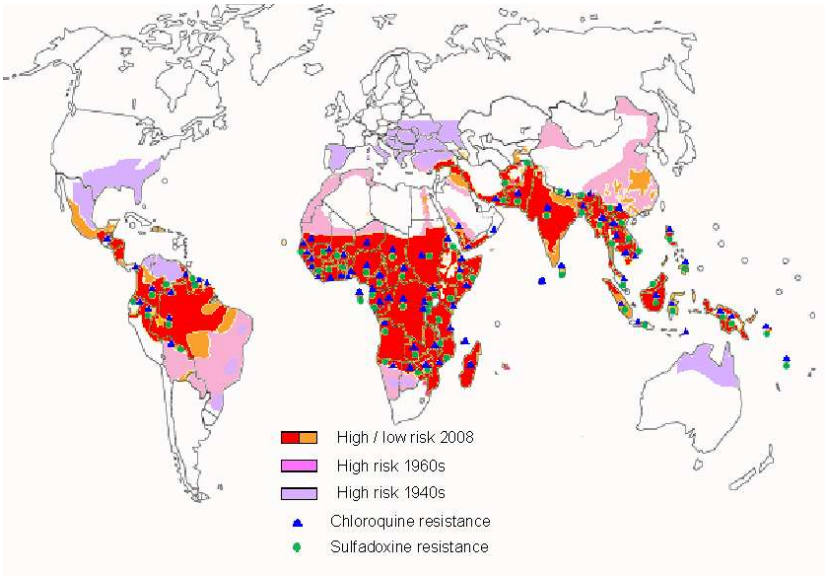

Figure 1.1: *Risk of malaria infection worldwide and historical evolution. Risk estimation: high malaria risk indicates endemic (constant incidence on the population) of the area. Low risk indicates that child infection prevalence is below 10%. Data from the WHO World Malaria reports 2005 and 2008 (WHO, 2008).*

Besides being a humanitarian and sanitary scourge, malaria is also a great socioeconomic burden for the affected countries (Sachs and Malaney, 2002). Malaria and poverty are intimately connected and feed one each other, although the correlation between prevalence of the epidemic and of poverty is not straightforward due to the effects of ecological and geographic factors. Growth is better correlated (with negative correlation) to the presence of the disease. Countries with intensive malaria have income levels 33% those of similar countries without malaria. The countries that eliminated paludism in the past half century had an economic growth in the 5 years after eliminating malaria substantially higher than growth in the neighboring countries. In average, countries with intensive malaria grow 1.3% less per person per year, and a 10% reduction in malaria can be associated with 0.3% higher growth. The mechanisms through which malaria affects productivity remain unclear, nevertheless it can be speculated that malaria has a negative impact on tourism and foreign investment because, unlike other diseases, it makes no

distinction between rich and poor people, and also because has cumbersome or inefficient treatments. For the same reason, it hinders internal movements, thus stunting the transfer of technologies and knowledge between the industrialized and developed areas and the countryside. In addition, it causes recurrent or continual absenteeism both at work and mainly in school, strongly affecting education and compromising future competence (Gallup and Sachs, 2001).

The response of the international community to this disgrace can be represented by three general initiatives that afford more resources and presumably guarantee major efforts. Firstly, the eradication of the malaria was included in the United Nations Millennium Declaration: "Target 7: Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases" (UN Millenium Project, 2009). Secondly, the Roll Back Malaria initiative was launched in 1998 to coordinate the actions against malaria of more than 500 partners and institutions, specially including those nations where the disease is endemic. As quoted in their foundational declaration: "RBMs overall strategy aims to reduce malaria morbidity and mortality by reaching universal coverage and strengthening health systems. The Global Malaria Action Plan defines two stages of malaria control: (1) scaling-up for impact (SUFI) of preventive and therapeutic interventions, and (2) sustaining control over time (Roll Back Malaria, 2009). Methods for malaria control include vector control, interruption of vector-human contact (isolation of ill people, use of mosquito nets, and improving the access to and compliance with treatments).

Finally, the Medicines for Malaria Venture (MMV) was founded in to discover, develop and deliver new affordable antimalarial drugs through effective public-private partnerships (Medicines for Malaria Venture, 2009). Malaria is currently treated with blood schizontides they are used to treat acute infections and to quickly relieve the clinical symptoms, and may also act as chemo-prophylaxis for the parasite's RBC propagation for uninfected human hosts. Among them, the most widespread are chloroquine, quinine, artemisin, sulfadoxine and their derivatives. The parasite has recently developed resistance to drugs as a consequence of recurrent infections, mixing of populations due to human increased mobility and inefficient or poor drug treatments. The strategic response to this problem requires the discovery of new targets for antimalarial action, drug design and development of a treatment. Better understanding of the culturing protocols can eventually lead to the optimization of the resources, thus facilitate and reduce costs of research in medicine development.

1.2 Malaria and *Plasmodium falciparum*

1.2.1 Life cycle and pathogenesis

Malaria is initially characterized by fever, chills and paroxysm, headache, weakness, and other symptoms that mimic a viral disease such as flu. These symptoms reappear recurrently in continual or discrete episodes, eventually leading to severe or acute syndrome with generalized metabolic acidosis and severe anemia. *P. falciparum* acute symptoms include spleen, renal or multisystem failures and cerebral malaria. The later entails impaired consciousness with non-specific fever, generalized convulsions and neurological sequels. If not treated, cerebral malaria leads to a coma that persists for 24-72 hours, initially raisable and then unraisable, and it is fatal after 72 hours (Miller et al., 2002). The incubation period (the time from the initial malaria infection until first symptoms appear) generally ranges from 7 to 14 days for *P. falciparum,* 12 to 18 days for *P. vivax* and *P. ovale* and 18 to 40 days for *P. malariae*. Occasionally the incubation period may reach from 8 to 10 months for *P. vivax* and *P. ovale*.

All four species of *Plasmodium* that infect humans exhibit a similar hybrid life cycle with minor variations (see Figure 1.2). Human infection begins when **sporozoites** are injected with the saliva in a mosquito bite. The sporozoites enter the circulatory system and within 30-60 minutes will invade a liver cell (**hepatocyte**). After invading the hepatocyte, the parasite undergoes an asexual replication. This replicating stage is called pre-erythrocytic schizogony. It does not entail symptomatology and may last several days. Schizogony refers to the asexual replicative process in which the parasite forms a progeny of ready-to-invade parasites. The progeny, called **merozoites**, are released into the circulatory system following the rupture (*a.k.a.* **lysis**) of the host cell.

During blood schizogony, the intraerytrocytic cell cycle occurs and malaria symptoms manifest. Merozoites released from the infected liver cells invade erythrocytes (**RBCs**). After entering the RBC, the parasite undergoes a trophic period, during which it consumes the infected red blood cell (**IRBC**) contents and strongly modifies its membrane to allow an increased uptake of substances from the extracellular medium, followed by an asexual replication. The recently invaded IRBC (0-18 hours) is often called a **ring form** due to its morphology in Geimsa-stained blood smears. As the parasite increases in size, this 'ring' morphology disappears and the IRBC is called a **trophozoite** (18-35 hours). Nuclear division marks the end of the trophozoite stage and the beginning of the **schizont** stage (35-48 hours). Erythrocytic schizogongy consists of 3-5 rounds (depending on species) of nuclear replication followed by a budding process. Late stage schizonts in which the individual merozoites become discernible are called **segmenters** (42-48 hours). After

approximately 48 hours, the IRBC breaks down and 8 to 32 (average $\sim$ 16) merozoites are released. These merozoites invade new erythrocytes and initiate another round of schizogony (see Figure 1.3).

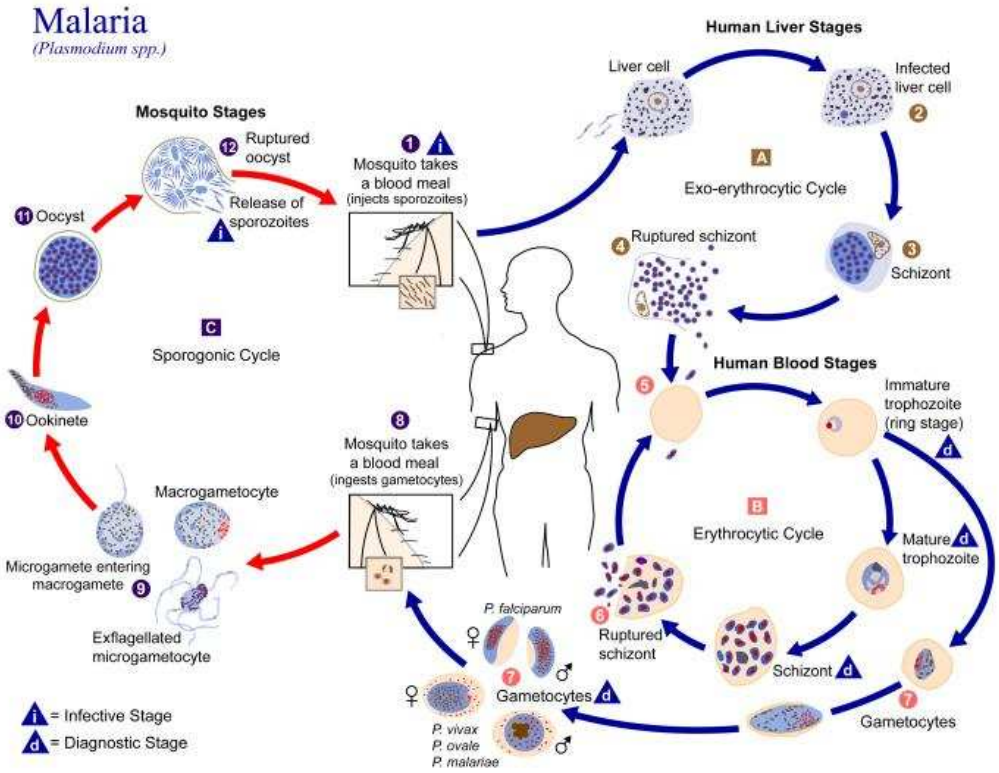

Figure 1.2: *Life cycle of Plasmodium spp. During a blood meal, a malaria-infected female Anopheles mosquito inoculates sporozoites into the human host (1). (A) Exoerythrocytic cycle: sporozoites infect liver cells (2) and mature into schizonts (3), which rupture and release merozoites (4). (B) Erythrocytic cycle: merozoites infect red blood cells (5). The ring stage trophozoites mature into schizonts, which rupture releasing merozoites (6). Some parasites differentiate into sexual erythrocytic stages (gametocytes) (7). (C) Sporogonic cycle: the male and female gametocytes are ingested by the mosquito during a blood meal (8). While in the mosquito's stomach, male and female merge generating zygotes (9). The zygotes become motile and elongated ookinetes (10) and invade the midgut wall, where they develop into oocysts (11). The oocysts grow, rupture, and release sporozoites (12), which make their way to the mosquito's salivary glands. Inoculation of the sporozoites into a new human host perpetuates the malaria life cycle (1). Reprinted from the Public Health Image Library (http://phil.cdc.gov/phil/home.asp). Image and caption from da Silva J. and Moser M., 2002.*

The blood-stage parasites of most of the *Plasmodium spp.* within a human host usually undergo a synchronous schizogony. The most obvious advantage of such a coordination is related to transmission, and to the nocturnal biting habits of the *Anopheles* mosquito (McKenzie and Bossert, 2005). *P. falciparum in vivo* synchronisation is not so clearly observed (Naughton and Bell, 2007).

It has been speculated that the synchronism among parasites is mediated by the host immune system through its circadian rhythms (Hawking, 1975), the rise of temperature associated with fever (Gravenor and Kwiatkowski, 1998), and specific responses to certain stages of the parasite (Rouzine and McKenzie, 2003). The paroxysms of malaria are associated with the rupture of IRBCs and the concomitant release of antigens and waste products. The dumped wastes arouses the immune system response, thus provoking the intermittent fever paroxysms associated with malaria. Fever, in turn, is assumed to affect the later stages of the IRBCs more, thereby leaving a population comprised solely of young ring-stage cells.

Blood stage schizogony in *P. falciparum* differs from the other human malarial parasites in that trophozoite- and schizont-infected erythrocytes adhere to other RBCs, forming aggregates of several RBCs (**rosettes**) and to capillary endothelial cells, disappearing from the peripheral circulation (**sequestration**) (Wahlgren et al., 1994). In severe *falciparum* malaria, the clumps formed by IRBCs and aggregated RBCs can obstruct capillaries and post-capillary venules due to the combined effect of rosetting and sequestration. This may lead to local hypoxia (deficiency of oxygen reaching the tissues of the body) and to the accumulation of toxic products. Obstruction of the micro-circulation in the brain (cerebral malaria) and in other vital organs is thought to be responsible for these severe complications of the disease (WHO, 2000).

Sequestration may be the factor responsible for the lack of synchronism observed in *P. falciparum in vivo*, as the IRBC remain stuck in the capillary vessels, safe from the increased response of the immune system triggered by the rise in temperature, and also avoiding going through the spleen (where blood is filtered and IRBCs would be destroyed). Other factors may play a central role in the kinetics of the infection *in vivo (White, 1998; McKenzie and Bossert, 2005). P. falciparum* merozoites are not as selective with the age of erythrocyte host as the other *Plasmodium* spp. are. Malaria parasites have a preference for young RBCs (Simpson et al., 1999), but *falciparum* merozoites are not so finicky. This means that the infection is not so strongly limited by the availability of suitable red cells (Ginsburg and Hoshen, 2002). Furthermore, *P. falciparum* IRBCs can produce cytokines (immunoregulatory proteins) and thus might have some control over the host immune response (Allan et al., 1993).

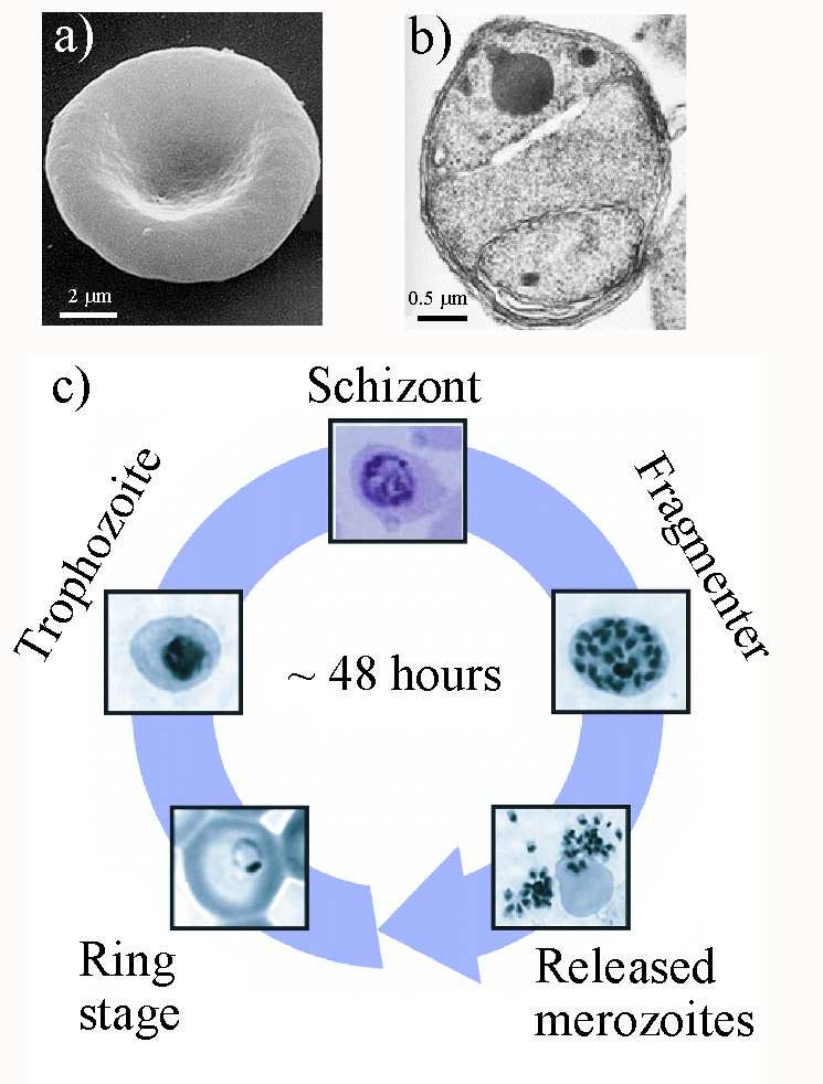

Figure 1.3: *a) Scanning Electron Micrograph (SEM) of a human RBC. From the Public Health Image Library (PHIL) ref. 11693. b) SEM of a Plasmodium falciparum merozoite. Kindly provided by Dr. Peter David, Unité d'Immunologie Moléculaire des Parasites, Institute Pasteur. c) Geimsa stained smears showing the major morphological stages throughout the intraerythrocytic cycle of Plasmodium falciparum (Bozdech et al. 2005).*

All in all, many factors may affect the virulence and severity of the disease, both host-related (control of the parasite population by the body is affected by the fitness and antigenicity of the immune system, the treatment and drug resistance) and parasite-related (multiplication capacity, red cell selectivity, control of the host immune response). The cardinal rules governing the intra-host dynamics of the infection (timing and magnitude of peak parasite density) are yet not fully understood (Heddini, 2002). The development of models and the use of statistical techniques for their fitting and comparison to real data may provide quantitative understanding and predictive capability of the pathogenesis of such a complex disease (Mideo et al., 2008).

1.2.2 Specific features of the intraerythrocytic schizogony

General description of the intraerythrocytic stages

RBCs are the hemoglobin (HgB) containing cells responsible for the transport of oxygen to the tissues in the human body, and comprise the **hematocrit**, which represents around 40% in volume of the circulating human blood. They take the form of flexible biconcave disks that lack a cell nucleus and organelles, and they cannot synthesize proteins. They are continually produced in the bone marrow of long bones at a rate of 10^6 RBCs per second and live for about 120 days. During the first 7 days, the developing RBCs are known as **reticulocytes** and comprise about 1% of circulating red blood cells. Their aging entails changes in their plasma membrane, making them susceptible to phagocytosis in the spleen, liver and bone marrow. Many of the important breakdown products are recirculated in the body (Guyton, 1987).

The cell division cycle of *P. falciparum* starts with the invasion of the host RBC by 1 or more haploid merozoites, and ends with the release of up to approximately 32 new merozoites that can go on to invade fresh cells (Granham, 1988; Tori and Aijkawa, 1998). Invasion of the host RBC consists on the forced invagination of the RBC plasma membrane and is mediated by the apical organelle of the parasite. The inner membrane of the extracellular merozoite rapidly degrades and disappears. During the next 48 hrs, the intracellular parasite develops surrounded by two membranes: the plasma membrane of the infected RBC (IRBCM) and the invaginated membrane, closely attached to the parasite itself, the parasitophorous vacuolar membrane (PVM) (Chitnis and Blackman, 2000).

During the first $\sim 20\ h$ of development, the young parasite is observed in light microscopy as the **ring stage**. Rings result from the extension of parasite cytoplasm. The main activity occurring during this stage is the adaptation for further feeding: the mod-

ification of the IRBCM and the production of the enzymes required for the intracellular metabolism. This stage corresponds to the quiescent or G_0 phase of the cell division cycle and beginning of the growth phase G_1.

Initially, the parasite lies apparently dormant. From $\sim 16\ h$ post-invasion, there is a progressive increase in the metabolic and bio-synthetic activity, i.e. growth of the parasite, degradation of the host cytosol and hemoglobin and major changes in the IRBCM. Digestion of hemoglobin occurs in the digestive vacuoles of the parasite and results in the production of amino acids and an insoluble pigment called hemozoin, which accumulates in the IRBC and is characteristic of the mature **trophozoite stage** ($\sim 20\ h - 35\ h$ post-invasion). The digested cytosol and hemoglobin serve as a source of the building blocks for further parasite duplication. Trophozoite stage corresponds to G_1(growth), S (DNA synthesis) and G_2(growth) phases of the cell division cycle.

During the last 14 hours of the cell division cycle, DNA replication and cellular differentiation of 8 to 32 daughter cells or merozoites occurs by asexual division (mitosis, or M phase). The first nuclear division occurs at $\sim 35\ h$ post-invasion (Leete and Rubin, 1996). Approximately 3, 4 or 5 divisions (presumably interspersed by periods of further DNA replication) are completed in a common cell cycle before the nuclei and other organelles are partitioned into nascent merozoites. The metabolic activity decreases all through the M phase. The IRBC in former M phase is known as the **schizont stage** ($\sim 35\ h - 42\ h$ post-invasion); the latter partitioning events are recognized as the **fragmenter stage** ($\sim 42\ h - 48\ h$ post-invasion).

Membrane transport and metabolism

The exchange of substances through any cell membrane is a diffusion-mediated process constrained by four factors: (1) the number and capacity of the diffusion channels; (2) the intrinsic rates of diffusion and reaction in the intracellular medium; (3) the geometrical shape and dimensions of the cell; and (4) the concentration of the molecular species in the immediate extracellular medium, just across the cell membrane. The transport of substances through the RBC membrane is considered to be limited by the activity of membrane transporters and usually modeled with more or less sophisticated versions of the Michaellis-Menten model for enzyme activity. A simpler alternative is to consider only the geometrical limitations and assume Blackmann kinetics (Ferrer et al 2008a).

To reach the parasite, the extracellular substances must cross the IRBCM and the PVM. The activity and number of the transport systems of the IRBCM undergo marked alterations through the infection cycle. The homeostatic state in the IRBC cytosol changes consequently, and the geometry of the PVM is extremely complicated, with mul-

tiple invaginations that directly connect the trophozoite with the extracellular medium (Kirk, 2001). In conclusion, uptake kinetics for the IRBCs can not be easily modeled even when models of the transport through the healthy RBC membrane are already available. In this case, Blackmann kinetics are also a good alternative.

RBC metabolism has been well-studied in the last four decades and extensive biochemical data have been accumulated (Kauffman et al., 2002). Schematically, the RBC mainly uptakes glucose from the environment and processes it through the glycolysis pathway to generate ATP and lactate. The ATP molecules are consumed mostly for maintenance of RBC homeostasis (cation transport, mainly through the $Na^+ - K^+$ pump, in order to maintain the cell volume, electroneutrality and osmotic balance). Other important metabolic pathways are those involved in the control of the oxidation state of HgB, the anabolism of nucleotides and the continual production of sugars through the pentose phosphate pathway, related to both the maintenance and reconstruction of the cell structure.

The uptake rate as a function of the extracellular concentration alone is not fully understood for RBCs. The available models for carrier mediated facilitated diffusion, such as the Michaellis-Menten models are insufficient. The transport of glucose in RBCs has recently been reported to consist of two different phases for the transport of glucose (fast and slow) that operate under different cellular conditions (low or high intracellular nutrient concentration). The metabolic state of the RBC and the internal availability of ATP play a central role in the modulation of these two phases (Leitch and Carruthers, 2007).

Metabolism of IRBC has been extensively studied and described in the search for potential targets for drug activity (Sherman, 1998a; Rosenthal and Meshnik, 1998; Sherman, 1998b). Schematically, the parasite requires a great amount and variety of mononucleotide precursors which can not be bio-synthesized to carry out DNA duplication. Some of these substances can be obtained from the digestion of the RBC cystosol and HgB but many others (such as purine or pyridine) must be obtained from the extracellular plasma. For this reason, during its trophic stage the IRBC uptakes macromolecules from the surrounding medium. In order to do so, the IRBCM is profoundly altered. An increased permeability allows the ingestion of macromolecules but completely perturbs the osmotic equilibrium of the IRBC with the extracellular surroundings. To maintain homeostasis, the parasite digests an excess of HgB and cytosol and compensates the increased permeability with adapted concentration gradients and pump activity (Lew et al., 2003). This results in an increased need for ATP. Glucose uptake (and lactate production) varies through the infection cycle and may reach 100-fold the average uptake of healthy RBCs (Roth, 1990).

During the last decades, increasingly detailed kinetic models of RBC metabolism have been developed (Rapoport et al., 1976; Palsson et al., 1989). They use large arrays of coupled nonlinear ordinary differential equations that represent the biochemical reactions occurring within a single cell and its immediate environment (Kuchel, 2004). Such an approach, known as systems biology, has produced a group of testable models that simulate cell behavior and reproduce experimental results. Steady state solutions and predictions of the *in silico* models of the RBC can be compared to data obtained through NMR spectroscopy (Jamshidi and Palsson, 2006). Recently, a systems biology model of the IRBC has been developed and published (Mauritz et al., 2009). Its results corroborate that the IRBC approaches lytic rupture as the infection cycle goes by, becoming increasingly fragile.

Models from systems biology are too complex to be included in macroscopic models of microbial communities. They usually operate at temporal scales ranging from 10 μs to 1 ms and spatial scales around the micrometer. In addition, these models are not spatially explicit and consider homogeneous populations (with the same average behavior for every cell) and homogeneous spatial distributions of the modeled cells.

Merozoite release and invasion of RBCs

At the end of each asexual reproduction cycle ($\sim 48\ h$), ~ 16 merozoites egress from their host cell ready to invade healthy RBCs. This process is physically hindered by two membranes: the PVM and the IRBCM. The long-running debate about the mechanisms that underpin the egress of merozoites from the host cell basically offers three alternatives:

(i) Membrane fusion. The IRBCM and the PVM fuse, creating an opening through which clustered merozoites are ejected to a single point in the immediate neighborhood $\sim 5\ \mu m$. The IRBC ghost first remains in place, then degrades over time (Winogard et al., 1999).

(ii) Membrane breakdown: The IRBCM and the PVM degrade so that merozoites remain in the place previously occupied by the IRBC. Remnants of the IRBC disperse as segregate macromolecules (Wickham et al., 2003).

(iii) Exploding flowers: An increase in merozoite size and concurrent breakdown of the RBC cytoskeleton cause a physical pressure that blows the IRBC apart, scattering both the invasive merozoites and the clotted remnants of the IRBC to a local range up to $\sim 25\ \mu m$ (Glushakova et al., 2005).

Recent video recordings suggest that model b) best fits observations (Rayner, 2006).

Extracellular merozoites, like all apicomplexans, have motile systems that allow them to glide- that is: to move towards a fixed destination, maintaining their orientation and without cell deformation. However, the merozoite has not been seen to glide, although it is not known how it behaves in the confines of a venule crowded with adherent schizonts. *In vitro*, merozoites do not move across a substrate at all, until they contact the surface of a healthy RBC (Pinde et al., 2000).

Once the surface coat of the extracellular merozoite contacts the surface of a healthy RBC, it provisionally attaches to it. Then, it reorientates so that the apical complex faces towards the RBC membrane. Once the reorientation is completed, a stable junction between the apical complex and the RBC membrane is set. Then, the attachment becomes irreversible and the invasion of the RBC takes place: the membrane of the RBC increasingly invaginates, and the parasite pulls itself to the center of the RBC via the proteins allocated on its surface coat. The invasion process lasts ~ 30 seconds and is not always successfully carried out (Chitnis and Blackman, 2000). Merozoites are short lived in the extracellular medium $\sim 20 - 30$ minutes (Barnwell and Galinski, 1998). Beyond this time span, they seem incapable of RBC invasion. As a result, the spreading of merozoites through the hematocrit in *in vitro* cultures occurs solely after the lysis at the end of an infection cycle or after a frustrated invasion attempt. This spreading lasts a few minutes and has a short range.

RBCs containing more than one parasite (two, three and eventually four) are commonly observed in blood smears taken from patients and also in *in vitro* cultures. The tolerance for multiple invasion can be explained because changes in the IRBCM take place slowly (all through the ring stage), and also because the local geometric constraints on the IRBCs under the conditions in which the propagation of the parasite takes place (mainly in venous capillaries, with no blood circulation) ensures that immediate neighboring RBCs are surrounded by an excess of potential invaders.

The study of the multiple-infected RBCs *in vitro* can provide information regarding how the RBC susceptibility to infection varies with RBC age (Simpson et al., 1999). Alternative measurements also show that *Plasmodium falciparum* has preference for retyculocytes, but that its selectivity is not as strong as for other *Plasmodium spp.* The susceptibility to invasion of retyculocytes is between 2 and 4-fold the susceptibility of grown RBCs (Pasvol et al., 1980).

Physical modifications of the IRBCM and interactions among RBCs

Many changes affect the RBC during the intracellular development of *P. falciparum.* Many of these changes mimic those of apoptosis (programmed cell death) with the difference that the latter are ATP-dependent, meaning that they require the activation of specific cellular mechanisms by the RBC and are triggered by osmotic shrinkage. In contrast, the osmotic ability of IRBCs is maintained by the parasite and any one-time scarcity of energy (glucose) may lead to cell lysis, the disintegration and dissolution of the cell(Sherman et al., 2004). The alteration of the IRBCM is completed by the end of the ring stage, and the osmotic compensation requires an increasingly excessive digestion of host HgB and cytosol after $\sim 24\ h$ post-invasion, which in turn requires an additional energy demand (Lew et al., 2003). All in all, it is reasonable to assume that IRBCs are more fragile and sensitive to perturbations as they approach the end of the infection cycle. This argumentation is consistent with the observation of an increased rate of lysis of the mature IRBCS, and it has been speculated to be responsible for the synchronization of the parasite population after fever paroxysms (Kiatkowski, 1989). Paradoxically, low temperatures also seem to affect parasite development only after $\sim 30\ h$ post-invasion (Rojas and Wasserman, 2007). This suggests that the fragility of IRBCs is not related to the denaturing of the enzymes by febrile temperatures, but to a complex counterbalance of the alterations produced by the parasite.

The shape and surface of the IRBCs are substantially altered. The biconcave RBC disc becomes more spherical and rigid, and it also acquires a sickled curvature, responsible for the name *falciparum.* The surface becomes wrinkled and covered with sub-microscopic protuberances called **knobs**. Knobs are elevations of the plasma membrane of $\sim 25\ nm$ height and 100 nm diameter, with increased electron density. The number and size of knobs vary with the development of the parasite. Rings have no knobs and, while the early tropozoite has numerous small knobs, the schizont has fewer and bigger protuberances. The function of knobs is basically to produce adhesiveness. This adhesiveness serves the parasite in two ways. Firstly, it causes cytoadherence: the IRBC sticks to the endothelium of the circulatory vases, giving rise to the sequestration in capillaries. Secondly, it is responsible for the aggregation of IRBC with many RBCs in clumps called **rosettes** (Wahlgren et al., 1994).

Higher parasitemia is associated with rosetting phenotypes, so the role of rosetting is to enhance parasite growth and survival by facilitating invasion or promoting immune evasion *in vivo* (Rowe et al., 2002). However, rosetting does not play a role in the invasion or targeting of parasites into uninfected cells *in vitro,* even when the culturing conditions are similar to microvascullar flow (Clough et al., 1998). This suggests that the role of

Entity	Characteristic	Value	Source
RBC	density	1.1 g/ml	Guyton, 1987
	volume	85 μm^3	
	diameter	$6-8\,\mu m$	
	width	$1-2\,\mu m$	
	max. age	120 days	
IRBC	infection cycle	48 h	Torii and Aikawa, 1998
	ring stage	18 h	
	trophozoite stage	17 h	
	schizont stage	7 h	
	fragmenter	6 h	
Merozoite	volume	2 μm^3	Torii and Aikawa, 1998
	endurance	30 minutes	
Plasma	Hematocrit	$5x10^6$ $RBCs/ml$	Guyton, 1987
	[glucose]	65 mg/dl	
	[oxygen]	1.4 g/l	
	pH	7.4	
	proteins	8% volume	

Table 1.1: *Some characteristic features of the healthy erythrocytes (RBCs), infected erythrocytes (IRBCs), extracellular merozoites and blood plasma.*

rosetting is not to guarantee the proximity of infectable RBCs, but to prevent immune response.

Rosettes usually bind a few RBCs to a single IRBC, but they may contain from 10 to 20 IRBCs and up to 50 uninfected RBCs (Wahlgren et al., 1994). Rosetting adherence forces resist the physiological shear forces that are encountered in venous circulation $0.1-0.5\ Pa$ (Chotivanich et al., 2000). Healthy RBCs, in turn, are also subject to small attractive forces. They are known to form aggregate structures in culture called **rouleaux**. The adhesion binding energy between two healthy RBCs has been estimated to be from $6\cdot10^{-4}\,N/m$ to $1\cdot10^{-3}\,N/m$ (Hochmuth and Marcus, 2002). This inter-cellular attraction produces a net inwards pull on the cells of the border of the hematocrit, so these will exhibit a property similar to the surface tension of liquids (Foty and Steinberg, 2005) (see Figure 1.4).

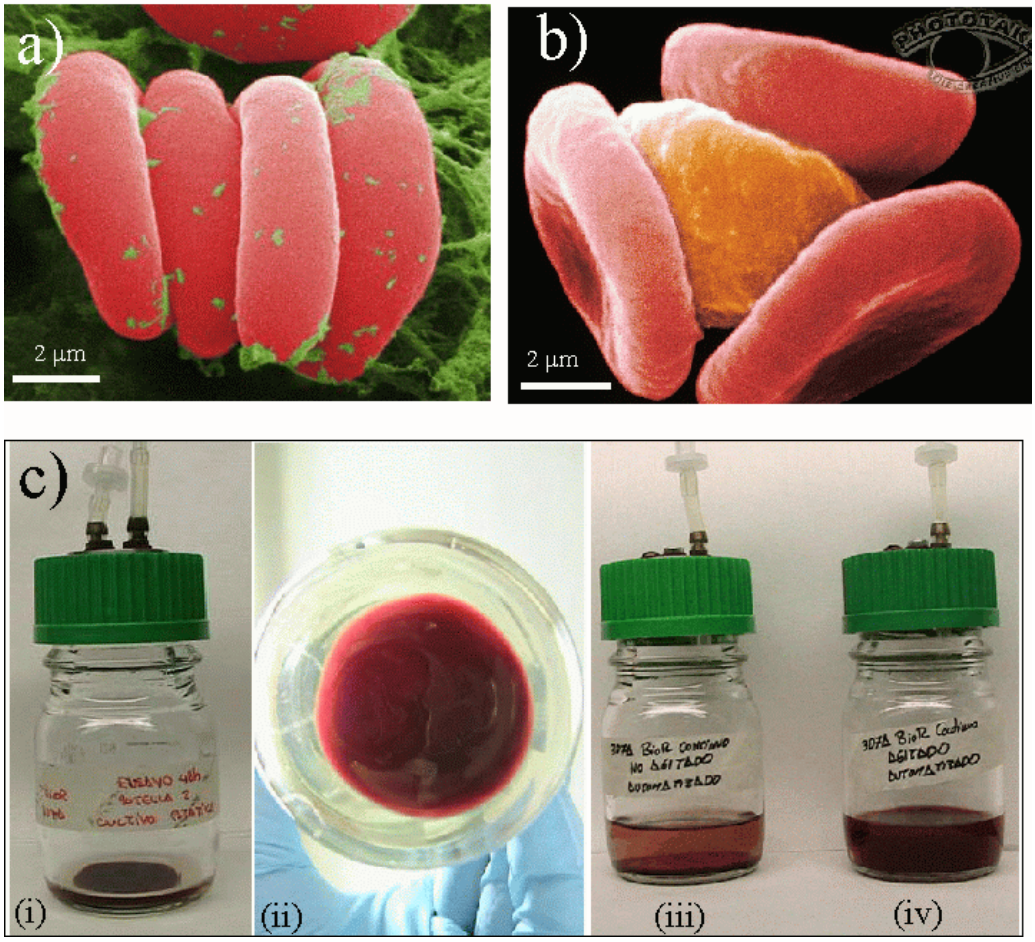

Figure 1.4: *a) Scanning electron micrograph (SEM) of erythrocytes grouped in a Rouleaux formation. ref.: PHIL-1092. b) SEM of a rosette formation comprised of an IRBC surrounded by three healthy RBCs.*
c) Static hematocrit layer shaped like a sessile drop: i) side view and ii) top view. Attractive forces that bind settled hematocrit RBCs are weak, so hematocrit easily deaggregates with minimum perturbation: iii) after medium renewal and iv) in a suspended culture. Images kindly provided by Dr. J. Vidal, Experimental Microbiology Group, GlaxoSmithKline.

1.3 Cultivation methods

1.3.1 *In vitro* cultivation of erythrocitic stages

Most species of malarial parasites and most stages of *P. falciparum* cannot be routinely maintained in cell culture. Even the erythrocytic cycle of *P. falciparum*, which can be cultured, is very slow, labor intensive, and expensive to maintain and propagate.

P. falciparum-infected erythrocytes were first preserved *in vitro* with subsequent infection capacity in 1974 (Pavanand et al., 1974). IRBCs were preserved for at least 64 hours with re-invasion and no loss of viability. Different storage times of the RBCs were tried. The current methods in use for *in vitro* propagation of the parasites are based on the *candle-jar* method, which was first developed by Jensen and Trager (1976).

This method keeps a thin layer of RBCs incubated in culture plates (petri dishes), at 38^{o}C, in 20 ml of artificial culture medium (RPMI1640; see table 1.2 for an outlined specification.) mixed with human serum, at around 10% hematocrit in volume and in a low oxygen atmosphere. **Subcultivation** of the population is required approximately twice a week to limit culture infection ratios below 10% in the parasite load. This means that a sample of the culture containing both healthy and infected RBCs is derived to another culture vial with just healthy RBCs. Samples are diluted to a 1:3 - 1:4 fraction with healthy hematocrit suspensions. This ratio corresponds approximately to the one that results in a $0.5\% - 1\%$ of parasitized cells (*a.k.a.* parasite load or **parasitaemia**). The medium is renewed daily, but not the day after subcultivation, through a gentle drain that does not perturb the hematocrit layer. The easiest way to get the low oxygen atmosphere is to keep the whole system hermetically closed together with a candle. After each time the system is manipulated, the candle is lit and allowed to extinguish itself, thus consuming most of the oxygen.

In the same article, the authors presented the *continuous-flow* method. It differs from the *candle-jar* method in that the culturing medium is continually renewed instead of replaced at discrete events. This achieved through a constant input flow with a peristaltic pump operating at an input rate of 50 ml/day.

Both methods allow preserving the cultures for long-term stable cultivation with parasite growth rates reflecting 6-7-fold increases per cycle, and with full invading capacity and no observable cellular damage of the IRBCs. This opened the door to the *in vitro* cultivation of the parasite for pharmacological and immunogenic trials. The ratios between the cell and medium volumes, the thickness of the cell layers (Jensen and Trager, 1978), and the effect of using different RBCs sources and storage times (Capps and Jensen, 1983) or strains of *P. falciparum* (Chin and Collins, 1980)were studied at the time.

In later years, different studies regarding the same method were carried out. The role of pH, glucose and lactate (Jensen et al., 1983), the synchronization of the cellular cycle through the damage induced by drugs (Lambros and Vanderberg, 1979) or by the effect of febrile temperatures (Kiatkowski, 1989), and the variations found in the growth rate when using different substitutes of the serum were analyzed.

With the ultimate goal of building semi-automated, bio-reactor-like culturing methods, the harvests from static and agitated culture systems were compared (Butcher, 1981; Butcher, 1982). The results were not conclusive and the heuristic assumption that agitation would harm or be deleterious for parasite development remained (Jensen, 1988). At the same time, many semi-automated methods based on the continuous flow protocol were proposed. However, apparently none of the proposed apparatus offered a sufficient increase in the performance to warrant the cost of building and maintaining it (Trager, 1994). Nowadays, there are some published standardized protocols for culturing *Plasmodium spp.* (MR4, 2008). These methods are published by the *Malaria Research and Reference Reagent Resource Center* (MR4), and they are periodically updated.

Some of the experiments presented above have been used to build, calibrate and validate the model. Their detailed description can be found in Section 2.3.

1.3.2 Cultures performed by EMG-GSK

The bulk of the simulations were compared to the experimental work carried out by the EMG-GSK. They had particular questions that required the application of the computational model (e.g. what are the best storage timings for the RBCs, which is the optimal design for the static cultivation of the parasite, how to assess the effect of adrug when there are very small populations, among others). Additionally, some experimental trials were designed to test the predictions and patterns produced by our model, in order to check its validity. Unless specified otherwise, *P. falciparum* infected RBCs in *in vitro* cultures were raised under the culturing conditions according to the MR4 protocols published in 2005. These general culturing conditions are outlined in table 1.2. A list containing all the experiments performed by EMG-GSK is presented in Appendix B.

The measurements performed on the culture system may give information regarding different aspects of the infection course. Parasitaemia represents the extent of the infection. Number of lysis and damaged RBCs can give information regarding the stress suffered by RBCs and IRBCS. Age structure provides information regarding the degree of synchronicity of the culture. Finally, the percentage of multiple infections can yield information both regarding the spreading of the parasite through the hematocrit and the

Input measurements	Initial hematocrit ($\%H_0$), volume and shape of the culture system, initial parasitaemia ($\%I_0$)
RBC source	Human AB-, supplied by Spanish Red Cross. Variable storage time.
Initial hematocrit ($\%H_0$)	5 % - 15% Hematocrit
Parasite source	*P. falciparum* 3D7A supplied by the MR4
Initial Parasitaemia ($\%I_0$)	0.1 % - 3 %; Not applicable to recrudescence trials.
Culture medium:	RPMI 1640, 25 mM HEPES, 10% human serum, 0.15 mM Hypoxanthine.
C_{gluc}	2.67 mM/l
C_{lact}	0 mM /l
pH	7.2
Culturing method:	static, agitated, suspension or recrudescence
Medium renewal	Continuous medium renewal or daily punctual procedure
Culturing conditions	Temperature: 37º C; Athmosphere: 5% O2, 5% CO2, 90% N2
Subcultivation:	discrete events: daily or each 2, 3, or 4 days
Dilution ratio (ν):	1/5 - 1/3 of culture volume
Output measurements	parasitaemia ($\%I$), Growth ratio (GR), age structure of IRBCs, % lysis or damaged IRBC counts, % multiple infections

Table 1.2: *General conditions of the cultures performed by the EMG-GSK. Detailed specification sheets for RPMI1640 culturing medium and for human serum may be found at commercial suppliers.*

preference for certain RBCs (young RBCs). They are all obtained by extracting samples from the culture and counting the observed forms of the parasite and RBC.

The **growth ratio** per infection cycle (GR) measures the proliferation of the parasite or propagation of the infection. It is obtained dividing the final by initial parasitaemias through an infection cycle. This measurement can only be carried out for very synchronous cultures. Growth ratios at fixed time spans are used instead to assess the propagation of the infection. They are calculated from the parasitaemia measurements at consecutive samplings. Typically, they comprise 12-hour growth ratio (GR_{12}), daily growth ratio (GR_{24}), or ratio observed at each 48 hours (GR_{48}), through the average duration of the parasite intraerythrocytic life cycle.

1.3.3 Pending issues

More than 30 years of collective experience in the cultivation of *P. falciparum* have provided a vast amount of heuristic knowledge and a broad set of rules that are considered to be commonsense when tackling *in vitro* cultures. Some of this evidence is revisited by INDISIM-RBC in Chapters 2 and 3. Additionally, some limitations in parasite development have been observed in experimental static culture systems, in particular, and are not yet fully understood.

a) The culturing medium must be replaced daily by fresh medium, except the day after subcultivation. Medium renewal does not entail the agitation of the hematocrit layer. Variations in the total amount of culture medium do not drastically modify the culture requirements for renewal.

b) Subcultivation of red blood cell is carried out in order to control the parasitaemia. Custom static in vitro culture systems usually support 5% parasitaemia. Whenever the parasitaemia exceeds a 15% threshold, IRBCs undergo excessive stress and cultivation is no longer feasible. It has been claimed that increasing the culture medium to hematocrit ratio allows for higher parasitaemias (MR4, 2008); recent publications propose methods for achieving such high infections (Radfar et al., 2009).

   Furthermore, cultures where the red blood cells are maintained in suspension (through the continual agitation of the culture system) do not present this limitation on the parasitaemia. Subcultivation must still be carried out every three or four days to maintain the viability of the culture, but the ratio of IRBCs to healthy RBCs may reach up to 60% in this case.

c) Culture systems must be of an appropriate size. Culture volumes typically are $\sim 2-100\ ml$ and each volume must be cultured in a dish of appropriate surface. The hematocrit is typically set in layers up to $3\ mm$ depth.

The usual explanations for the observed limitations are commonsense yet quite vague. For instance, the medium must be replaced because it is exhausted, RBC population must be subcultured because otherwise no healthy RBCs would be available for contagion, and the size of the system must be appropriate for optimal handling.

A more thorough examination of these and other limitations would increase the insight concerning the system, and could contribute to improving the current standard protocols. For instance, to have a bioreactor to cultivate malaria (and extensively other infectious diseases) would aid pharmacological research. Such a device should be easy to use and reliable, and allow for automated cultivation of the parasite with minimal RBC and medium expenses. In order to build an automated bioreactor-like culturing device, the following questions (conceptual and applied) regarding the current protocols require detailed answers:

1. Why is the culture medium continually replaced? Are there alternative methods for maintaining an appropriate environment for the RBCs?

2. Why are IRBC populations subcultured every 4 days at most? Can this limitation be skipped?

3. Why is there a maximum threshold of culture viability at $10\% - 15\%$ parasitaemia for static cultures but $\sim 60\%$ for suspended cultures?

4. How does the storage of RBCs prior to cultivation affect the culture yield? Which is the optimal storage period?

5. How do IRBC populations get synchronized or desynchronized? How does the degree of synchronicity affect the population dynamics?

6. How does the size of the culture system affect the parasite performance? Which is the optimal size for static cultivation?

7. How does the agitation of the culture system affect the propagation of the infection? Whihc is the optimal agitation regime?

Answering these questions requires a detailed scaffolding, a model that directly tackles the system formed by interacting healthy and infected RBCs.

Better understanding of the culture systems is also required in many other domains regarding the study of *P. falciparum.* An example is the current interpretation of the experimental results of drug trials. The observed relation between the dose and exposure to drug treatments and the subsequent lag in the infection detection times in treated cultures is not yet fully understood. The amount of ill cells that are destroyed or harmed, and the recovery times for the harmed surviving IRBCs can not be extracted from the treatment outcomes. As a consequence, the effect of drugs *in vitro* is hardly translated into an expected effect in the systems *in vivo*, which may lead to suboptimal choices on which drugs should be selected for animal test trials. Appropriate models of the data results are required to facilitate decision-making in this field.

Model and simulation of the behavior of IRBCs is also fundamental for the interpretation of more complex systems. One particular case is the response of experimental murine models. These animal models consist of genetically modified mice with a reduced immune system that can be engrafted with human erythrocytes. Later, infection can be inoculated into each mouse and the response of the subject to specific treatments can be assessed. The *in silico* model could be a tool to provide estimations for relevant parameters in the animal model that can not be easily inferred from measurements on the real systems.

1.4 Models of the infection process

1.4.1 Levels of description and other classification criteria for mathematical models

Mathematical models are simplified representations of the reality used in natural sciences. They use mathematics to describe and understand the features (dynamics, structure and phenomena) of the systems under study. Models in microbiology may tackle microbial ecosystems at different levels of description. Three different levels of description can be distinguished depending on their specific spatial (L) and temporal (T) scales: molecular, $L \sim 10^{-9}m$ and $T \sim 10^{-6}s$ (e.g. metabolic reactions, membrane transport, protein folding); cellular, $L \sim 10^{-6}m$ and $T \sim 10^{-3}s$ (e.g. whole cell systems); and populational, where $L \sim 10^{-3}m$ and $T \sim 1s$, at a minimum. Last, it must be stressed that the characteristic scales of the populational or community level of description may vary depending on the system and process under study. For instance, the appropriate scales to study certain aspects of bacterial population dynamics in planktonic ecosystem might be on the order of $L \sim 10^{3}m$ and $T \sim 1\ year$.

Besides the level of description, models in microbiology can be classified into dual

categories, following other criteria:

Mechanistic-Empirical: Depending on whether the model attempts to unravel the mechanisms that drive a particular observed behavior (e.g. how gravity and air friction determine the trajectory of a cannonball) or simply tries to describe it (e.g. best fit to experimental data).

Spatially Explicit-inexplicit: Depending on whether the model explicitly considers (or not) the position of the modeled items. The simplest way to avoid making the spatial configuration explicit is to consider a mean field approximation, to replace all local interactions with an average or effective overall interaction.

Dynamic-Static: Depending on whether the model accounts for a process (e.g. the evolution of landscape in successive stages of deforestation) or the outcome of a steady or one-time situation (e.g. the best bet in a roulette game).

Stochastic-Deterministic: Depending on whether the model considers randomness or not, both as uncertainty in the description of the processes and as diversity in the values of the variables.

Continuous-Discrete: Depending on whether the model considers a continuum or a set of one-time actions occurring at discrete events and at distinct sites.

Analytic-Numerical: Depending on the techniques used to assess the outcome of the model. Whether an exact result of the model can be obtained or it must be solved using approximate numerical calculations, often with a computer.

These classification criteria do not pretend to be exhaustive, but just orientational. A more thorough description of the presented categories can be found in the thesis of A. Standaert (2007).

1.4.2 Strategies in modeling complex systems

Two alternative strategies can be followed to structure information and build models: the **top-down** approach (comprised by **Population-based Models** or "PbMs" in the field of ecology) and the **bottom-up** approach (consisting of **Individual-based Models** or IbMs in the field of ecology).

The top-down approach is the most widespread in science. It is also referred to as decomposition or analysis and builds models through the breaking down of the system and the describing of each of the parts that compose it. In the context of ecology, this

approach aims to disclose the behavior of individuals and the rules that govern them from the observations of the whole system or parts of it. Models built at a population level of description are a particular type of **System-based Models**: they consider variables that characterize the population and a set of laws governing them. These rules are usually formalized with differential equations which are ultimately based on assumptions regarding the behavior of individuals. System-based modeling consists on: defining the relevant variables of the system and proposing a set of rules governing them, applying these rules (i.e. solving the equations), and assessing the validity of the model through the comparison of its results with experimental observations.

The bottom-up approach is synthetic thinking, i.e. the combination of simple assumptions to give rise to a complex whole. It consists on piecing together detailed bits of information to subsequently ascertain how the linked system behaves. **Individual-based modeling** is an established discipline in theoretical ecology. It constitutes both an approach and a set of techniques that differ from the ones used in the traditional system-based models (Grimm and Railsback, 2005). Basically, Individual-based Models in microbiology consider each cell a discrete entity characterized by individual variables, besides their age and spatial location. IbMs constitute a shift of paradigm and follow a bottom-up approach. They focus on the emergence of aggregate (and sometimes unexpected) behaviors observed in ecosystems from the detailed description of their components. The cornerstone of IbMs is that simple rules defined at an individual level can generate complex behaviors at a macroscopic level of description (Uchmanski and Grimm, 1996).

Models built at an individual level are a type of **Agent-based Models** (AbMs): these are computational models that simulate the actions and interactions of the autonomous components of a system, in order to assess their collective behavior. Individual-based modeling requires using heuristics and specific knowledge to define the individual entities and the rules governing them, programming tools to put together the modeled rules and mechanisms, and statistical analysis to extract information from the simulation outcomes and compare them to experimental observations (Bonabeau, 2002). In recent years, the Individual-based approach has been increasingly used in the field of microbial ecology (Hellweger and Bucci, 2009).

In this thesis, the bottom-up approach is most commonly used: the heusristic assumptions of the experimentalists are sistematically put to the test because the foremost aim of the thesis is to gain understanding of different culture systems. Nevertheless, continuous models are also used to check the consistency of the IbM structure and predictions and to connect the IbM with macroscopic observations (see Section 3.5). The problem

of choosing and developing a modeling strategy in the domain of microbial population dynamics is tackled in depth in Chapter 4.

1.4.3 Models of the spreading of diseases

The evolution of the parasitic infection can be regarded as the population dynamics of two coexisting populations: the parasite and the host cells. In this case, tackling the potential factors that regulate the spreading of the disease within a host or through an *in vitro* culture may resemble solving a traditional problem in ecology: population regulation in an ecosystem composed of two competing species (Haydon et al., 2003). This kind of approach allows posing general questions on the mechanisms that govern parasite proliferation and to predict the expected population dynamics for not too complex systems and for steady states. An example of such an approach is the Nicholson-Bailey Model (Bailey, 1933; Nicholson, 1933), which is closely related to the logistic model of Lotka-Volterra for the population dynamics of two interacting populations (Lotka, 1925).

An alternative approach is to consider the sub-populations of healthy and infected individuals and model the disease as a transference from one group to the other. This approaches assume a given model for the transfer between the two groups in order to study the evolution of the system as a whole, with the ultimate aim of improving the procedures and treatments to control the disease. For instance, models of the disease *in vivo* are used to: (i) infer the effective multiplication rate and degree of sequestration in the microvascullary system from the observed parasitaemias in the peripheral blood (White et al., 1992), (ii) describe the regulation of the parasitic population by the febrile temperatures and *vice versa* (how the population may give rise to fever) (Gravenor and Kwiatkowski, 1998), or (iii) study the response to drug therapy and determine the optimal susceptible targets for further drug development, among others (Chyiaka et al., 2008).

The building block of any epidemiological model is the Kerman-McKendric model or SIR model, which computes the theoretical number of individuals (*a.k.a.* agents) infected with a contagious illness in a closed population over time. The name of this class of models derives from the fact that they involve coupled equations relating the number of susceptible agents $S(t)$, infected agents $I(t)$, and recovered (or dead and removed from the infection process, but still counted within the population) agents $R(t)$.

The most basic SIR model assumes that the population size is fixed (i.e., no births or deaths by natural causes), that the incubation period of the infectious agents is instantaneous, and that the duration of infectivity is equal to the span of the disease. It also assumes a completely homogeneous population with no age, nor spatial or social structure. In this case, the model consists of a system of three coupled nonlinear ordinary

differential equations:

$$\frac{dS(t)}{dt} = -\beta SI \tag{1.1}$$

$$\frac{dI(t)}{dt} = \beta SI - \gamma I \tag{1.2}$$

$$\frac{dR(t)}{dt} = \gamma I \tag{1.3}$$

where β is the infection rate, and γ is the recovery (or mortality) rate.

The key value governing the dynamics of such a system is the **reproduction number** (R_0), which can be regarded as an epidemiological threshold: $R_0(t) = \frac{\beta S(t)}{\gamma}$. It represents the subsequent potential infections caused by a single infection, and is determined by the total number of "contacts" that occur since one agent gets infected until it recovers or dies. The value of R_0 is crucial in this model: when $R_0 < 1$, each agent that contracts the disease will infect fewer than one individual before dying or recovering, so the outbreak will run down until it gets extinguished. When $R_0 > 1$, each person that gets the disease will infect more than one person, so the epidemic will spread until reaching every single agent. In this case, the number of infected individuals exponentially increases until healthy individuals are no longer available and $I(t)$ follows a sigmoidal curve. The greater R_0, the steeper the slope of $I(t)$. A simple but surprising result drawn from this model is that for the population as a whole it is better to have epidemics with high mortality, because they dissipate more easily (Anderson and May, 1979).

The SIR model can be made more realistic by introducing increasing complexity. One first step would be to introduce continuous or discrete inputs (for instance, a growth rate of healthy agents, μ) and/or output of agents (for instance, a constant death rate of agents, equal to the growth rate, μ). This leads to stable, oscillating or chaotic population dynamics in the steady state (Kuske et al., 2006). Studying the temporal patterns observed in these models gives information regarding the characteristic temporal scales of the propagation of the infection.

More elaborate models can consider that the recovered individuals become infectable after a certain period (i.e. SIRS models) (Aiello et al., 2000), or define additional stages in the infection process, e.g. SEIR models S-susceptible, E-exposed, I-infective, R-removed (Li et al., 2001), where the exposed stage corresponds to those already infected individuals that are not yet infective themselves. Alternatively, the model can be structured to explicitly represent different parts of the system, for instance to consider different ages for the agents (White et al., 1992). The population is then divided into a set of interacting sub-populations, each one governed by a set of equations 1.1, 1.2 and 1.3. The set of

equations are coupled with each other. One could split hairs to add all these and more refinements into a model to get more detailed and structured models with bigger matrices of coupled equations, but there is the risk of missing the forest for the trees.

A qualitative step in the complexity of the model is to consider the spreading of the infection as a random process occurring between individuals. This is assuming that the single effective contact between an infected and a susceptible individual, as well as the death or recovery of an infected agent, are stochastic phenomena occurring with a given probability. Then, the spreading of the disease can be described as a Markov process (where the future state of the system at a given time is a stochastic function of the state at the present time). Such processes can be studied with computational techniques to deal with random processes, such as the Monte-Carlo simulation methods (Metroppolis and Ulam, 1949). This way of tackling the problem is called the Reed-Frost model, and must be introduced when we consider finite populations (Abbey, 1952). It provides qualitative differences from the Kerman-Mckendrick model outcome mainly when small populations are taken into account. The most relevant difference is that disease can persist even when $R_0 < 1$ and may peter out although $R_0 > 1$. The nonrandom Kerman-McKendrick model can be considered as the average outcome of the Reed-Frost model when populations are big enough and when we consider an statistical ensemble of simulations, when the model is run several times with different seeds for the random numbers (Hoppensteadt and Peskin, 2001).

The models presented to date all consider homogeneous distributions of the infected agents. They consider that any individual has the same probability of coming into contact with the infection. This is generally true when the system under study is mixed (i.e. infection occurring in circulating blood, or agitated culture systems) and when the information regarding the spatial structure of the system is not available. More complicated spatially explicit or network-based models can be developed to account for the spatial structure in the spreading of diseases (Rhodes and Anderson, 1996).

Alternatively, models can be built using bottom-up strategies: Cellular Automata (CA) and Individual-based models (IbMs). CA consider the spatial structure of the system alone and define local rules for the interaction between sites. CA have been applied to study epidemics, giving similar results to the ones obtained with the SIRS models and extending their versatility to other situations (Ahmed and Agiza, 1998).

The connection between PbMs and IbMs in the modeling of infectious diseases has recently been analyzed in depth (Sharkey, 2008). More particular, IbMs have been developed to better understand the between-host spreading of malaria (Dietz, 1988), the intra-host population dynamics of *P. falciparum* (White et al., 1992), and both processes

simultaneously (McKenzie and Bossert, 2005).

Chapters 2 and 3 in the current study are particular applications of the IbM approach to the study of the infection processes occurring in *in vitro* cultures of *P. falciparum* infected RBCs.

1.5 Aim, objectives and outline of the thesis

The motivation for the research carried out by MOSIMBIO is to extract general understanding on microbial populations through the study of different specific applications. Therefore, we address particular aspects regarding the role of the population structure in communities consisting of a large number of cells, usually at the request of field experts. As a result, we develop practical tools that can be used to connect theoretical microbiology with social and industrial demands.

The research in malaria is issued upon request of the EMG-GSK and gives rise to the development of the tool INDISIM-RBC. The aim of this thesis is to acquire better understanding of the long-term *in vitro* cultivation of human RBCs infected with *Plasmodium falciparum* under different conditions.

Alongside this particular application, the thesis also intends to revise, compile and formalize the current know-how regarding the modeling of microbial populations in general, in order to improve current strategies and techniques for studying microbial ecosystems and cell cultures.

The specific goals to be achieved can be classified into three categories:

A) conceptual goals, which aim to improve the comprehension of the culture systems, decode the underlying mechanisms and explain the observed behaviors,

B) methodological goals, which refer to those advances regarding the procedures and techniques that need to be achieved when raising, building and analyzing the model; and

C) applications and practical benefits that can be obtained from the use of the methodology, the model and the simulator INDISIM-RBC.

A) Theoretical approach

The increase in understanding of complex systems requires building mechanistically rich models that capture the essential features for explaining the observed behaviors. Such models should be appropriate to:

A1) Key the experimental measures that provide operative information relevant to the study of malaria-infected RBCs in *in vitro* cultures.

A2) Identify the individual features and local mechanisms that govern the proliferation of the malaria infection in *in vitro* cultures.

A3) Determine how macroscopic conditions and external manipulations of the culture system influence the dynamics of the infection.

B) Methodological approach

Building a mechanistically rich and (temporally and spatially) explicit IbM of the experimental cultures that allow for a quantitative comparison of the simulation outcomes with real measurements is more easily made when following a plan. A methodological roadmap to do so is:

B1) Develop minimal models and gradually increase their complexity following the PoM (Pattern-oriented Modeling) strategy (Grimm et al., 1996).

B2) Select the most appropriate approach to tackle each specific problem (either PbM and IbM), and apply the SWOT analysis (Strengths-Weaknesses-Opportunities-Threads) to modeling in microbiology.

B3) Develop tools to canvas the experimental data and the simulation outcome.

B4) Adapt standardized methods for describing, analyzing and communicating IbMs, such as the ODD (Objectives-Design concepts-Details) protocol, to modeling in microbiology.

C) Applications of INDISIM-RBC

The use of the theoretical and methodological advances should lead to effective applications that serve to improve the understanding and management of specific real systems. The expected benefits are to:

C1) Define methods and criteria to compare different RBC storage periods, parasite strains and commercial sources of culture medium, among others.

C2) Determine the connection between mechanisms operating at the cellular and the macroscopic levels of description.

C3) Detect the advantages and limitations of different culturing protocols and call the prospects for improving them.

C4) Set the geometry and dimensions of the culture vials that lead to optimal static cultivation of the parasite.

Outline of the thesis

The specific results obtained with INDISIM-RBC to cover the objectives listed above regarding the application of INDISIM-RBC appear recurrently in Chapters 2 and 3 of this thesis. The fundamentals of INDISIM methodology are presented in Section 2.1. The first robust version of INDISIM-RBC, the two-dimensional model, is presented in Section 2.2. The results obtained with it are presented in Section 2.3. Section 2.4 outlines some of the problems that could not be tackled with the version of the model in 2D. The model was presented in Ferrer et al (2007).

The three-dimensional version of the model is presented in Chapter 3. The observed patterns that could not be explained by the 2D model are described in depth in section 3.1. Section 3.2 describes the modifications made on the model to deal with these pending issues described above. Section 3.3 presents the first results from this version. The geometric limitations on the static cultivation of the parasite that result from local interactions are studied in depth in Section 3.4. The contribution of diffusion to the local limitations is examined in Section 3.5. Finally, some lines that remain open for further research are proposed in Section 3.6. The application of INDISIM-RBC to static cultures was presented in Ferrer et al (2008b).

The view obtained by MOSIMBIO when modeling different types of microorganisms and phenomena results in an heuristic program that is formalized in Chapter 4. A strategy for raising realistic and mechanistically rich models, Pattern-oriented Modeling (PoM), is outlined and adapted to microbial communities and cell cultures in Section 4.1. Its application to deal with malaria cultures is put forward as an example. Two complementary strategies can be adopted to build models of microbial communities: bottom-up and top-down approaches. The suitability of each strategy -specifically, models based on individuals (IbMs) and models describing whole populations (PbMs)- is examined in Section 4.2. The discussion is illustrated with examples from predictive microbiology and microbial ecology in general. Finally, some tools for analyzing microbial systems are presented in Section 4.3. Some strategies and techniques adopted from thermodynamics and statistical physics are applied to the analysis of populations, in order to provide criteria to deal with real or simulated cultures. This discussion was compiled in two publications:

Ferrer et al. (2008a, and 2009).

The thesis concludes with Chapter 5, which consists of a summary of the general conclusions (Section 5.1), an scheme of the specific results obtained (Section 5.2), and an exposition of the perspectives for further work (Section 5.3).

In order to complement the information here presented, some clarifications, technical details and additional data have been included in the Appendices. A list of the abbreviations employed in the document is presented in Appendix A. A detailed description of the experimental work carried out by the EMG-GSK is presented in Appendix B. An outlined description of the statistical tools and programs used is presented in Appendix C. Finally, a list of publications related to the work presented is offered in Appendix D. This appendix also includes a link to the web page of MOSIMBIO, where the interested reader can find links to some sample simulations in 2D and 3D, to the open sources of the simulation code in Visual Fortran 6, and to other lines of investigation pursued by the group.

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 by using alternative cultivation methods. *Plasmodium falciparum in vitro* cultures (see Section 1.3) have been modeled using INDISIM-RBC, an Individual-based model (see Section 1.4) that has been framed 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 of the results obtained 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 modeling 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 they are 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 pseudo-random sequences. These 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 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 of 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 further extended. Although INDISIM originally arose from MC and MD models, it has its own specific features.

For this reason, the model is better seen in the context of System Dynamics, the

branch of physics that tackles 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 of the application of the IbMs approach to microbial populations is provided 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 ecological populations. Simulations were used to show how the application of fundamental theoretical principles can give rise to the behavior observed in real communities (Wagensberg et al., 1988a; Wagensberg et al., 1988b). 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 as 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 the first time and its general methodology was explained in detail (Ginovart et al., 2002a). From then on, INDISIM evolved in the study of 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 yogurt processing with *S. thermophilus* and *L. bulgaricus* (Ginovart et al., 2002b), bacterial ecosystems in soil dynamics (Ginovart et al., 2005; Gras, 2004), 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 a spatially explicit IbM. The modeled space is split into a set of regular divisions called spatial cells. The modeled events occur at finite and regular intervals, 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 of 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 modeled 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 analyzed.

The matrices $P_N(t)$ and $G_Q(t)$ are usually very large ($N \sim Q \sim 10^4 - 10^6$) 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 implemented as in-

dependent sub-models 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 with each other. Such a procedure provides the staggered study of complex systems and allows determining the appropriate degree of complexity 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 occurring 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 following 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.

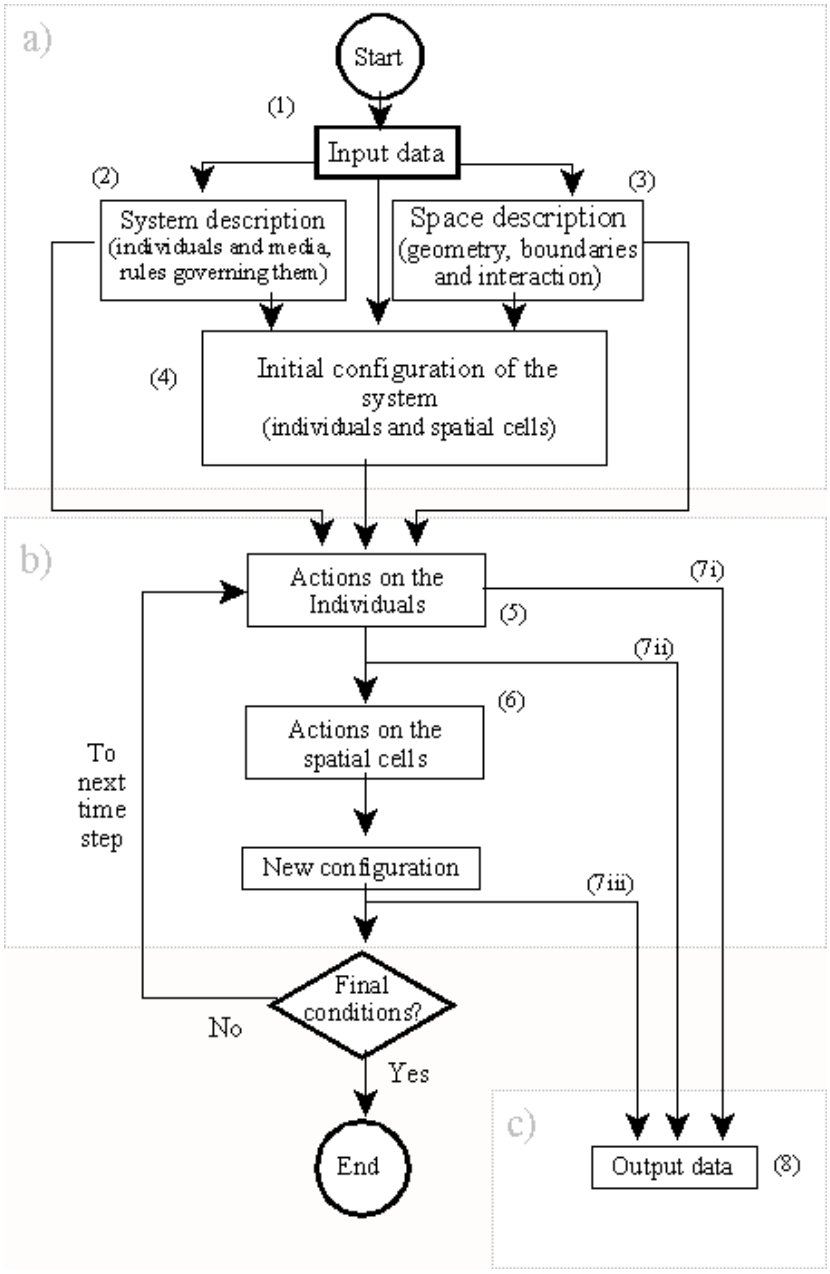

Figure 2.1: *Flow chart of the simulator INDISIM. There three main parts of the Montecarlo Method are distinguished. a) Initialization: (1) External data is input (2) to characterize the individuals and spatial cells, define the rules governing them, and (3) set the spatial structure and scaffold of the model. This information is used (4) to set the initial configuration of the system and to run the simulation (5,6). ; b) Main loop: repetitions at each time step that recurrently apply the rules governing (5) the individuals, and (6) their local environment; and c) Outcome: (7) data are obtained through the simulation (7i: during the course of individual actions, 7ii: after all the population has acted and 7iii: after the action on all spatial cells); (8) output information is stored and analyzed.*

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 comprises 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 *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 behaviors 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 particularities 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 focuses especially on how system-level behaviors 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 represents a small fraction of the hematocrit. In the first version (*2Dv.1*), the space is modeled 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 modeled 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 this 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. In other words, most of the surface of the culture system consist of thousands of replicas of the model tiling it (exceptionally, the regions next to the walls of the culturing flasks are not in contact 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 re-enter through a point symmetrical to its exit. Such boundary conditions reduce 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 a set 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. These 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. Also there are properties that depend on the cellular state, such as the infection susceptibility (P_{inf}), which depends on the age of the RBC, or the accidental death rate during the infection cycle (P_{death}), which depends 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 requisite 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 ($ICL = \{0, R, S, T\, or\, F\}$), 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 ($n_{RBC}\left(x, y\right)$), and of each of the infection stages ($n_{ICL}\left(x, y\right)$, $ICL = 0 : F$), number of merozoites ($n_{mero}\left(x, y\right)$), and total amount of glucose ($C_{gluc}\left(x, y\right)$) and lactate ($C_{lact}\left(x, y\right)$).

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: 1) processes affecting the RBCs, 2) processes affecting the spatial cells and 3) 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 the RBC is placed and modifications in the spatial cell due to RBC actions are updated as each individual action proceeds. Once a RBC has finished all its actions,

another RBC starts to act. Output data regarding any of these processes are labeled as *7i* in Figure 2.1.

RBC actions alter the characteristic variables of each spatial cell. After every RBC has acted, processes affecting the spatial cells take place. The non-homogeneities caused by the diversity in biological activity produce fluxes to relax the appearing gradients. Output data reflecting these transport phenomena are updated in paral·lel, once every spatial cell has acted, and they are labeled as *7ii* in Figure 2.1.

Finally, processes 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 of the whole system are implemented after actions on the individuals and spatial cells have occurred. They represent a modification of the scenario for the next time step. Output data showing the configuration of the system after all actions are labeled as *7iii* in Figure 2.1. The modeled processes are listed below following this scheduling:

1. Processes affecting the RBC (including IRBCs).
   (a) Motion: RBCs randomly shift positions in agitated cultures and don't move in static cultures (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 neighboring cells, so
      ii. any healthy RBC neighboring an IRBC at its lysis may become infected with a fixed probability.
   (e) Death: dead RBCs are removed from the model. Death may becaused by
      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 the infection cycle.

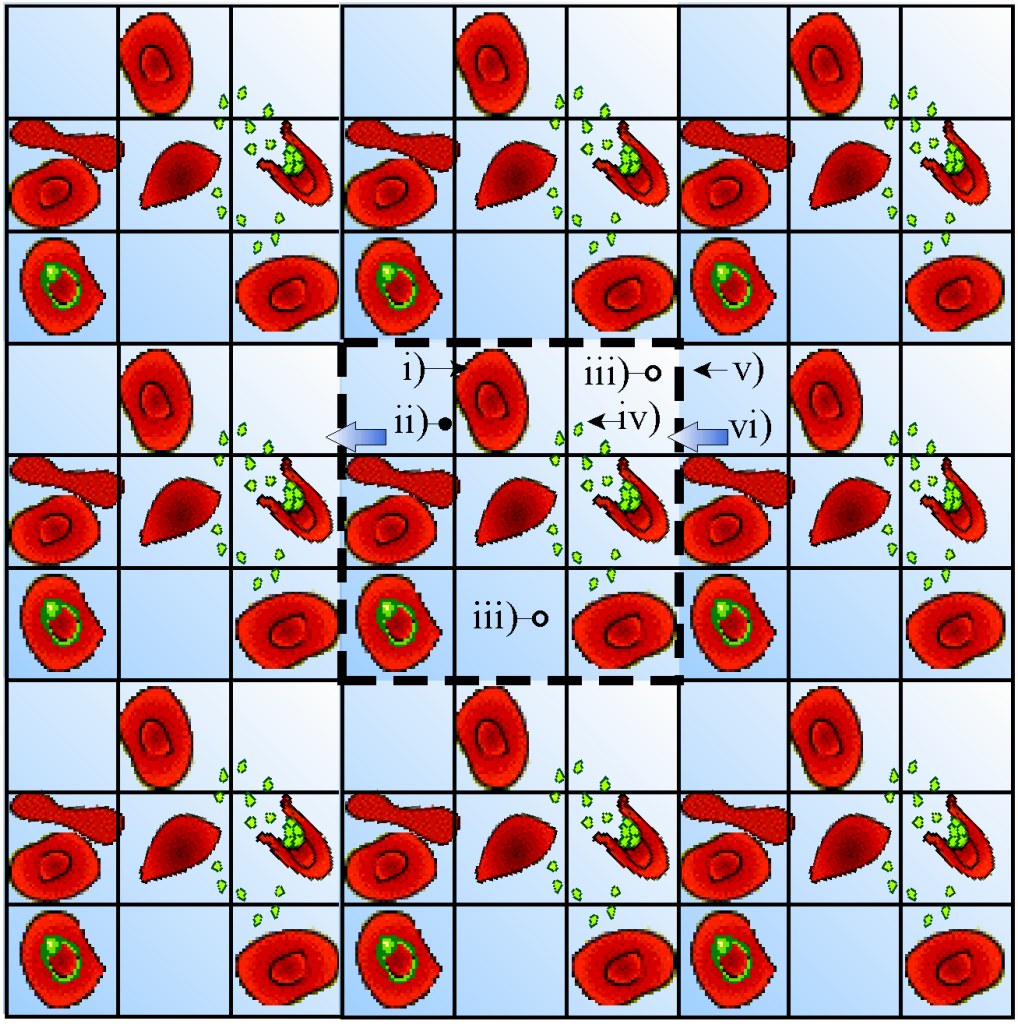

Figure 2.2: *Depiction of the entities, variables and spatial structure of the 2D versions of INDISIM-RBC. Figure shows a 3x3 square grid containing i) RBCs and ii) Spatial cells. Up to 2 RBCs can be contained in a spatial cell; iii) concentration of substrate glucose is indicated as the gradation in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

Parameter description	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,
Maximum time for infection	$\bar{t}_{MSUS}$	170 (*1700*) ts	1983
Average glucose uptake rate	$\overline{U}$	20 su / ts	Pasvol et al. 1980
Maximum metabolic stress index	IM_M	3	Guyton, 1987
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}	
- Healthy RBC	*0*	- 0	
- Ring	R	18 (*180*) ts 0	Sherman, 1998a
- 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 %	
Initial mean post-invasion time	$\bar{t}_{INF}(0)$	0 - 48 (*480*) ts	Trager and
Initial glucose per spatial cell	$\overline{C}_{gluc}(0)$	1200 su	Jensen, 1976
Initial lactate per spatial cell	$\overline{C}_{gluc}(0)$	0 su	
Subcultivation period	-	48-96 (480-960) ts	
Subcultivation ratio	-	66% - 80%	MR4, 2009

Table 2.1: *Characteristic 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 versions 2Dv.2 and v3D of the simulator that operate at reduced time steps.*

2. Processes affecting the spatial cells (SC).

   (a) Diffusion of substrate: Substrate propagates from cell to cell following 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) Resetting 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) Sub-cultivation: 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 spatial patterns emerge from the individually 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) depends on these features, and thus also emerges from the individually defined rules.

Adaptation/Fitness/Prediction/Sensing:

RBCs perceive just their immediate surroundings ($\sim 10\ \mu m$). Rules governing individuals take into account the number of glucose and lactate particles, extracellular merozoites and RBCs in the same spatial cell. Rules governing each spatial cell $(i,\ j)$ take into account its immediate neighborhood (nn) which consists of eight cells, four side-to-side spatial cells -$(i,\ j \pm 1)$ and $(i \pm 1,\ j)$- and the four diagonal neighbors $(i \pm 1,\ j \pm 1)$.
The principle of parsimony suggests that, since cellular behavior 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. Nor do RBCs have predictive capability, and they 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 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 -sub-model *1d.ii-* or accidental death -sub-model *1e-*).
It must be stressed that the inclusion of stochasticity is essential in order to obtain the appropriate emergent behaviors. 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 nutrients and accumulating harmful substances. Both interactions are local and range to the first nearest neighboring cells.

Observation:

The graphical interface of the model shows data collected at the end of each time step (label 7iii in Figure 2.1). It is comprised of 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, measurements 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. Measurements of local characteristics of the real systems were not performed during the preparation 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 screen-shot of the *2Dv.2* version of the model.

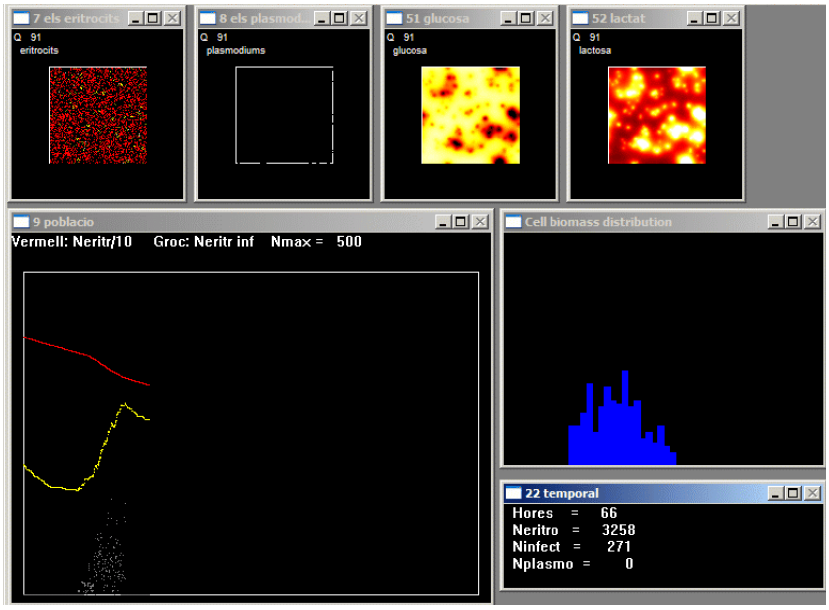

Figure 2.3: *Screen-shot of the version 2Dv.2 of the 2D model. This version operates with a time step set to six minutes. The size of the spatial grid is 100x100 cells. The initial population is 20000 RBCs and the initial parasitaemia 1%. Windows show: 1) spatial distribution of healthy RBCs (red) and IRBCs (yellow); 2) spatial distribution of merozoites (indicating where lysis occur); 3) spatial distribution of glucose concentration; 4) spatial distribution of lactate concentration; 5) temporal evolution of RBC population (red, rescaled), parasitaemia (yellow) and merozoites (grey); and 6) age structure, histogram of the post invasion times among the IRBC sub-population. The numerical display shows the time step and corresponding time for real systems and total number of RBCs, IRBCs and extracellular merozoites.*

Details

Initialization:

Cultures are set with random distributions of RBCs, infected RBCs and glucose. No lactate and no extracellular 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 truncated ($\overline{X} - 1.96\sigma_X^2 < X < \overline{X} + 1.96\sigma_X^2$) 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 behaviors. 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 modeled in accordance to the protocols proposed by Trager and Jensen (1976) and the MR4 recommendations (MR4, 2008). It consists of the removal of most of the population and replacement by healthy RBCs. Depending on the protocol, this occurs each 48 ts- 96 ts, and the fraction of removed population varies from 66% to 80%.

Medium renewal is modeled as the (partial or total) resetting of the glucose and lactate concentrations. This 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 homogenization of glucose and lactate concentrations. This occurs together with subcultivations or at fixed spans. In particular, suspended cultures are modeled as being agitated at 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 exposure 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:

Three kinds of submodels are distinguished: 1) those referring to the erythrocytes (RBC), 2) those refering to the spatial cells (SC), and those refering to the culture system as a whole (WS).

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 the first turn of action. Randomization of the sequences of action i is carried out using disordering tables that can reshuffle any given list. These tables are created at the beginning of the simulation (step 3 in Figure 2.1).

1a) RBC Motion. At the beginning of the simulation and after an agitation, RBCs are randomly distributed throughout 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 is required by its 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 greater energy demands 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).

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

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 when the RBC starves. 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 nucleus divisions 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 updated ($t_{INF} = t_{INF} + 1$) 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 origin spatial cell itself (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_{CULT}, see Submodel 3a) after the threshold of optimal susceptibility (t_{MSUS}).

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

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 origin 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 as 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 one of the nearest neighboring cells after each unsuccessful invasion attempt.

1e) RBC Death. Each RBC may die due to i) accidental death probability P_{death}, ii) exceeding its maximum age ($t_{RBC} > t_{MRBC}$), or iii) not fulfilling its metabolic demands ($IM > IM_{MAX}$). IRBCs, in particular, die at the end of their infection cycle (iv, when $t_{INF} > t_{MINF}$). Whenever an 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.

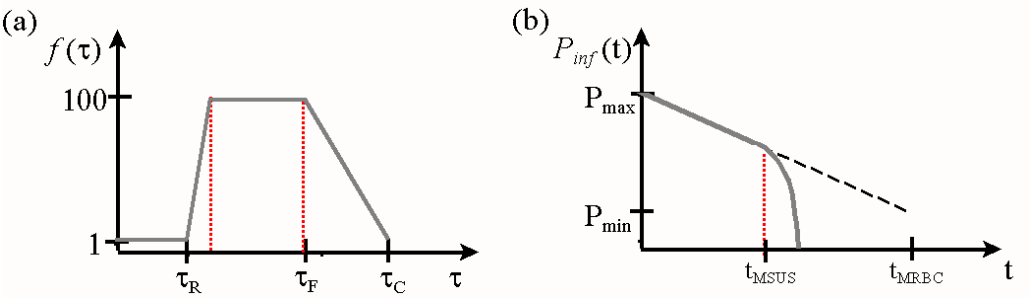

Figure 2.4: *a) Factor of individual metabolic requirements for infected RBCs ($f(\tau)$ as a function of the post-invasion time (t_{INF}). $\tau_R = t_{MCL}(R)$: end of the ring stage of the infection cycle. $\tau_F = t_{MCL}(R) + t_{MCL}(T)$: beginning of the fragmenter stage of the infection cycle. $\tau_C = t_{MCL}(R) + t_{MCL}(T) + t_{MCL}(S) + t_{MCL}(F)$: 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.*

2a) SC Diffusion. The 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 further compensated through diffusion, according to Fick's law. For each spatial cell and at each time step, the model evaluates the difference in concentration with the neighboring 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 immediate 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 $w_{k,l}$ weighs the contribution of each spatial cell, which is inversely proportional to the distances between 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 at $\widetilde{D} = 0.6$, 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 the version of the diffusion model here presented should be improved. This problem is further 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 modeled beyond merozoite egress at the end of an infection cycle. In contrast, in version *2Dv.*2 merozoites can move to the nearest neighbor spatial cells during the next five time steps after egress. At the end of each time step, merozoites remaining in the medium randomly shift their position to one of the nearest neighbor cells 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 pre-fixed periods or be triggered whenever a pre-fixed 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 ($N_{input} = N_{RBC}(t = 0) - N_{RBC}(t = t*)$, where $t*$is the time at subcultivation), 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 at zero ($t_{CULT} = 0$). Therefore, the time in culture of an 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 on the day after a subcultivation, and it entails the complete renewal of medium. Continuous 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}.

2.3 Analysis of experimental observations with the 2D model

2.3.1 Short-term preservation of IRBCs. Adjustment of INDISIM-RBC.

The reproduction of the first short-term culture of malaria infected RBCs *in vitro* was used to check the relevant parameters to set the calibration of the program. The reference employed is the experiment of Pavanand et al. (1974) for preserving *P. falciparum* in culture for more than a complete cycle of infection. Infected RBCs are maintained in static cultures for 64 hours with effective parasite proliferation. Cultures of 0.2 ml are incubated at 37ºC in a mixture of culturing medium and human sera, with no medium renewal. Concentration of the hematocrit is 9.36% in volume. Size of the culturing aliquots and shape of the hematocrit layer is not specified. The initial glucose concentration is 2g/l and initial lactate concentration is 0 (Diggs et al., 1971). Initial parasitaemia is approximately 1%. The experimental measurements are parasite counts per 1000 RBCs. Three forms of the IRBC are distinguished by their infection stages: rings, trophozoites and schizonts. No distinction is made between schizonts and fragmenters.

The experimental data regarding the parasite counts of each of the infection stages $(0, R, T \; and \; S + F)$ are used to build and calibrate the model. The initial parasitaemia $\%I_0$ can be inferred from the experimental data assuming no accidental death rates for the healthy RBCs and Ring stage forms. The values for the probability of accidental death per time step and infection stage $P_{death}(0, R, T, S \; and \; F)$, the maximum susceptibility to invasion of healthy RBCs P_{max} and the initial post-invasion time distribution for IRBCs

are set to adjust experimental data (post-invasion times are uniformly distributed from 1 to a maximum initial post-invasion time).

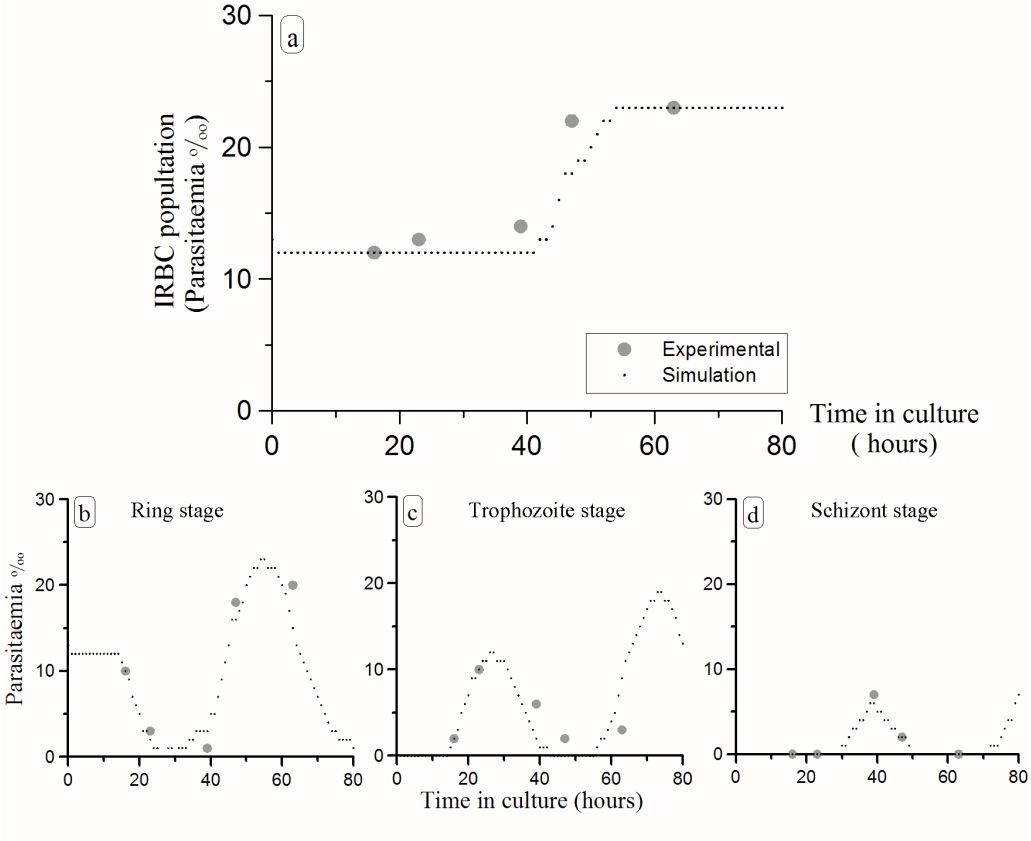

Figure 2.5: *Reproduction of the experimental results for short -term in vitro preservation of P. falciparum* (Pavanand et al., 1974) *with the version 2Dv.1*. Grey dots represent experimental results. *Small black dots represent simulation results. The statistical variance is not shown. a) Total parasite count. b) Ring-stage IRBCs. c) Trophozoite-stage IRBCs. d) IRBCs at schizont and fragmenter stage.*

Simulation results obtained with version *2Dv.1* are shown in Figure 2.5. Results obtained with version *2Dv.2* with the same input parameter sets (with the values of P_{death} and $P_{prop} = 0.05$ readjusted in accordance with the modification of the time step) do not show significant differences.

The variance of the model is assessed performing 100 repetitions of the simulations with a fixed set of parameters and varying the random seed. It has been found that,

once every parameter is fixed, simulations are highly determined, and the variance is $\sigma^2 < 10^{-3}$.

The comparison of the model outcome with experimental measurements is usually carried out qualitatively because the use of statistical tests can be deceptive and must be handled with care. For instance, a Student's t-test suggests that the two data sets are not significantly different ($p - value = 0.08$), but the Kolmogorov-Smirnov test gives only a non-significant probability that both samples come from the same distribution function ($r^2 = 0.7143$). For an outlined description of the statistical tools appropriate to comparing each set of samples see Appendix C.

Input parameter	Value			
$\%I_0$	0.12%			
$P_{death}(-)$	0	R	T	$S+F$
	0	0	0.005	0.025
P_{max}, P_{min}	0.8 , 0			
$t_{INF}^{MAX}(t=0)$	4 ts			

Table 2.3: *Values for some initialization parameters in the 2Dv.1 simulations of the experiments for short -term in vitro preservation of P. falciparum* (Pavanand et al., 1974). *Measured initial parasitaemia ($\%I_0$), and values set to best fit experimental outcome: probability of accidental death per time step and infection stage ($P_{death}(0, R, T, S \text{ and } F)$), maximum susceptibility to invasion of healthy RBCs (P_{max}) and maximum value of the post-invasion time of the inoculated IRBCs ($t_{INF}^{MAX}(t=0)$).*

The values for the input parameters have been searched through the systematic and oriented trial of different sets of values (Prats, 2008). The values that best fit the experimental observations are outlined in Table 2.3. The sensitivity of the model to slight modifications on the values of the parameters has been qualitatively assessed. For instance:

1. the increase in parasitaemia (or growth ratio GR) is more affected by variations in the accidental mortality rate P_{death} than by those in the infection probability P_{max}. In fact, varying P_{max} from 0.7 to 1 does not significantly alter the simulation outcomes, while small modifications of $P_{death}(S + F)$ from 0.25 to 0.255 make a substantial difference.

2. The initial distribution of post-invasion times is determinant to the best fit to the experimental data on the parasite counts of the different stages of IRBCs. The best-fit function is a uniform distribution of post-invasion times truncated for post invasion times greater than $t_{INF}^{MAX}(t=0)$.

The analysis of sensitivity of input parameters is not exhaustive. The model outcome could be more accurately fitted to the experimental observations, yet the current qualitative fitting of the model indicates that the modeled infection dynamics and the structure of the post-invasion times among IRBCs are consistent with the observations of real cultures.

A thorough assessment of the sensitivity of parameters is important for well-established models, in order to improve their predictive capability. Models with fewer parameters and a simpler structure are more suitable for such a use. A discussion of the role of different kinds of models is put forward in Section 4.2.

2.3.2 Long-term cultivation of IRBCs. Readjustment of INDISIM-IRBC.

The experiment carried out by Trager and Jensen (1976) is used to validate the qualitative behavior of the model when simulating cultures with multiple cycles of infection. Healthy and infected RBCs are maintained in static cultures for several weeks with effective parasite proliferation. Cultures of 1 ml with 5% hematocrit are incubated at 38^{o}C under a low-oxygen atmosphere and culturing conditions referred to in Table 1.2, in flat bottomed vials. The depth of the hematocrit layer (HLD) is set to a value $HLD \sim 0.1\ mm$. Initial concentrations of glucose and lactate are 2.667 mM and 0 mM, respectively.

Subcultivations are performed every three our four days, depending on the parasitaemia. Variable dilutions are effected to obtain a parasitaemia around 0.5% after subculture. In particular, the dilution factor of the sample to the subculture is 3-fold, 3.5-fold or 4-fold, this last being the most usual. Some of the subcultures are indicated in the experiment, while others are not specified but may be inferred from the indicated dilution shown in the experimental data sets. Four forms of the asexual intraerythrocytic stages of the parasite are distinguished in this experiment ($0, R, T, S\ and\ F$).

Two different protocols are presented: the *continuous method* and the *candle-jar method*. In the first one, IRBCs are maintained for 48 days out from an inoculum with $\%I_0 = 0.32$ in a semi-open system with a continuous flow of culture medium. Medium renewal is performed continuously at a rate of 50 ml/day out of a reservoir of 120 ml. The second protocol achieves the preservation of parasitized red blood cells for more than three months in a closed culture with discrete daily medium renewal. The medium renewal is carefully executed, so no disturbance of the RBC population is expected and no agitation needs to be considered due to medium renewal.

Both protocols use the same blood source, parasite strain and culture medium prepa-

ration, and maintain the hematocrit under similar geometrical conditions. Two culture lines (lines A and B) are cultured under different atmospheric conditions, B being apparently more suitable for RBC development (Jensen, 1988). At variable time spans, (1-day to 5-day) the culture is sampled and the parasitaemia and age structure are measured.

Both lines are simulated with the same model. The experimental parasite counts of each of the infection stages ($0, R, T, S$ *and* F) are used to check the capacity of the model of simulating long-term culture systems. The maximum susceptibility to invasion of healthy RBCs P_{max} is common for both trial lines, as they share parasite strain and RBC source. Their values are set to best adjust experimental data. The initial parasitaemia $\%I_0$ and the input parameters describing the external manipulations on the system are set according to the specifications of each experimental procedure. The initial post-invasion time distribution for IRBCs (mean value $\bar{t}_{INF}(t=0)$) and the values for the probability of accidental death per time step and infection stage $P_{death}(0, R, T, S \text{ and } F)$ are set independently for each trial to best fit the experimental data. Every value is determined using the procedure explained in Section 2.3.1. They are presented in Table 2.4.

Of note, the inoculum of line B is much more synchronized than the one of line A (the initial distribution of post-invasion times is narrower in line B). This could be explained by a hypothetical rupture of mature parasite forms during the transfer from one culture system to the other.

The only difference between the models of line A and B is that the former has a continuous replacement of part of its culturing medium. This is implemented as the replacement of a fraction of the culture medium at each time step ($f = \frac{50\ ml/day}{120\ ml} \cdot \frac{1\ day}{24\ ts}$). This replacement entails the removal of extracellular merozoites. Thus, each merozoite has a probability of 0.017 of being removed from the medium.

Input parameter	Line A		Line B	
$\%I_0$	0.32%		0.14%	
$\bar{t}_{INF}(t=0)$	22 ts		2 ts	
$P_{death}(-)$	S	F	S	F
	0.006	0.01	0.005	0.012
P_{max}, P_{min}	0.85, 0			

Table 2.4: *Values for some initialization parameters in the simulations of the experiments for long -term in vitro cultivation of P. falciparum* (Trager and Jensen, 1976). *Measured initial parasitaemias ($\%I_0$) and best fit values for mean value of the post-invasion time of the inoculated IRBCs ($\bar{t}_{INF}(t=0)$) and probability of accidental death per time step and infection stage ($P_{death}(S \text{ and } F)$, for healthy RBCs , Rings and Trophozoites, $P_{death} = 0$) for each line of the culture. Maximum and minimum susceptibility to invasion of healthy RBCs (P_{max}, P_{min}) are common for both lines.*

The analysis of the variability and sensitivity of the model provides similar results to those presented in the last section. Two differences with respect to the previous conclusions are:

1. In this case, the post-invasion time distributions among inoculated IRBCs that best fit the initial measurements of the age structures are normal distributions around a mean value $\bar{t}_{INF}(t = 0)$ with standard deviation $\sigma_{tINF} = 0.1$. The model shows very dissimilar results for different input distributions of the post-invasion times, when multiple consecutive cycles are taken into account.

2. The model is highly sensitive to variations in the parameters that regulate the culturing protocols. Small variations in the subcultivation span or dilution factor may lead to significantly different courses of the infection. The experimental subculturing protocols are not described with enough detail. In fact, the moment of external manipulations are not specified and must be deduced from the results.

An increased number of parameters (the initialization parameters presented in Table 2.4 plus the input parameters that describe the external manipulations of the system) now govern the dynamics of the model in a complex and interrelated way. These parameters represent magnitudes or events that can not be directly measured, and must be fixed to better adjust the evolution of the infection and of each of the IRBC stages.

The degree of adjustment of the model outcome to the experimental data is assessed by:

1. qualitatively checking the similarity of the temporal evolution of the IRBC population structure (see Figures 2.6 and 2.7, total IRBC evolution and for each of the infection stages); and

2. comparing the average infection growth ratios between two consecutive samplings $< GR_{sample} >$ throughout the culture (see Table 2.5).

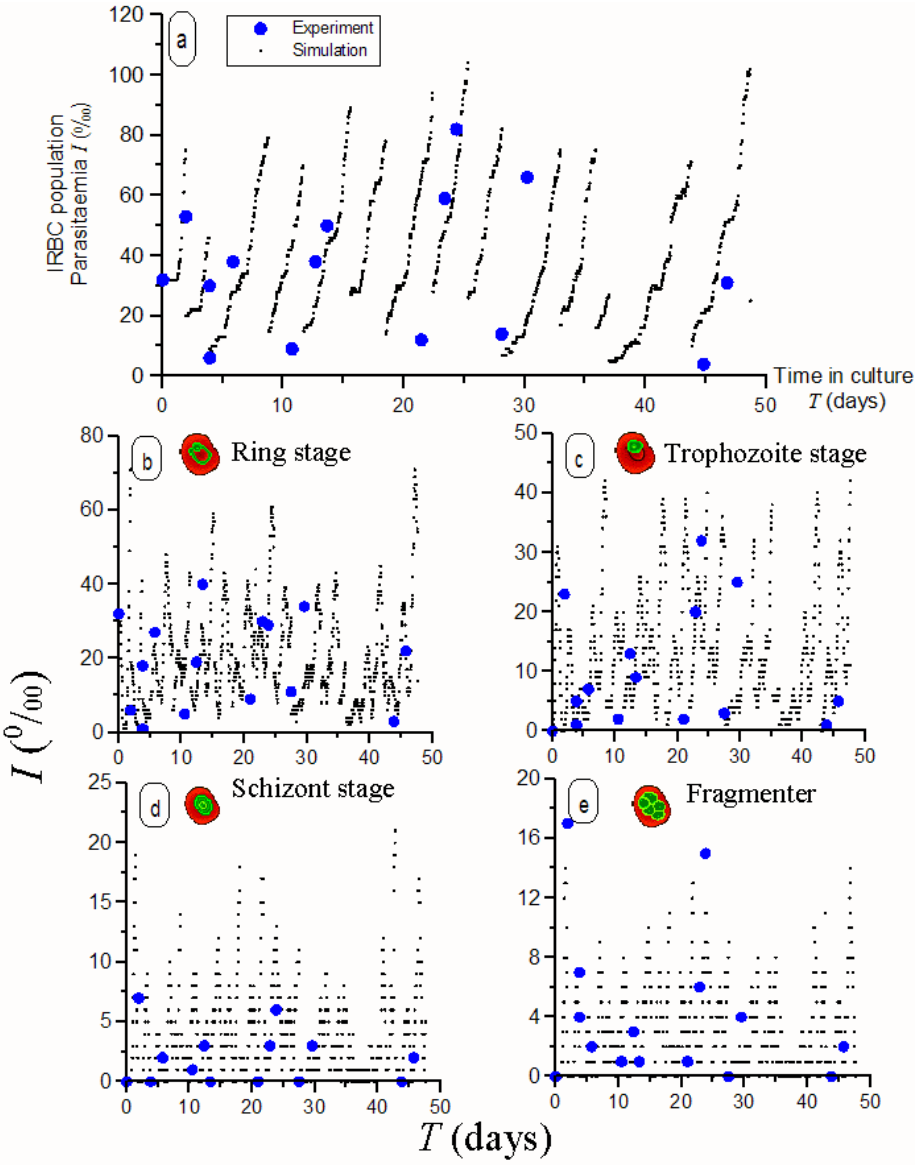

Figure 2.6: *Simulation results compared to experimental data of the infection course in the Continuous-flow cultivation (Line A) of P. falciparum IRBCs (Jensen and Trager, 1978). Blue dots stand for experimental measurements. Small black dots represent simulation results. Leap in the black dot trail indicates a subcultivation. a) Total parasite count. b) Ring-stage parasite count. c) Trophozoite-stage parasite count. d) Schizont-stage parasite count. e) Fragmenter-stage parasite count.*

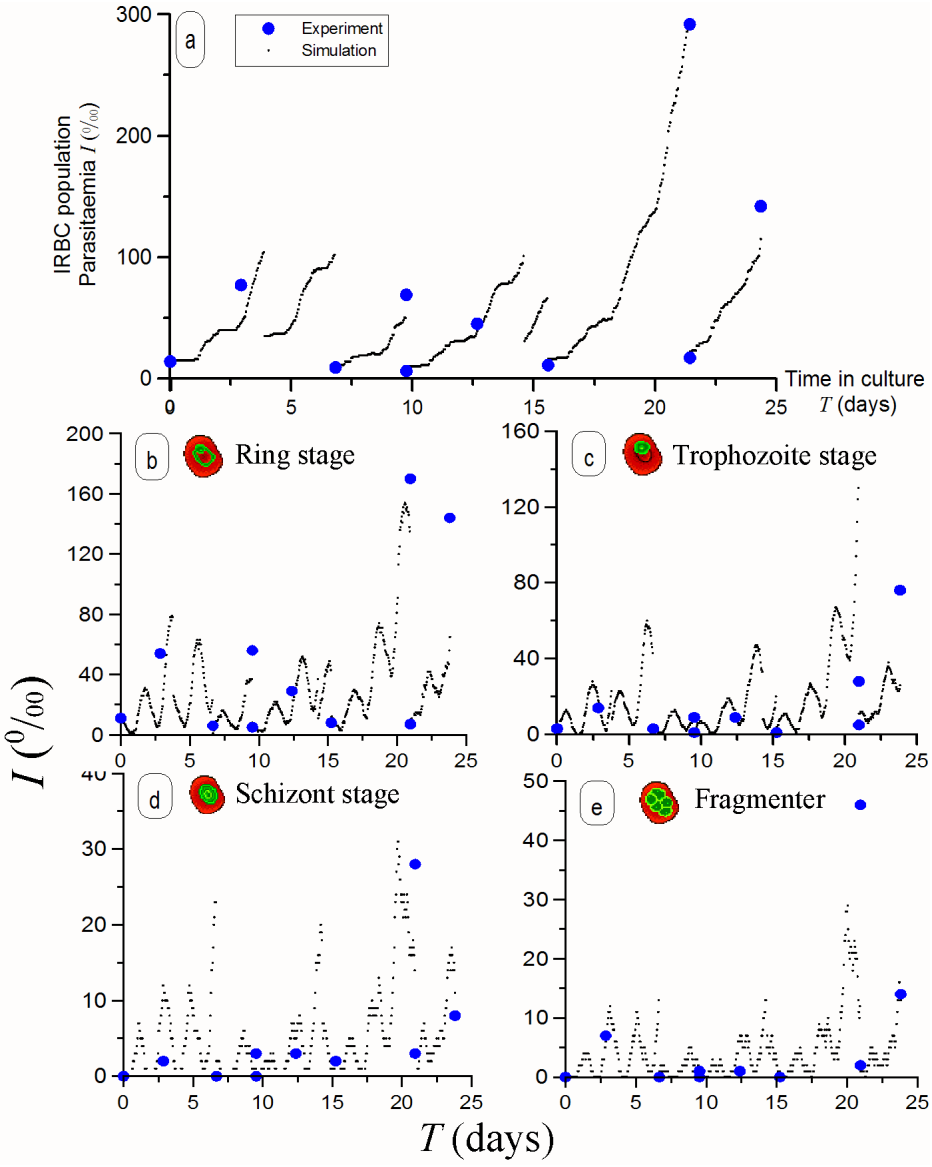

Figure 2.7: *Simulation results compared to experimental data of the infection course in the Candle-jar cultivation (Line B) of P. falciparum IRBCs (Jensen and Trager, 1978). Blue dots stand for experimental measurements. Black small dots represent simulation results. Leap in the black dot trail indicates a subcultivation. a) Total parasite count. b) Ring-stage parasite count. c) Trophozoite-stage parasite count. d) Schizont-stage parasite count. e) Fragmenter-stage parasite count.*

$< GR_{sample} >$	Simulation Outcome	Experimental results
Line A	7 ± 3	8 ± 5
Line B	12 ± 6	10 ± 4

Table 2.5: *Temporal average growth ratio of the parasitaemia during consecutive samplings* $< GR_{sample} >$*. They stand for approximately two complete intraerythrocytic cycles of the parasite. Simulation results are consistent with experimental measurements.*

Simulation results obtained with version *2Dv.1* are shown in Figure 2.6. Results obtained with an equivalent model in version *2Dv.2* do not show significant differences. The outcome of the model when performing several repetitions of the simulations varying the random seed show highly dissimilar long-term evolutions of the IRBCs population structure, but no significant differences in $< GR_{sample} >$.

The comparison between simulated and experimental growth ratios allows the conclusion that the model is consistent with the system it represents. However, it must be stressed that the uncertainty of GR_{sample} is very high both in experiments and simulations. Again, a more accurate quantitative fit of the simulation outcome to the experimental data would be very toilsome and would not provide much useful information. Additionally, the qualitative reproduction of the experimental dynamics with the simplest model assumptions provides a certain validation of the model: minimally, the model and the experimental cultures behave in a similar way.

2.3.3 Patterns observed in *in vitro* cultures.

The accurate quantitative reproduction of several particular data sets from single experimental trials positively allows us to extract some conclusions. First, it indicates that the model includes enough information at an individual level to reproduce the complex observed behavior. Second, it suggests that the rules in the model are a good representation of the processes occurring in the real culture systems, because the same rules serve to simulate different cultures. And third, it provides an estimation of the values (or at least order of magnitude) of some magnitudes that can not easily be measured (such as the RBC death rate and the infection probability).

However, it does not provide enough arguments to assess the validity of the model on its own. First, because it can't be disproved that simpler models could also reproduce the observed behavior. Second, because the coincidence between the model outcome and experimental observations does not ensure that it accounts for *real* underlying mechanisms. And third, because obtaining the best fit requires adjusting the values of several parameters simultaneously, and different sets of parameters could provide similar outcomes. The

exhaustive exploration of the parameter space would be extremely time-consuming and has been carried out just to a reasonable extent here.

So, what else can be done to assess the validity of the model? The reproduction of a group of independent patterns observed in real culture systems at different scales and regarding distinct phenomena can provide increasing confidence in the capacity of the model to capture the essence of the experiments. If the model reproduces different uncorrelated phenomena and features of the real systems, it means that it correctly represents the reality. Then, the modeled rules can be better regarded as indicators of the essential underlying processes occurring in real cultures. Some independent patterns are reproduced below. A more detailed discussion of the arguments presented in this paragraph may be found in Section 4.1.

Strains of *P. falciparum*

The performance of *in vitro* cultivation of *P. falciparum* strongly depends on the strain, ranging from the infeasibility of setting cultures to highly effective harvests (Chin and Collins, 1980). The variability in the performance of a fixed strain is also very high and the evolution of observed cultures shows that the multiplication ratio of the number of IRBCs per infection cycle does not solely depend on the strain.

The results obtained by Butcher (1982) reflect this intrinsic variability. Three different strains of *P. falciparum* were maintained under static *(candle-jar)* and suspension conditions. The parasite was maintained in continuous cultures with the same blood source and conditions for approximately one month. Renewal of medium was performed daily but was skipped sometimes between two consecutive subcultivations. Samples were extracted when setting new subcultures every three days. The growth ratio (GR_{72}) was then measured. Growth ratios where calculated as the ratio between the observed parasitaemias at the beginning and end of each subculture. The experimental results are presented in Table 2.6.

GR_{72}	Strain A		Strain B		Strain C	
trial	Static	Susp.	Static	Susp.	Static	Susp.
1	6.9 *	1.8 *	1.8 *	14.4 *	5 *	23.8 *
2	10.6 *	2.0 *	1	4.5	0.7 *	9.0 *
3	8.8 *	8.2 *	0	8.0	4.4 *	10.8 *
4	9.1	11.6	7.1	22.0	6.8 *	9.2 *
5	7.1	6.0	1.8 *	2.9 *	7.7 *	19.7 *
6	10.6	16.5	9	25.3	-	-
7	5.1 *	3.3 *	-	-	-	-
$< x > \pm\sigma$	8 ± 2	7 ± 5	3 ± 4	12 ± 9	5 ± 3	14 ± 7

Table 2.6: *Experimental average* 3-day growth ratios (GR_{72}) *for Static (candle-jar culture conditions) and continuously agitated cultures ("Susp." stands for RBCs in suspension) for 3 different strains of P. falciparum. The row* $< x > \pm\sigma$ *indicates the mean and standard deviation of* GR_{72} *averaged over a month. The symbol * indicates that no medium renewal was carried out between two samples. Reprinted from (Butcher, 1982).*

The first remark on the experimental observations is that the intrinsic variability of any real culture system usually allows drawing only tentative conclusions regarding the observed behaviors. The averaged values are often accompanied by such high uncertainties that they can not be said to be significantly different. Using statistical tool packs such as ANOVA (a collection of statistical models to perform the ANalysis Of VAriance of datasets, i.e.: Student's t-test, Kruskal-Wallis test, Kolmogorov-Smirnov test, among others; see Appendix C) facilitates the discrimination of data, but it is still not enough to distinguish different strains.

Even discriminating between different culturing protocols is controversial. For instance, results from Strains B and C suggest that suspended cultures are more efficient than static cultures but this is not supported by the observations of Strain A. Furthermore, results from Strains A and B, which combine subcultures with and without medium renewal, suggest that the lack of medium replacement hinders the development of the parasite. However, Strain C, whose medium was not renewed between subcultures shows a qualitatively increased performance.

The simulations of the different set of experiments carried out using INDISIM show similar statistical behaviors and their analysis presents analogous problems. One strategy to follow in order to analyze the model is to focus on the effect of one single parameter, while the rest set of parameters are fixed.

For instance, modifying the strain can be associated with varying the infection capability of the merozoites P_{inf}. Variations in those parameters used to set P_{inf}, within the ranges $0.7 < P_{max} < 1$, and $0 < P_{min} < P_{max}$ entail fine variations of the simulated

outcome growth ratios (GR_{48}). Of note, this small modification of GR_{48} can eventually lead to great dissimilarities between cultures under similar harvesting protocols, because the effect of consecutive cycles is a non-linear feedback, and also because the external manipulations of the system are triggered at fixed moments and can affect systems with slightly different configurations in very unlike manners.

An additional remark specifically regarding the effect of varying the probability of invasion in the simulations points to a different topic. During the exploration of the different values of P_{inf}, it has been found that the behavior of the model is consistent with the description provided by population-based epidemiological models. A minimum threshold for the value of the parameter P_{inf} discriminates between the infections that extinguish and those that affect the whole system. In this sense, this minimum threshold is analogous to the reproduction factor (R_0) employed in continuous models. More specifically, P_{inf} plays the same role as the infection rate (β) in the SIR model, therefore $R_0 \propto P_{inf}$ (see Section 1.4.3).

Source of RBCs

Several studies indicate that *P. falciparum* has a preference for infecting young RBCs. Firstly, age fractionation of RBC populations prior to cultivations shows that the growth ratios of the infection are significantly greater for those cultures performed with samples containing more retyculocites (young RBCs aged < 7 days) (Pasvol et al., 1980). Secondly, studies of multiple infections confirm that the susceptibility to invasion of retyculocytes is 4-fold the susceptibility of older RBCs (Simpson et al., 1999).

Human RBCs for malaria cultures are usually obtained from outdated blood that is considered not to be usable for blood transfusions. This blood can be kept frozen (at temperatures that may reach -80^{o}C) for several years without lost of harvest viability. Cultures with blood frozen for 19 months and 24 months do not show significant differences when compared with control cultures from fresh blood (Pavanand et al., 1974).

In addition, whatever the source and manner of storage of RBCs, they must be prepared prior to the cultivation. They are washed and mixed with culture medium and usually subsequently stored at temperatures that range from 1^{o}C to 4^{o}C, prior to the cultivation. The time spent in these conditions affects the culture's performance. The blood storage regimes have influence on the RBC suitability when storage exceeds 3 or 4 weeks (Capps and Jensen, 1983). Cells can be used for up to 5 weeks of storage at 4^{o}C and are typically maintained for 4 days to a week once washed (Trager, 1994).

To sum up, storage spans greater than the age of retyculocytes have no significant effect on RBC susceptibility to infection, but the performance of the infection still depends on

this storage. Besides, significant differences in the parasite proliferation are observed between cultures of young and old RBCs.

The hypotheses assumed in the model are: i) the parasite has a preference for young cells, and also detects their freshness, and; ii) storage ages RBCs, reducing their freshness but without changing them. Presumably, low temperatures prevent the retyculocytes from turning into mature RBC forms. These assumptions are introduced in the model of the temporal evolution of P_{inf} in two ways:

1. as the abrupt decay in P_{inf} after a period in culture t_c that stands for the differences in the infection susceptibility observed for retyculocytes and older RBCs. This decay is introduced as an exponential factor modulating P_{inf}, with both the offset value and time constant equal to the 7-10 days (maximum age of retyculocytes). See Equation 2.4. Note: $t_{CULT} = t_{RBC} - t^*_{RBC}$, where t^*_{RBC} is the age at subcultivation; t_{CULT} does not include the storage time.

2. as a linear decay of the infection susceptibility with the RBC age (t_{RBC}, including storage time), from a maximum value P_{max} at age 0, until reaching a minimum value P_{min} at $t_{RBC} = t_{MRBC} = 120$ *days* (Guyton, 1987), which stands for the decrease in freshness of the RBC. See Equation 2.4. As mentioned above, variations in P_{max}, P_{min} and t_{MRBC} entail very small differences in the overall GR of the system.

A trial of virtual experiments to test this model consists in launching simulations similar to the *candle-jar* cultivation of the parasite where the subcultures are fed with RBCs that proceed from sources stored for different time spans. Two series of static cultures are simulated, with subcultivations that set the parasitaemia at the new culture to 0.5%, each 48 h and each 96 h, respectively. Initial parasitaemia is set to $\%I_0 = 0.5\%$ and infection probability to $P_{max} = 0.85$, $P_{min} = 0$, respectively.

Each trial explores culturing protocols where the subcultivation are carried out using increasingly old RBCs (Storage spans are set to: 0 days, 4 days, 8 days, 12 days, up to 3 months). The average parasitaemia of long-term cultures maintained for up to 3 months is compared for the different renewal protocols. The obtained results are presented in Figure 2.8.

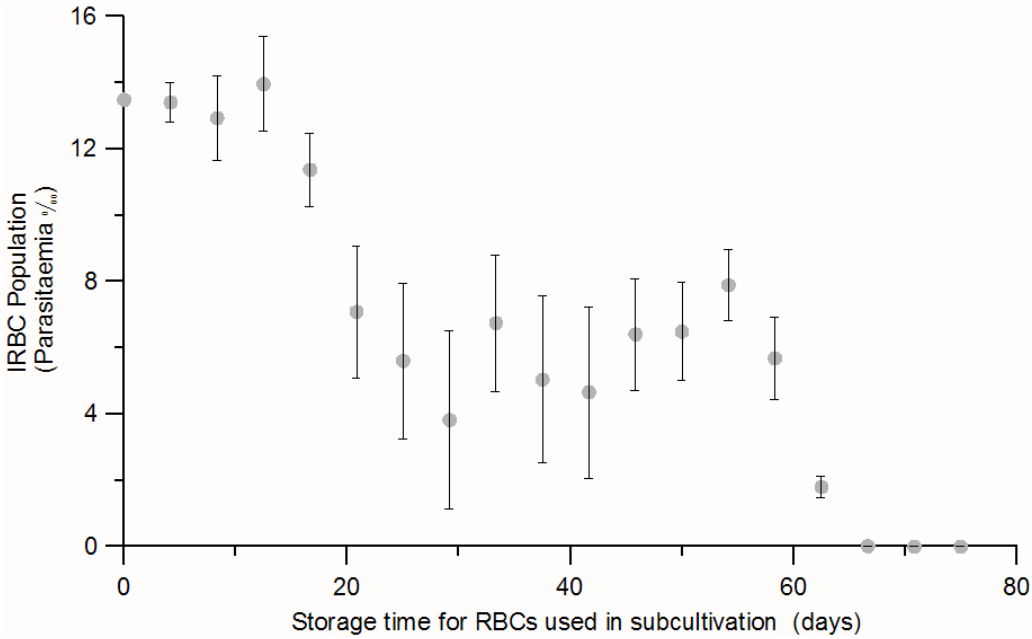

Figure 2.8: *Simulations showing the modeled effect of the storage of RBCs on the culture performance. Average parasitaemia through the culture decreases with the storage time for the source of RBCs. Grey dots show the mean values obtained from two sets of simulation trials that included subcultivations each 48 h and 96 h. Errors bars reflect the variability in the simulations.*

These results show that with storage times < 16 *days*, the parasite can grow unhindered. After 20 *days* of RBC storage, the parasite can be maintained in culture with a parasitaemia slightly greater than the inoculum (indicating that there is no effective propagation of the infection, i.e. $GR \sim 1$). If the storage time is greater than 60 *days*, the viability of cultures is severely limited.

The rules presented above allow reproducing the temporal evolutions observed in several cultures and provide an explanation for the reduced performance of the cultures in which RBCs have been stored for a long time, consistent with the experimental observations.

No virtual experiments on the multiple infection have been carried out with this version of the model because a rule limits the maximum number of multiple invasions on an RBC to 3 merozoites per RBC. This rule was introduced in response to the local limitations caused by the spatial structure of the model: up to 32 merozoites can be released from an IRBC at the end of the infection cycle, but each IRBC has 8 healthy RBC neighbors at most. Not introducing a maximum threshold for the number of multiple invasions led

to IRBCs infected with 5 or even 7 multiple invasions, which have never been observed in real systems.

Merozoite egress from IRBCs

At the end of each infection cycle merozoites are released to the extracellular medium ready to invade RBCs. The three alternative mechanisms for merozoite egress described in Section 1.2.2 are represented by three mutually exclusive rules (see Figure 2.9):

(i) Membrane fusion: all the merozoites are released into a single spatial cell, chosen from one of the nearest neighbors of the IRBC ($\sim 10\ \mu m$).

(ii) Membrane breakdown: each of the merozoites has a probability of being released into one of the nearest neighboring spatial cells including the same cell ($\sim 10\ \mu m$).

(iii) Exploding flowers: egressing merozoites are randomly distributed among the neighbor spatial cells up to a range of 2-3 spatial cells distance ($\sim 25\ \mu m$). The probability of being dumped into a particular spatial cell linearly decreases with the distance from the source IRBC. The set of probabilities is normalized.

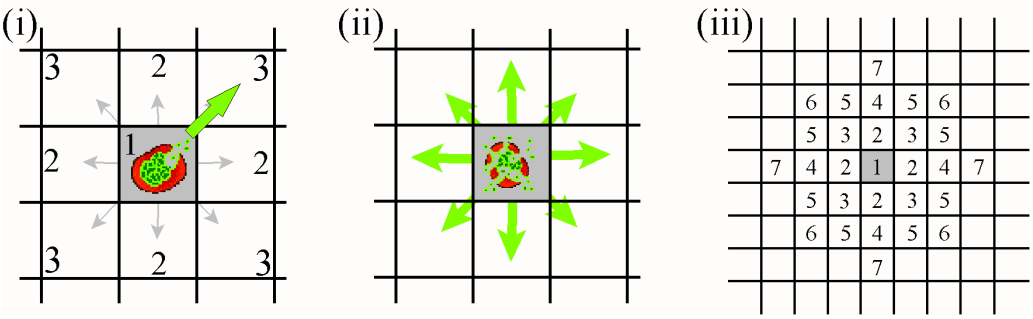

Figure 2.9: *Stencils of the models of merozoite egress from IRBCs at the end of the infection cycle (lysis occurring at the grey central spatial cell). i) membrane fusion: all merozoites are released to one of the nearest neighbor cells with a probability P(1)=0.4, P(2)=0.1 and P(3)=0.05; ii) membrane breakdown: merozoites are scattered around, each merozoite is dumped into one of the nearest neighboring cells with the same probabilities as (i); iii) exploding flowers: Long range egress. The probabilities of a merozoite being dumped into each cell are: P(1)=0.212, P(2)=0.053, P(3)=0.037, P(4)=0.026, P(5)=0.022, P(6)=0.019 and P(7)=0.018.*

Each mechanism of merozoite egress has been introduced into the model. The outcome of the 8-days simulation of a closed system was carried out assuming that the values for all

the parameters are fixed and similar to the ones used in the simulation of the *candle-jar* experiments ($\%I_0 = 0.1\%$, semi-synchronous inocula, $P_{max} = 0.8$ and $P_{death}(S) = 0.005$, $P_{death}(F) = 0.01$).

In addition, the effect of the spatial structure on the infection process was assessed for each mechanism. The scale of the spatial cell side was modified to allow up to 1, 2 3 or 4 RBCs per spatial cell. 10 different simulations of 8-day static cultures were carried out for cultures with different RBC densities and no accidental death rate ($P_{death} = 0$). The obtained results are as follows:

i) The punctual egress of merozoites leads to infections that do not propagate through the hematocrit, $GR_{24} \leq 1$. Unless $P_{death} = 0$, the parasitaemia fades out until there are no IRBCs, $GR_{24} < 1$. Increasing the RBC density (ρ_{RBC}) leads to an increase in the parasitaemia and in the infection daily growth ratio, $GR_{24} > 1$. The maximum observed growth ratios are achieved with 4 RBCs/spatial cell: $GR_{24} = 1.8$ (between day 7 and day 8 in Figure 2.9a and $GR_{48} = 2.3$ (between day 6 and day 8 in Figure 2.10a, respectively. The growth ratio per infection cycle increases with cell density and rapidly becomes saturated around $GR_{48} = 2.5$ (see Figure 2.10b.

ii) The egress of merozoites at one of the nearest neighboring spatial cells reproduces the infection propagation rates. Average simulation growth ratios range from $GR_{48} = 4.5$, when there is 1 RBC per spatial cell (which is consistent with the behavior observed in real cultures), to $GR_{48} = 10$ for 10 RBCs/sc.

The effect of the spatial structure on parasite proliferation is depicted in Figure 2.10.b. The growth ratio increases with the number of RBCs/sc but becomes saturated around $GR_{48} = 12$. This is consistent with the expected behavior in real systems if there were no limitations due to the local spreading of merozoites or the availability of RBCs, since ~ 16 merozoites on average are released in the lysis at the end of each infection cycle and the probability of invading a healthy RBC is (at the most) less than one.

iii) When merozoites are released to a fairly extensive area around the IRBC (for instance, the two first layers of first neighboring cells) the simulation outcome shows growth ratios that reach values far greater than those observed in real systems $GR_{24} = 14$ and $GR_{48} = 25$. The average value obtained with the 10 simulations with 1RBC/sc is $< GR_{48} >= 15$. This parasite multiplication ratio already exceeds the behavior observed in real cultures, so no further simulations increasing the number of RBCs per spatial cell were carried out.

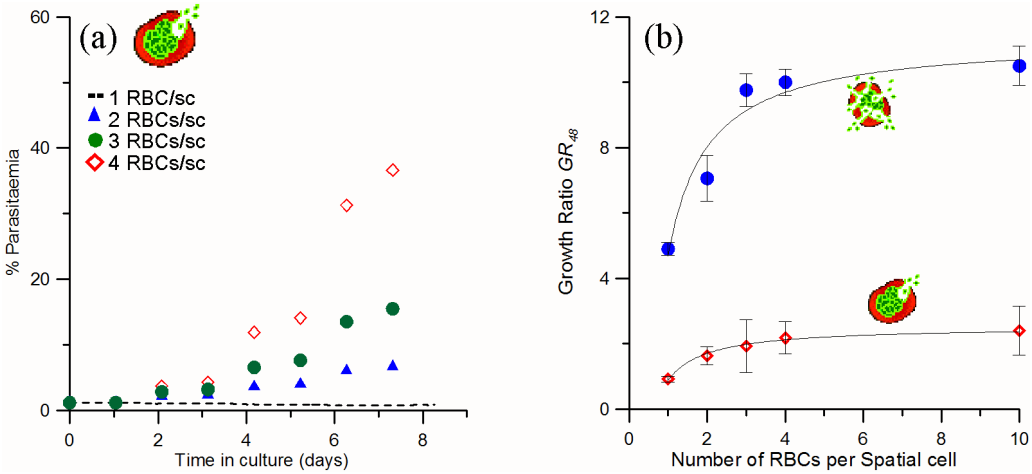

Figure 2.10: *Simulation results of the different models for merozoite egress. a) Temporal evolution of the parasitaemia for different cell densities according to the membrane fusion model (punctual egress of merozoites). Dashed line (–) represents 1RBC per spatial cell (RBC/sc); triangles (▲) represent 2RBC/sc; solid dots (●) represent 3 RBC/sc; diamonds (◇) represent 4RBC/sc. b) Comparison between the growth ratios predicted for different cell densities by two different egress models: dots (●) stand for the membrane breakdown model (scattered merozoites) and diamonds (◇) stand for membrane fusion model. The former are in greater agreement with experimental results.*

The results provided by the model suggest that membrane breakdown is the mechanism through which merozoites egress from IRBCs at the end of an infection cycle. These results are consistent with recent experimental observations (Rayner, 2006).

Serum and culture medium

The cultivation of *P. falciparum* requires the use of human serum to provide the proteins and amino acids to the parasite inside the IRBC. As it is a scarce resource, proposals for alternative serum sources (natural and synthesized) can be found in the literature (Jensen, 1988; Trager, 1994). Today, several companies provide commercial serums with different specifications.

In order to assess the appropriateness of different commercial culture mediums, the group EMG-GSK performed a series of experiments consisting of 1-week cultivation of the same strain of *P. falciparum* with the same blood source under the experimental conditions detailed in Table 1.2 and with different culture mediums and degrees of synchronism of the inoculum. Initial parasitaemia was set to $\%I_0 \sim 1\%$ and the degree of synchronism

was qualitatively assessed only distinguishing between synchronous and asynchronous inocula. The experimental results are outlined in Table 2.7. A more detailed description is presented in Appendix B.1.

Days in culture	Medium A				Medium B		Medium C			
	S.		*A.*		*S.*		*A.*		*S.*	
	$\%I_{obs}$	$\%I_{\text{calc}}$	$\%I_{obs}$	$\%I_{\text{calc}}$	$\%I_{obs}$	$\%I_{\text{calc}}$	$\%I_{obs}$	$\%I_{\text{calc}}$	$\%I_{obs}$	$\%I_{\text{calc}}$
0	1	1	1	1	1	1	0.95	0.95	1	1
1	1.3	1.3	3.3	3.3	1.3	2.8	2.3	2.3	1.9	1.9
2	7.8	7.8	$3.4_{(\frac{1}{4})}$	13.6	4.4	4.4	5.7	5.7	1.9	1.9
3	7.6	7.6	-	-	4.8	4.8	$3_{(\frac{1}{4})}$	12	$3.8_{(\frac{1}{4})}$	11.5
4	$4_{[\frac{1}{4}]}$	16	$2.9_{(\frac{1}{4})}$	47.7	$2.9_{(\frac{1}{4})}$	11.4	7.7	30.8	11.6	32.8
5	-	-	3.8	60.3	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-
7	10.2	40.8	-	-	7.9	31.8	-	-	-	-
$GR \pm \sigma$	3 ± 2		3.0 ± 1.2		2.5 ± 0.9		2.9 ± 0.7		2.7 ± 1.4	

Table 2.7: *Cultivation of P. falciparum using 3 different commercial culture mediums. S.: Synchronous inoculum; A.: Asynchronous inoculum; observed ($\%I_{obs}$) and calculated ($\%I_{calc}$) parasitaemias; GR: average observed growth rate and deviation (σ_{GR}). The subscript fraction in brackets in the observations column indicates the dilution carried out during the subcultivation prior to measurement, i.e. the fraction of replaced RBCs; - indicates no experimental measurements. Experiments performed the EMG-GSK.*

The statistical analysis of the calculated parasitaemias $\%I_{calc}$ with ANOVA (or the simpler comparison of daily growth ratios GR_{24}) shows: i) that the choice of one of these culture mediums does not significantly affect the propagation of the infection, and ii) that synchronous and asynchronous inocula behave in a similar way, both when comparing the behavior for each culture medium separately and when considering the subsets of synchronous and asynchronous trials.

INDISIM-RBC does not include information regarding the exhaustive composition of the extracellular medium, nor a detailed description of RBC metabolism. The effect of the different culture mediums is introduced into the model as variations in the values of the probability of accidental death (P_{death}). Post-invasion time distributions among the inoculated IRBCs are considered uniform, with an amplitude $t_{INF}^{MAX}(t=0)$ that decreases with the observed degree of synchronism. The initial values of $\%I_0$ are set according to the input measured values (at day 0), infection probability is set to $P_{max} = 0.85$, $P_{min} = 0$, and P_{death} is modified from culture medium to culture medium. A reproduction of three sample cultures is depicted in Figure 2.11.

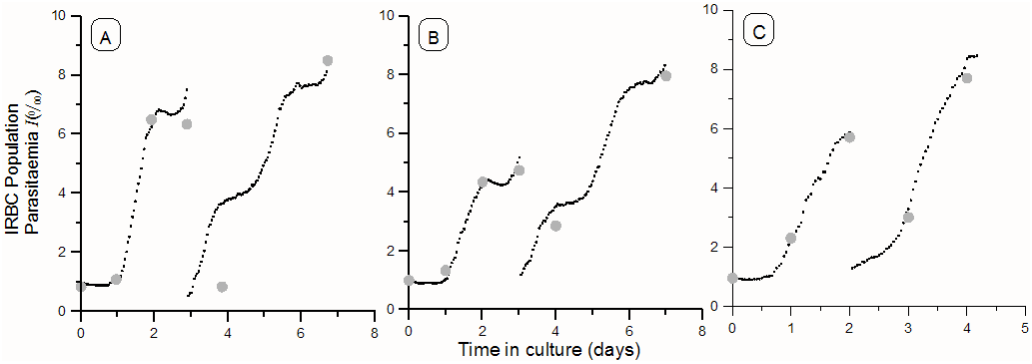

Figure 2.11: *Experimental results compared to simulation outcome for cultures using different culturing mediums. Grey dots correspond to a 7-day trial measurement. Black small dots correspond to simulation outcome. a) Culture medium A, synchronous inoculum with 100% Rings;* $P_{death}(T) = 0.001$, $P_{death}(S) = 0.007$, $P_{death}(F) = 0.01$ *and* $t_{INF}^{MAX}(t = 0) = 20\ ts$; *b) culture medium B, synchronous inoculum with 90% Rings;* $P_{death}(T) = 0.001$, $P_{death}(S) = 0.01$, $P_{death}(F) = 0.01$ *and* $t_{INF}^{MAX}(t = 0) = 25\ ts$; *c) culture medium C, asynchronous inoculum with* $< 80\%$ *Rings;* $P_{death}(T) = 0.001$, $P_{death}(S) = 0.005$, $P_{death}(F) = 0.02$ *and* $t_{INF}^{MAX}(t = 0) = 35\ ts$.

2.3.4 Synchronization of the infection process

The degree of synchronism of the inoculated IRBCs has little effect on the long-term cultivation of *P. falciparum* because, typically, *in vitro* synchronization of the parasite population disappears during the infection course (Kiatkowski, 1989). The statistical analysis of the simulation results that represent the short-term cultures in Table 2.7 show similar conclusions to those drawn from the analysis of experimental data: no significant differences can be found in the culture performance (GR) between cultures with synchronous and asynchronous inocula. Nevertheless, simulations show different behaviors depending on the degree of synchronism of the inoculum ($t_{INF}^{MAX}(t = 0)$). For instance, the lag period until the disease spreads through the hematocrit (τ_{lag}) is $\tau_{lag} = 48 - t_{INF}^{MAX}$.

Furthermore, *P. falciparum* schizogony is stably synchronized for long-term *in vitro* cultures only when specific manipulations are carried out, e.g.: sudden increase in temperature or treatment with sorbitol, among others *(MR4, 2008; Kiatkowski, 1989).* These treatments mostly affect IRBCs in mature stages of the infection cycle and leave young ring-stages (post-invasion times $t_{INF} <\sim 15\ ts$) unharmed. In particular, the procedure to synchronize parasites with sorbitol requires repeating the treatment twice in $48h$ to guarantee that only ring forms are left. To keep the parasites synchronized, the sorbitol treatment must be performed once a week.

Modeled IRBC populations naturally become desynchronized as a consequence of the variability in the duration of the infection cycles (see Figure 2.12). Synchronous IRBC populations can be obtained through the one-time removal of the mature part of the IRBC population to represent specific synchronizing treatments. In this case, a rule introduced *ad hoc* forces the system as a whole to synchronize.

Of note, stage-dependent death probabilities $P_{death}(0, R, T, S \text{ and } F)$, together with periodic subcultivation, also limit the degree of desynchronization of IRBC population (see Figures 2.12 and 2.13). In this case, the preservation of certain synchronism emerges from rules that do not directly impose the coordination of the population.

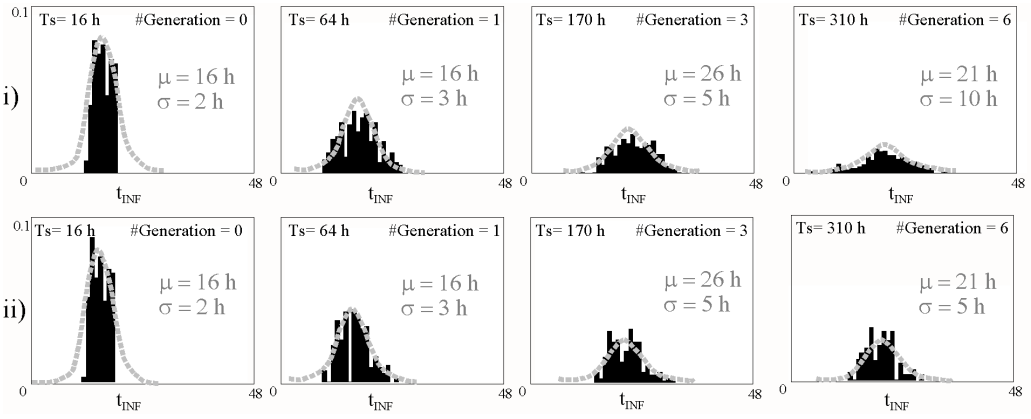

Figure 2.12: *Simulated post-invasion-time distributions among the IRBCs (outcome from model 2Dv.2). Relative frequency vs. post invasion time (t_{INF}). Ts indicates the time in culture of the sample, #Generation stands for the number of completed infection cycles in the culture. Synchronous homogeneous inoculum: $t_{INF}^{MAX}(t = 0) = 20$ ts, 2 hours. Variability in the duration of the infection cycle: $\sigma = 2$ hr. i) IRBC population gets desynchronized in cultures with $DP(\forall) = 0$ and no subcultivation; ii) synchronism is maintained by the staggered death of IRBCs ($P_{death}(F) = 0.001$) and periodic 96 hr subcultivation. Grey dashed lines* show *the best- fitting normal distributions, with mean value μ and standard deviation σ.*

Apart from synchronizing the IRBC population, the use of a staggered P_{death} applied at each time step entails the reduction of the viability of the infection. The number of IRBCs in INDISIM is analogous to $I(t)$ in the SIR model, and increasing P_{death} is analogous to an increase in the recovery/mortality rate γ (see Equation 1.2 in Section 1.4.3). This can be derived from the individual scope: the reproduction factor of the infection (R_0) is proportional to the number of IRBCs that successfully undergo a complete infection cycle ($P_{survival}$).

$$R_0 \propto P_{inf} \cdot P_{survival} \tag{2.7}$$

The probability of survival for IRBCs undergoing a complete infection cycle is the joint probability of the IRBC surviving each time step $p_{survival}(ts)$, throughout the infection cycle $t_{INF} = [1, t_{MINF}]$.

$$P_{survival} = \prod_{t_{INF}=1}^{t_{MINF}} p_{survival}(t_{INF}) \tag{2.8}$$

In turn, surviving is the complementary event of dying. Then,

$$p_{survival}(ts) = 1 - P_{death}(ts) \tag{2.9}$$

Combining 2.8 and 2.9, we find:

$$P_{survival} = \prod_{t_{INF}=1}^{t_{MINF}} [1 - P_{death}(t_{INF})] \tag{2.10}$$

$$P_{survival} = P_{death}(R)^{t_{MCL}(R)} \cdot P_{death}(T)^{t_{MCL}(T)} \cdot P_{death}(S)^{t_{MCL}(S)} \cdot P_{death}(F)^{t_{MCL}(F)}$$

The custom values of P_{death} in version *2Dv.1, 2Dv.2* and *v3D* (see Table 2.1) entail the effective survival of the IRBCs $P_{survival} = 0.9$. Particularly, the probabilities obtained using the values used to simulate the short-term preservation (see Table 2.3 in Section 2.3.1) and the long-term cultivation (see Table 2.4 Section 2.3.2) of the parasite are: $P_{survival}(short - term) = 0.6$ and $P_{survival}(long - term) = 0.9$, respectively.

A simple mathematical model describing the population structure in *P. falciparum* IRBCs was developed to understand infection dynamics *in vivo (White et al., 1992).* The White-Chapman-Watt model (WCW) considers that IRBC post-invasion times are normally distributed and uses two parameters only; the mean (μ) and the standard deviation (σ_μ). At any time (ts), the post-invasion time distribution is represented with the function $q(t_{INF})$. If the infection is fairly synchronous ($\sigma < 12\ h$), this function is composed of the underlying intrinsic distribution of parasite stages of the current generation (a normal distribution $f_{ts}(t_{INF})$) and the contribution of the previous generation ($p_{ts}(t_{INF})$ represents the left tail of the last infection cycle) and next generation of parasites ($n_{ts}(t_{INF})$ represents the right tail of the next infection cycle). Such shape is no longer observable if the infection cycle is highly asynchronous (see Figure 2.13a). The WCW model predicts

a log linear rise in the parasitaemia for asynchronous cultures that becomes increasingly terraced with the degree of synchronism of the IRBC population (see Figure 2.13b).

Long-term cultivation of synchronous IRBC populations with periodic subcultivations at 96 hours has been compared to the WCW predictions on the parasitaemia-time curves of an unimodal IRBC population. The observed left-handed bends on the peaks of the parasitaemia are consistent with the skew to younger IRBC forms induced by a staggered P_{death} (see Figure 2.13c). The simulations are carried out using the version *2Dv.2* operating at conditions similar to the *candle-jar* protocols, with fixed death probabilities $P_{death}(S) = 0.002$and $P_{death}(F) = 0.005$. These values for P_{death} result in a probability of completing the infection cycle per IRBC equal to $P_{survival} = 0.64$.

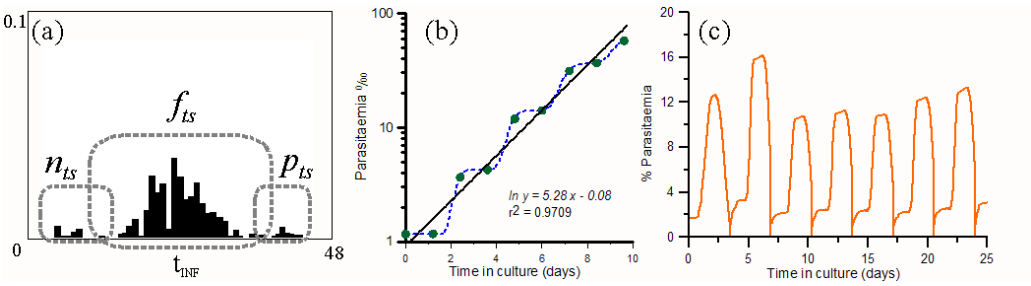

Figure 2.13: *Outcome of the Individual-based model compared to the predictions of a continuous model (White et al., 1992); a) The post-invasion time distribution is composed of the intrinsic distribution of the IRBCs of an infection generation (f_{ts}) plus the contributions of past (p_{ts}) and following (n_{ts}) generations; b) log linear rise of parasitaemia in an unlimited culture. Simulation outcome at each time step (dashed line) and daily samples (•). Terraces and steps smooth in successive infection cycles. Corresponding asynchronous evolution (solid line). Corresponding daily growth ratio, $GR_{24} = 5.28$, is compatible with the values typically observed in real systems; c) long-term cultivation of a synchronous culture with a 4-day subculturing period.*

2.3.5 Substrate availability

Glucose and lactate concentrations in the RBC culture medium are $C_{gluc} = 2.6\ mM$ and $C_{lact} = 0\ mM$, respectively. Maximum rates of metabolized glucose and lactate produced are: $\frac{d[gluc.]}{dt} = \frac{-(122\pm34)\ \mu M}{10^9 RBCs \cdot day}$ and $\frac{d[lact.]}{dt} = \frac{(143\pm47)\ \mu M}{10^9 RBCs \cdot day}$. Finally, adequate *in vitro* parasitaemia yields are obtained when $C_{lact} = 12\ mM$ (Jensen et al., 1983). Considering these data together, global limitations on the parasite development caused by the scarcity of glucose or the excess of lactate are expected not before 20 days.

Nevertheless, heuristic knowledge of the handling of *in vitro* cultures suggests that the culture medium should be replaced every two or three days at most, and it is commonly accepted that this requisite is brought about by the substrate limitations on the parasite. Moreover, it has been claimed that unrestricted parasitaemia yields can be obtained by increasing the fraction volume of medium per volume of the hematocrit, which indicates that the concentration of substrate is somehow determinant for parasite viability (MR4, 2008). Specific trial cultivations carried out by the EMG-GSK show that, under standard static culturing conditions, closed systems can be maintained up to 72 hours, at most (see Appendix B).

Trials with continuous cultures assuming different regimes for the medium renewal show that no significant differences are found between cultures with periodic subcultivations each 48 h and each 72 h (see Appendix B). Figure 2.14 shows the time course of both patterns in two replicas of 10-day cultures. Replicate cultures show identical evolutions. Similar values for the maximum parasitemia ($\%I_{max}$) and the growth ratio are obtained in both subcultivation protocols: $\%I_{max} = 4.9$ % and $GR_{24} = 2.7 \pm 1.0$ for the 48 h pattern; and $\%I_{max} = 4.6$ % and $GR_{24} = 2.2 \pm 0.9$ for the 72 h pattern.

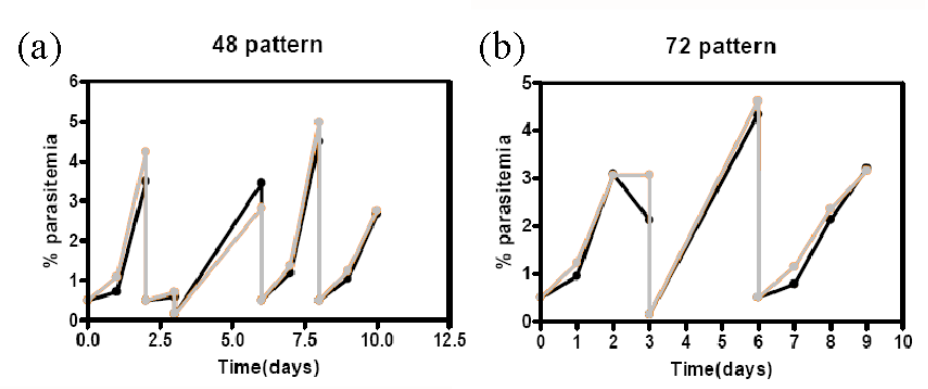

Figure 2.14: *Temporal evolution of the parasitaemia for different subcultivation protocols. a) Periodic subcultures performed each* 48 *h, and b) periodic subcultures performed each* 72 *h. Grey and black solid lines and dots correspond to the two replicas of experimental data.*

The nutrient limitations in closed *in vitro* cultures have been studied through the exploration of the outcome of INDISIM-RBC, which considers glycolysis as the only metabolism for RBCs and IRBCs. Results show that in culturing conditions glucose is the limiting factor, rather than lactate. Simulations of closed systems recreating the current culturing medium to hematocrit ratios suggest that generalized scarcity of glucose

resulting in massive death of IRBCs does not take place (see Figure 2.15a). Threshold parasitaemia yields appear after more than 10 days in culture and at parasitaemias around $\%I = 60\%$. Increasing 100-fold glucose concentration allows reaching $\%I = 100\%$ parasitaemias, and massive extinction occurs after more than 10 days, in this case as a consequence of excessive lactate. The massive death of RBCs when these threshold values are achieved does not entail the decrease of the observed parasitaemia.

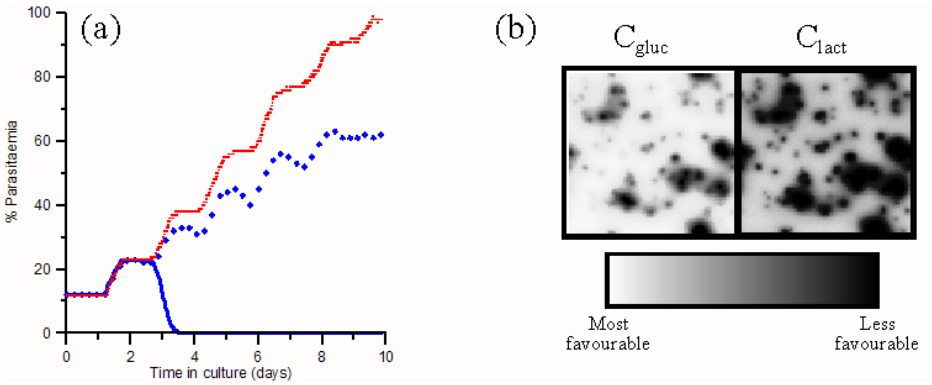

Figure 2.15: *Simulated nutrient limitations on the propagation of the infection; a) Parasitaemia evolution in closed culture systems for different overall initial concentrations of glucose; diamonds (♦) represent* $C_{gluc}(t = 0) = 2.6\ \mu M$*, dotted line (· · ·)* $C_{gluc}(t = 0) = 0.26\ mM$ *and solid line (-)* $C_{gluc}(t = 0) = 0.026\ \mu M$*; b) Auto-scaled visualization of local concentration of glucose and lactate show that glucose is a more limiting factor on the parasite proliferation.*

The results obtained with INDISIM-RBC do not reproduce the observations of real systems. The decay of the infection after three days that is observed in real cultures only appears in the simulations when considering glucose concentrations around 100 times smaller than those of real cultures. These results suggest that the model of the individual metabolism leaves out a key factor that limits the proliferation of the parasite. Despite this shortcoming, results shown in Figure 2.15a can be positively used: INDISIM-RBC is consistent with overall limitations appearing after three days in culture, caused by the scarcity of some limiting solute other than glucose or lactate. Remarkably, such limitations appear when the parasitaemia is around $\%I \sim 20\%$, which is consistent with the experimental observations.

Limitations at a local level appear when the overall conditions are not favorable enough. INDISIM-RBC shows that conditions near the focuses of the infection are less favorable to RBC viability. Figure 2.15b shows that around the IRBCs glucose is scarce and lactate accumulates. RBCs are more susceptible to not being able to fulfill their

metabolic needs near an infected region. The advantage of focusing on local limitations is that they can be skipped with simple manipulations of the culture that modify the spatial structure of the hematocrit (see Section 2.3.6).

INDISIM-RBC-*2Dv.*2 was designed to better account for the factors at a local scale that might be responsible for the limitations observed in real systems. In this version, the temporal scale is reduced with the aim of better capturing the diffusion of substrate and the spreading of the parasite. The results obtained with the improved version still do not show local limitations on the IRBC viability with the real glucose and lactate concentrations. Again, this suggests that either i) there is another metabolic limiting factor, ii) the algorithm used to represent the diffusion of substances through the hematocrit at the current scales is still insufficient, or iii) the current two-dimensional approach obscures the phenomenon occurring in the real hematocrit layer. The diffusion problem in 3D is revisited in Section 3.5.

2.3.6 Agitation of the culture systems

An alternative to the static cultivation of *P. falciparum* (where the RBCs remain settled forming a hematocrit layer at the bottom of culture vials) is the cultivation in suspension (where a slurry formed by RBCs and medium is maintained through the continual agitation of the culture system). Suspended cultures have been claimed to render more abundant harvests than static cultures (Butcher, 1981). Apparently, they show increased infection growth ratios (Butcher, 1982) and support higher parasite yields (Zolg et al., 1982) without unaffordable medium expenses. Furthermore, a method using low hematocrit and high glucose content apparently only required subculturing every 3 days without medium replacement in between (Fairlamb et al., 1985). These published experimental results do not provide strictly conclusive evidences to support the preference of IRBCs for suspended cultures, although they clearly point in this direction.

The expected behavior of the suspension systems is a compromise between two opposite trends:

1. on the one hand, the continual shift of positions of the RBCs favors the propagation of the disease because the number of potentially healthy RBC targets to be invaded by an egressed merozoite is no longer restricted to the immediate neighborhood of the lysing IRBC. In addition, the mix of the hematocrit layer homogenizes the concentration of solute substances, thereby eliminating the effect of local substrate limitations.

2. on the other hand, the reduction of cell density hinders the propagation of the

infection in suspended cultures. This can be understood as the decrease in the probability of encounter between a merozoite and IRBC due to a reduced density, which consequently reduces the rate of transmission.

In addition, the death probability is likely to be increased when the RBCs are submitted to the damaging methods usually employed to maintain cells in suspension (i.e. airlift columns or stirring, shaking or wiggling of the culture systems). Moreover, the continual movement of RBCs could hinder the attachment and invasion of healthy RBCs by merozoites, reducing the infection rate.

Discrete agitation regimes

The above arguments suggest that the agitation of static culture systems at discrete time spans would increase the performance of static *in vitro* cultivation of *P. falciparum.* Indeed, discrete gentle agitation of the hematocrit combines the increase in the number of potential healthy RBC targets (trend 1) with the high cell densities within the hematocrit layer (trend 2). In addition, the potentially harmful effects of agitation are reduced.

Data in the literature suggest that typical RBC sedimentation velocities are around $v_{sed}(RBC) = (2.10 \pm 0.18) \cdot 10^{-3} cm/s$ (Neu and Meiselman, 2001). Settling of RBCs after agitation would take between two and three minutes for custom culture systems with $H_{culture} \sim 2\ mm$ height (MR4, 2008). Measurements of the RBC sedimentation rate carried out by the EMG-GSK in cultures with several heights provided significantly different results, $v_{sed}(RBC) = (1.1 \pm 0.3) \cdot 10^{-5} cm/s$, which entail settling times on the order of $\sim 1\ h$ (see Appendix B). Nevertheless, culture systems submitted to gentle one-time agitations at big enough regular time spans ($T_{agitation} > 12\ h$) can be regarded as static cultures, because this assumption holds for more than 90% of the time in culture.

Discretely agitated static cultures have been implemented by the EMG-GSK. The comparison of the short-term cultivation of the parasite in closed systems under different discrete agitation regimes shows no significant differences between the cultures agitated each $24\,h$, $48\,h$ and $72\,h$ (see Appendix B). Closed cultures last for three days at most (no viable forms are observed in the sample at $96\ h$), show maximum parasitaeimas around $\%I \sim 5\%$ and growth ratios in the range $6 < GR_{48} < 10$. No correlation can be observed between the agitation regime and the growth ratio. For this reason, no experiments of the long term cultivation of discretely agitated cultures have been carried out.

Discretely agitated culture systems are represented by INDISIM-RBC as static cultures where the positions of RBCs are periodically shifted at each agitation. INDISIM-RBC shows no differences between the behavior of static cultures and discretely agitated cultures with different agitation periods. The results obtained in real and simulated

cultures suggest that either the local limitations (regarding the spreading of the parasite and/or the concentration of substances in solution) are not overcome with discrete agitation, or that there are no such local limitations.

Cultivation of *P. falciparum* in suspended cultures

The EMG-GSK has performed several trial experiments to assess the differences between static and suspension culture systems. Suspension cultures are carried out through the continuous agitation of the system with a magnetic stirrer operating at non-damaging velocities (see Appendix B). Trials in closed culture systems show that stirred cultures grow at significantly higher rates than static or discretely agitated cultures do. Nevertheless, like the other closed systems, suspension cultures are still submitted to overall limitations that do not allow the viability of IRBCs after 72 h. These results suggest that limitations on the infection exist both at a local and at a global level.

The positive results of cultures in suspension bring about the trial of long-term continuous cultivation under different agitation regimes (see Appendix B). The growth ratios observed in the continuous trials, $GR_{24}(stirred) = 3.7 \pm 1.7$ and $GR_{24}(static) = 1.8 \pm 0.6$, can be used to build a simple PbM for the spreading of the infection between two successive subcultivations. This model assumes unrestricted propagation of the parasite and fits an exponential function to the observed growth ratios. The parasitaemia at each time step after subculturing are obtained with:

$$\%I_{stirred}(ts) = 0.5\% \cdot e^{0.05/hr \cdot ts} \tag{2.11}$$

$$\%I_{static}(ts) = 0.5\% \cdot e^{0.023/hr \cdot ts} \tag{2.12}$$

As discussed in Section 2.3.5, these models are only valid for low parasitaemias because they assume unrestricted proliferation of the infection. Still, the empirical formulas 2.12 and 2.11 can be considered valid enough during the first 72 h and they can be used to predict the behavior of different subcultivation regimes maintained for a month. Two different subculturing protocols were trialed with prospects of using an automated culturing device. Both consist in performing periodic subcultivations each 48 h, but the man-made renewal sets parasitaemia to 0.5% after subculture, while the semi-automated renewal sets a fixed RBC dilution rate (ν) that leads to the same parasitaemia, according to Equations 2.11 and 2.12. Real static and stirred cultures are compared to the model predictions. Dilutions in semi-automated subculturing are set to $\nu_{stirred} = \frac{1}{11}$ and $\nu_{static} = \frac{1}{3}$, respectively. The predictions of the continuous model are not significantly

different from the experimental observations (Figure 2.16).

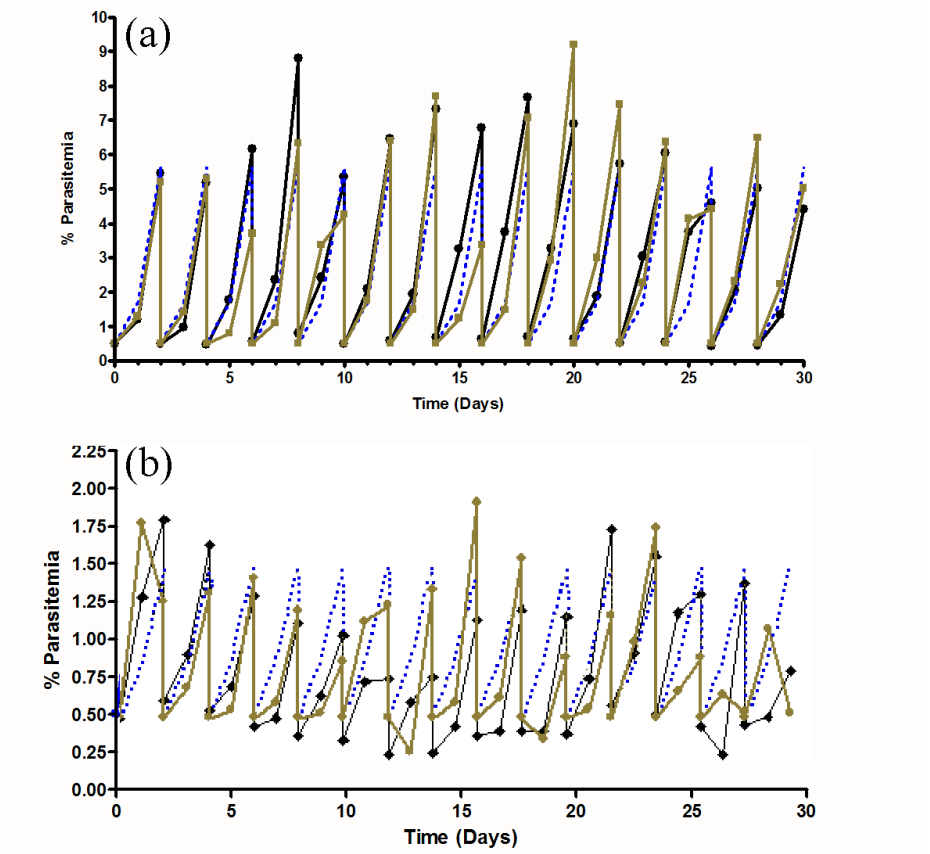

Figure 2.16: *Course of the parasitaemia, observed daily in stirred (a) and static (b) long-term cultivation of P. falciparum infected RBCs (see Appendix B) and compared to the predictions by equations 2.11 and 2.12, respectively. Green solid line corresponds to a culture with daily human subcultivation, to set the initial parasitaemia* $\%I_0 = 0.5$ *after subculture. Blue dotted line shows the expected behavior of automatized subcultivation with ratios* $\nu(a) = \frac{1}{11}$ *and* $\nu(b) = \frac{1}{3}$ *predicted by the model. Black solid line shows the measurements of real automatized cultures.*

Suspension can be represented by INDISIM-RBC as a custom static model with two modifications:

1. The system is continuously mixed. At each time step, the position of the RBCs is shifted at random, merozoites are scattered all through the spatial cells and substrate is continuously homogenized.

2. The RBC density corresponds to the volumetric hematocrit concentration (ranging from 5% to 10% in volume) rather than to the cell density observed in the hematocrit layer (where around 80% of the volume is occupied by RBCs). This marked reduction in cell density is introduced in the model as a decrease in the total number of modeled RBCs. For instance, in a model with 100×100 spatial cells, the representation of a static cultivation entails considering up to 10000 RBCs, while only 500 RBCs are included when modeling a suspended culture.

From the model outcome and the experimental results for the static and stirred cultures, one can state that the two processes can not be reproduced simultaneously with the same set of individual-related parameters: simulation results show that the combination of mechanisms 1 and 2 using the values for P_{inf} and P_{death} that best fit static cultures lead to non-viable cultures. Moreover, finding a pair of values that reproduce the behaviors of real systems is quite difficult and the simulation outcomes show much more uncertainty than what is observed for static cultures.

It is reasonable to assume that the values of the probability of infection and the damaged caused to the cells is different in static and stirred cultures. Firstly, invasion of a healthy RBC by a merozoite placed at its same spatial cell when both cells are settled in the hematocrit gently touching each other is a process that clearly differs from the infection when the two cells collide while drifting around. Secondly, cells submitted to any kind of agitation are necessarily exposed to more mechanical stress, and they can be damaged more easily. Several experiments carried out by EMG-GSK, which are not included in this study because they concern the development of a bioreactor, corroborate this assumption.

A more detailed study of agitated culture systems requires modifying the current versions of the model. Additional or alternative hypotheses regarding cell-cell interactions must be introduced into the model to tackle cultures in suspension.

2.4 Discussion and open questions

INDISIM-RBC proposes a set of rules governing individuals (RBCs and IRBCs) and their local environment (the spatial structure of the system, the extracellular parasite and substrate concentration in the culture medium) that can be used to better understand the behavior of *Plasmodium falciparum* infected erythrocytes in *in vitro* cultures. Several patterns observed in real systems have been used to guide the design and development of the model:

1. *Population structure:* Post-invasion time distributions among IRBCs in short-term and long-term culture systems are quantitatively reproduced by the model.
2. *Population dynamics:* Temporal evolution of the observed parasitaemia and average infection growth rates at different time spans (24, 48, 72 and 96 hours) are quantitatively reproduced by the model.
3. *Population variability:* The trends and variability observed for cultures with different population structures regarding healthy RBCs (age distribution associated with different storage times) and infected RBCs (post-invasion time distribution) are also found in the model outcome.
4. *Infection structure:* several mechanisms of local spreading of the parasite and their resulting modeled infection dynamics are compared to experimental results. The model outcome that is most compatible with the real system-level observations corresponds with the mechanism that is most consistent with the behavior of the parasite reported from individual imaging.
5. *Infection dynamics:* the relation between infection dynamics, IRBC population structure and culturing protocols is consistent with the observation of long-term culture system and with the predictions of accepted population-based models.
6. *Infection variability:* The variability observed for experimental cultures of different strains of *P. falciparum* with different infection capabilities is also found in the model outcome.

This provides a due degree of confidence in the proposed mechanisms and in the underlying structure of the model. Taking the idea a bit further, it can be stated that the quantitative reproduction of a single sort of experimental result alone would not be an appropriate strategy to analyze and validate INDISIM. This can be argued as follows: experimental *in vitro* cultures of *P. falciparum* infected RBCs are complex systems that show both predictable features (patterns) and erratic behaviors (variability). They are conditioned by different factors (strain of parasites, RBC source, culturing protocols...) that can not be disentangled one from the other, and which are hardly assessed through direct measurements. Even the simplest model that expects to capture the all the essential mechanisms that govern the culture evolution must be quite complex. INDISIM-RBC assumes several hypotheses at the same time, each of them concerning a particular aspect of the system and represented by a set of rules and parameters.The model has too many degrees of freedom to be compared to a single variety of experimental results (*a.k.a.* pattern).

An analysis of the sensitivity of the model entails the systematic adjustment of its outcome to a single observable pattern for different experiments and for different sets of parameters in the model. Yet, this analysis alone does not provide enough arguments to hold the model. Instead, several independent experimental patterns reproduced at different scales entail that, even if the mechanisms proposed by the model are just artifacts, they capture many aspects of reality. The robustness of the model, the consistency of its predictions regarding other patterns is also increased: when any experimental observation contradicts the model's predictions, any subsequent modification of the model must reproduce the patterns that are already correctly simulated.

The first versions of the model focus on the processes regarding the individuals and population structure, and on their relations with characteristic features of the culture system. Version *2Dv.1* was converted into a more detailed version, *2Dv.2*, which allows a more accurate description and analysis of processes such as synchronization, nutrient availability and merozoite spreading. Versions *2Dv.1* and *2Dv.2* provide compatible results, although some of the simulation parameters must be readjusted from version to version. In particular, the values for P_{death} and P_{inf} are tuned when shifting from a version to the other.

Several aspects of the simulator can be improved. The individual model could account for different metabolites to better reproduce the limitations observed in closed culture systems. The biological complexity of the RBC model could be increased so that tthe model was more oriented to the cellular level (i.e. whole cell models in system biology). However, the next step in this study was to analyze how the spatial structure of the system and the local environmental affect the infection process.

The approach in 2D does not allow for an accurate description of the local processes. A 3D model has been developed to better understand the spreading of the extracellular parasite and the limitations caused by the diffusion of substrate through the hematocrit layer. Chapter 3 presents this model to analyze how the spatial structure of the culture at a system level affects the course of static *in vitro* cultures.

Chapter 3

P. falciparum infected RBCs in static cultures

The first versions of INDISIM-RBC (see Section 2.2) proved their validity in representing the mechanisms governing the propagation of *Plasmodium falciparum* in RBC *in vitro* cultures, as well as their capacity to provide understanding concerning the observed population dynamics (see Section 2.3). However, a more accurate description of the processes affecting the local environment of the RBCs is required when comparing different culturing devices (see Section 2.4).

The current chapter focuses on modeling the spatial structure of static *in vitro* culture systems. In particular, on how the geometrical characteristics of the hematocrit layer of settled RBCs influence the prevalence and the course of the infection. Firstly, the experiments carried out to explore different culturing setups are presented (Section 3.1). Secondly, the modifications on the 3D version of the Individual-based model are outlined and an abridged ODD description of INDISIM-RBC.*v3D* is presented (Section 3.2). Thirdly, some general results used to validate the 3D version of the simulator are put forward (Section 3.3). Finally, two underlying mechanisms that may cause the behavior observed in real static *in vitro* cultures are analyzed: a) local limitations on the parasite proliferation (Section 3.4), and b) local degradation of the culture medium due to insufficient diffusion rates (Section 3.5). These potentially limiting factors are tackled through population-based models supported by the outcome of INDISIM-RBC.*v3D*.

3.1 Geometrical characterization of the static culture systems

3.1.1 Preliminary considerations on the application of INDISIM-RBC.*v3D*

The initial steps in the modeling of any real system consist in finding an appropriate representation of the system under study. INDISIM-RBC correctly reproduces the observed behaviors of the cultures and apparently captures some of the underlying mechanisms of real systems. It can be considered a good tool to simulate real systems.

The next step is applying the results of the model to better understand and control real systems. INDISIM-RBC.*v3D* was developed and is used as a support tool to facilitate the analysis of different culturing setups with the aim of designing improved culturing devices.

In this sense, it is important to recall that one fundamental aim of the collaboration between MOSIMBIO and EMG-GSK was to use the models for practical purposes. According to the EMG-GSK, the improvement of the *in vitro* culturing protocols initially followed two lines of investigation with different objectives:

1. One of the objectives was to better understand the factors and constraints on the static cultivation of malaria parasites in order to optimize its performance, better reproduce the conditions *in vivo* and interpret the results of drug trials.

2. The second objective was to develop an automated continuous suspension culture, an automatized bioreactor capable of maintaining steady conditions for the infected RBCs. Such an apparatus would be used to maintain the culture lines of different parasite strains and would also serve to perform batteries of drug trials.

The results concerning line (1) were presented in Ferrer et al. (2008) and constitute the main part of this chapter. They comprise a systematic study of the constraints arising from the geometry of the hematocrit layer, which covers hematocrit layer depths that range from 0.06 mm to 3 mm, and a separation between the walls of the culturing device that varies from 7.5 mm to 9 cm.

The detailed study of experimental systems that differ only in their three-dimensional configuration demands the development of an explicit 3D model. The setting up of the version *v3D* required some additional measurements regarding the spatial structure of real static culture systems, both on the local scale (e.g.. counting the number of multiply infected IRBC as an indirect measure of the propagation of the merozoite, or assessing

the average RBC density within the hematocrit layer) and at the level of description of the whole system (e.g. measuring the macroscopic geometry of the culture system and its effects).

The research concerning line (2) led to preliminary considerations on the design of a prototype of a suspended cell airlift bioreactor. However, the model outcome suggested that suspension cultures required exploring new roads. The conclusions drawn from the application of INDISIM-RBC.*v3D* are similar to those presented in Section 2.3.6: a different set of rules for the interactions must be developed to deal with suspended cultures.

The research revealed that it is also beyond the scope of the model to deal with some technical limitations found in the agitated experimental systems. For instance, how to handle continuous subcultivation of the parasite (the model cannot explain the observed loss of infection capability associated with this process) and how to control hematocrit and parasite leak through the filters and valves of the culturing device, among others. The study of suspension culture systems is left as a perspective for further work. Thus, the associated experimental trials are not presented here.

3.1.2 Experimental trials on static culture systems

The model in 3D attempts to capture the essence of the local environment of the RBCs, in order to analyze how the spatial configuration of the static cultures affects the infection dynamics. For this reason, the following set of experiments have been carried out to gain a more a detailed depiction of the hematocrit layer under typical culturing conditions.

Multiple infections in IRBCs

The simplest hypothesis regarding the invasion of RBCs by merozoites is to assume that it is a homogeneously random process, in the sense that every healthy RBC is equally susceptible to invasion. If this hypothesis holds, then the distribution of the number of malaria parasites per IRBC follows a Poisson distribution. However, if not all the RBCs are equally susceptible to invasion (i.e. if the pool of susceptible erythrocytes is smaller than the total number of RBCs, as occurs when the infection is a locally limited process), the population of IRBCs must bear an increased number of parasites. The result is an increased proportion of multiparasited IRBCs, which implies that the distribution of the number of invasions per IRBC has a longer tail than the one predicted by the Poisson distribution (Simpson et al., 1999).

The Poisson distribution ($f_P(k;\lambda)$) expresses the probability of a number of events (k) occurring in a fixed period of time if these events occur with a known average rate (λ). It is given by:

$$f_p(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{3.1}$$

A corrected Poisson distribution function is used to analyze the experimental results. This is a Poisson distribution that does not consider the probability of no events occurring ($k = 0$). It can be used to tackle the distribution of the number of invasions within the population of IRBCs. The probability of no events occurring can be associated with the population of non infected RBCs: $f_p(k = 0;\lambda) \sim 1 - \%I \sim 90\%$.

The observations of multiple infected parasites carried out by EMG-GSK (see Appendix B) have been compared to the corrected Poisson distribution function (Figure 3.1).

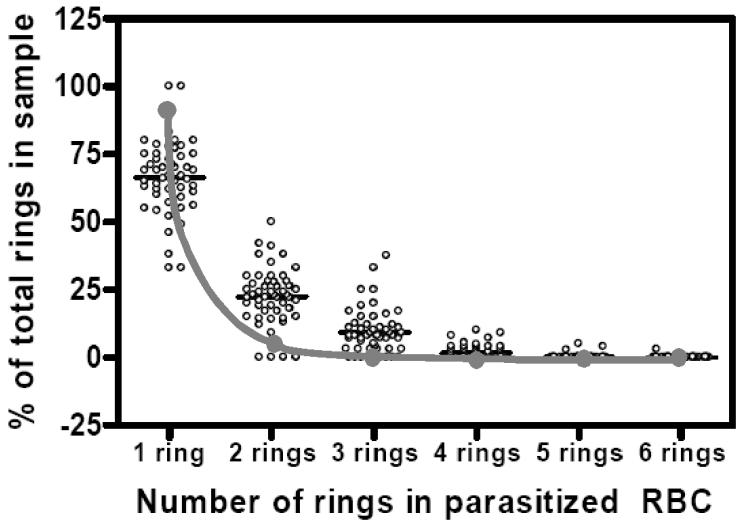

Figure 3.1: *Percentage of multiple infected RBCs. White dots (∘): observed experimental counts in over 51 thin smears, each containing 2500 RBCs; grey dots (•) and solid line (-): theoretical ratio obtained with a corrected Poisson distribution assuming* $\lambda = 0.1$.

The parameter λ in the Poisson distribution represents the average number of parasites per RBC. Assuming a maximum threshold parasitaemia ($\%I = 10\%$), the expected number of parasites is $\lambda \sim 0.1$. The values obtained with the corresponding Poisson distribution ($f_p(k;\lambda = 0.1)$ are significantly different from those observed in real cultures

(Figure 3.1). The correlation factor between observed (f_{obs}) and theoretical values is only $r^2 = 0.17$. More realistic parameters for the parasitaemia (i.e. $\%I < 10\%$) yield worse fits. In conclusion, it may be stated that infection of RBCs is not a completely random process.

The measurement of multiple infected RBCs *in vitro* has been carried out to better understand the process of invasion and to check the validity of the 3D-model for the local spreading of extracellular merozoites (see Section 3.3.3).

Cell density in the hematocrit layer

A determining factor for the propagation of the infection in the model is the average distance between RBCs (see Section 2.3.6), which is related to the cell density in the hematocrit. This is measured as the fraction of volume of the layer occupied by RBCs, *a.k.a.* RBCs average packing factor or just packing factor (pf).

Different culture devices are filled with culture medium plus healthy RBCs at 5% hematocrit, simulating an ordinary static culture. Their base surface (S) is small enough to allow an accurate measurement of the Hematocrit Layer Depth (HLD) and big enough to neglect capillarity effects. The real depth (HLD_{obs}) is measured once red blood cells are completely settled down. The depth expected if RBCs were completely packed without interstitial space (HLD_{calc}) is obtained by multiplying the culture depth (H_{total}) by the hematocrit fraction ($\%H = 0.05$). The real packing factor is calculated as the fraction of expected to real depths: $pf = \frac{HLD_{calc}}{HLD_{obs}}$. A description of the experimental methods is outlined in Appendix B. The results obtained are shown in Table 3.1.

Culture device	$S\ [cm^2]$	$H_{total}\ [cm]$	$HLD_{calc}\ [cm]$	$HLD_{obs}\ [cm]$	pf
Glass 1	2.41 ± 0.05	8.7 ± 0.1	0.435 ± 0.005	0.5 ± 0.1	0.87 ± 0.17
Tube 1	0.50 ± 0.05	42.0 ± 0.1	2.1 ± 0.005	2.4 ± 0.1	0.88 ± 0.04
Tube 2	0.28 ± 0.05	66.5 ± 0.1	3.325 ± 0.005	3.9 ± 0.1	0.85 ± 0.02

Table 3.1: *Estimated RBC density within the hematocrit layer, according to the measurements using three different culture devices: glass 1, a flat-bottomed test tube; tubes 1 and 2, plastic pipes similar to straws with a flat silicone seal at the bottom; S: base surface, calculated out from the measured diameter; H_{total}: measured culture depth; HLD_{calc}: expected depth of the hematocrit layer; HLD_{obs}: measured depth of the hematocrit layer; and pf: calculated packing factor of RBCs.*

The experimental RBC packing factor can be compared with the packing factors corresponding to theoretical spatial distributions of different objects. For instance, RBCs can be approximated to rigid spheres with the same volume ($r_{sphere} = 2.71\ \mu m$). These

spheres can be arranged in a perfectly ordered and optimally packed distribution, with all the spheres in contact with each other and leaving minimum interstitial space between them. This spatial distribution corresponds to the so called Face-Centered-Cubic configuration (FCC). Its theoretical packing factor is known to be $pf_{FCC} \simeq 0.74$. Or spheres can be arranged in a random 3D distribution, but still placed so that each sphere is in contact with its immediate neighbors (RS). This configuration leaves some gaps between the spheres, and the resulting theoretical packing factor is $pf_{RS} \simeq 0.64$ (Berryman, 1983).

Alternatively, RBCs can be approximated to rigid disks of the same volume ($d_{disk} = 7.4\ \mu m$; $h_{disk} = 2\ \mu m$). They can be arranged in a perfectly ordered and optimally packed honeycomb-like distribution (HD), with a packing factor $pf_{HD} \simeq 0.91$. Instead, disks randomly packed in two dimensional layers in contact with each other (RD) give a packing factor of $pf_{RD} = 0.82$ (Berryman, 1983).

The packing factor observed in real cultures is $pf_{exp} = 0.87 \pm 0.07$. The observations are most consistent with the ordered placement of rigid disk-like RBCs. Such a result tallies with the tendency of RBCs to form regular aggregate structures (ruleaux). Even so, a more disordered arrangement of RBCs is still compatible with observations in real cultures, if we consider RBCs as non-rigid bodies that can be distorted to better fit with each other.

The experimental results provide additional information: increasing HLD does not lead to an increase in the packing factor. This supports the hypothesis that there is no compacting of the hematocrit layer due to hydrostatic effects. It can be assumed that pf_{exp} remains constant throughout the hematocrit layer, no matter the value of HLD. This lack of compacting hypothesis is also compatible with theoretical arguments indicating that the net weight of the settled RBCs over any single cell is negligible when compared both to the cellular stiffness (Shung et al., 1982; Godin et al., 2007) and to the interactions among RBCs (Hochmuth and Marcus, 2002).

Shape of the hematocrit layer

The hematocrit layer and the free culturing medium behave as distinct phases in static cultivation: as long as the culture is not stirred, hematocrit and medium are regions with distinguishable physical properties. This jibes with the observation of a phenomenon typically associated with the surface tension of liquids: the formation of hematocrit puddles and the appearance of a meniscus in the interface between the hematocrit, the free medium and the walls of the culture vial.

The real shape of the hematocrit layer observed for certain geometries of the culturing vial is not strictly a thin flat deposit that covers all the bottom of the culturing device.

On a (non-wettable) glass surface, small volumes of culture show hematocrit layers that resemble a puddle and cover only a fraction of the total surface of the culturing device. In plastic culture vials (wettable surface), the hematocrit may stick to the walls and form a meniscus (see Figure 3.2). This observed behavior can be explained as follows.

Both healthy and infected RBCs are subject to some extent to cell-to-cell attractive interactions: rouleaux and other kinds of RBC-RBC aggregates provide an adhesion energy per unit surface (ϵ) estimated to be $\epsilon_{RBC} = 1 \cdot 10^{-4} \frac{N}{m}$ (Hochmuth and Marcus, 2002). The enhanced adhesiveness of IRBCs caused by knobs, which is exhibited in the formation of rosettes, contributes to increasing this stickiness (Chotivanich et al., 2000). The inter-cellular attraction produces a net inward pull on the cells placed on the border of the hematocrit, which can be measured as surface free energy (Foty and Steinberg, 2005).

Surface free energy (*a.k.a.* surface tension) resulting from molecular interactions is used in physics to describe the behavior of different phases of liquids. This analogy with liquids can be extended to define a capillary length (L_C): $L_C = \sqrt{\frac{\epsilon}{g \cdot \rho}}$, where ρ is the density of the liquid, and g is gravity. L_C is the characteristic length scale in which inter-cellular interactions are comparable to gravitational energy. It defines the shape of the puddles of liquids, as well as the extent of the observed meniscus.

An estimation of the order of magnitude of the L_C corresponding to the hematocrit can be carried out assuming that the surface free energy between the hematocrit and the culture medium is $\epsilon_{H20} \sim 1 \cdot 10^{-1} \frac{N}{m}$, and that the value of L_C of the culture medium is $L_c(H_2O) \sim 2\ mm$. This results in:

$$L_C(RBC) = \sqrt{\frac{\epsilon_{RBC}}{\epsilon_{H_2O}}} \cdot L_C(H_2O) \sim 0.06\ mm \tag{3.2}$$

L_C determines the maximum depth that the hematocrit layer can reach when RBCs are settled forming a puddle in the bottom of a glass vial without contacting its walls. It can be used to calculate the culturing surface (S) for small volumes of the hematocrit cultured in big vials. L_C is also associated with the range of the meniscus, where the hematocrit layer is affected by the presence of the walls of the culturing device. It can be used to interpret the effect of the distance between walls in static cultures.

The usual static culturing conditions imply dimensions much bigger than the capillary length, so the hematocrit layer may be considered a flat thin bed. Then, HLD is calculated from the total culture volume (V), the hematocrit concentration ($\%H$), the observed culturing surface and the fraction of the hematocrit layer occupied by RBCs (pf).

$$HLD = \frac{V \cdot \%H}{S \cdot pf} \tag{3.3}$$

This calculated HLD is used to characterize the effect of the thickness of the hematocrit layer on the parasite development in static *in vitro* cultures with different surfaces and volumes.

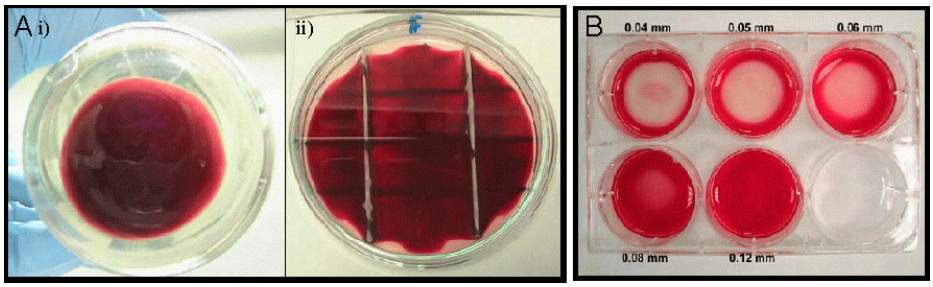

Figure 3.2: *Observed shape of the hematocrit layer in static culture systems. A) Glass culturing devices produce a convex meniscus. i) flat bottomed bottles used in the experiment* B *series. ii) petri dishes with plastic separators used in the* P *series (two separators). B) Plastic culturing devices produce a concave meniscus. Six-well plates used in the* W *series.*

Effect of the geometry of the hematocrit layer on the infection course

Static cultures are often referred to as 'thin-layer cultures', yet the exact meaning of 'thin' is not specified. Standardized protocols recommend two types of static *in vitro* cultures to carry out the *candle-jar* technique (MR4, 2008):

i) $V \sim 4\ ml$ with $\%H \sim 5\%$ in plastic flasks with a base surface $S = 25\ cm^2$, and

ii) $V \sim 12\ ml$ with $\%H \sim 5\%$ in flasks with $S = 75\ cm^2$.

If RBCs are settled in a perfectly compact layer, the thickness of the hematocrit layer is $HLD = 0.09\ mm$.

In most of the culturing protocols found in the literature, the hematocrit layer depth is not explicitly defined. Its calculated value ranges from $HLD \sim 0.2\ mm$ (in the *candle-jar* method: $V = 1.5\ ml$, $\%H = 10\%$, $S \simeq 10\ cm^2$ and $pf \simeq 0.85$) to $HLD \sim 1.2\ mm$ (in cultures in aliquots: $V = 0.5\ ml$, $\%H = 10\%$, $S \simeq 1.5\ cm^2$ and $pf \simeq 0.85$) (Trager, 1994).

For the smallest cultures, HLD is only one order of magnitude greater than the characteristic length of RBCs (the diameter of the erythrocyte is $D_{RBC} = 6 - 8\ \mu m \sim$

$1 \cdot 10^{-2}$ mm). In consequence, the HLDs proposed in literature represent significantly different backgrounds from the RBCs point of view. Such differences in HLD may be extremely important if the processes occurring at a local level (e.g. uptake of nutrients, local spreading of the parasites, among others) are limiting factors on the propagation of the infection.

The base surface of the culture systems found in the literature also varies from trial to trial, ranging from $S = 1.5$ cm^2 to $S = 75$ cm^2. In this case, the dimension of the system is always several orders of magnitude greater than the characteristic length of the RBC, so the relative differences at the RBC scale are not so evident.

The effect of the dimensions of the hematocrit layer (HLD, V) on parasite development (parasitemia $\%I$ and growth ratio GR) is analyzed through three series of experiments performed by EMG-GSK. Several static *in vitro* cultures are raised under similar culturing conditions and different geometric characteristics. Three trials compare different base surfaces, culture volumes and shapes of the culture vials. Their performance in parasitaemia and daily growth ratio is measured.

P-series) 2-day culture trials covering a range of base areas after fixing both the hematocrit layer volume and depth. They are carried out to evaluate the effect of the walls of the culture device. Cultures are set in glass Petri dishes, which have been previously split into detached sub-regions by gluing plastic separators onto the plate base surface. Different distances between the glass separators (L) are trialed. The growth ratio at each sub-region within a plate is evaluated separately. Each plate has a control sub-region with a fixed surface to rule out differences not associated with the variation of L.

W-series) 14-day trials varying the hematocrit layer depth by modifying the total culture volume with a fixed base surface. They are carried out to evaluate the effect of HLD. Cultures are set in plastic trays, each containing 6 culture wells, and use small culture volumes (from 1 ml to 10 ml) .

B-series) 18-day trials varying the hematocrit layer depth (HLD) by modifying both the total culture volume (V) and the culturing surface (S). They are carried out to evaluate the effect of HLD and of the total culture volume. Cultures are set in flat-bottomed glass bottles and use large culturing volumes (from 2 ml to 100 ml).

The geometric characteristics of the experimental sets are depicted in Figure 3.2. They are also presented together with the experimental results in Table 3.3. Specifications of the culture methods are detailed in Appendix B.

Trial name	V *(ml)*	V_{RBC} *(µl)*	L *(cm)*	S *(cm^2)*	V_H *(µl)*	HLD *(mm)*	$\%I$ *(%)*	GR_{48}
P1	16.5	825	0.75	6.8	129.2	0.19	0.35 ± 0.19	0.6 ± 0.8
P2	16.5	825	1	9.0	171.0	0.19	0.6 ± 0.2	1.3 ± 0.8
P3	16.5	825	2	16.0	304.0	0.19	0.7 ± 0.4	2.2 ± 2.1
P4	16.5	825	4	32.0	608.0	0.19	0.9 ± 0.3	3.3 ± 1.6
P5	16.5	825	9	63.6	1208.4	0.19	1.2 ± 0.5	4.1 ± 1.5

Trial name	V *(ml)*	V_{RBC} *(µl)*	D *(cm)*	S *(cm^2)*	V_H *(µl)*	HLD *(mm)*	$\%I$ *(%)*	GR_{48}
W1	1.0	49.4	3.5	9.62	73.73	0.060 ± 0.015	1.5 ± 1.2	5.2 ± 1.8
W2	1.5	74.1	3.5	9.62	110.60	0.09 ± 0.02	1.6 ± 1.1	5.4 ± 1.9
W3	3.1	155.5	3.5	9.62	194.03	0.19 ± 0.05	1.7 ± 1.3	6 ± 2
W4	5.8	290	3.5	9.62	232.09	0.34 ± 0.08	1.5 ± 1.0	4.9 ± 1.0
W5	10.0	500	3.5	9.62	388.06	0.59 ± 0.15	1.4 ± 1.0	5 ± 2
B1	2.6	130	3.1	7.54	432.84	0.18 ± 0.05	1.8 ± 1.2	6.3 ± 1.2
B2	5.2	260	3.3	8.55	432.84	0.34 ± 0.08	1.7 ± 1.2	6.4 ± 0.9
B3	5.8	290	3.4	9.08	746.27	0.32 ± 0.08	1.7 ± 1.6	7 ± 3
B4	10.1	506.5	3.6	10.18	755.97	0.56 ± 0.14	1.4 ± 0.8	5.1 ± 0.8
B5	10.4	520	3.5	9.62	776.12	0.60 ± 0.15	1.4 ± 1.0	5.1 ± 1.8
B6	20.8	1040	3.9	11.94	1552.24	1.0 ± 0.2	1.1 ± 1.5	3.3 ± 0.7
B7	26.0	1300	4.1	13.20	1940.30	1.1 ± 0.3	1.0 ± 0.4	3.0 ± 0.5
B8	51.5	2074	5	19.63	3841.79	1.5 ± 0.4	0.7 ± 0.2	1.6 ± 0.3
B9	68.6	3432	5	19.63	5122.39	2.0 ± 0.5	0.50 ± 0.19	1.6 ± 0.6
B10	74.1	3705	5	19.63	5529.85	2.2 ± 0.5	0.4 ± 0.2	1.0 ± 0.6
B11	98.8	4940	5	19.63	7373.13	2.9 ± 0.7	0.3 ± 0.16	0.8 ± 0.4

Table 3.3: *Macroscopic geometric characteristics and experimental results of P-, W- and B-trials. V: measured total culture volume, V_{RBCs}: calculated volume of packed RBCs, V_H: calculated volume of the hematocrit layer,L: measured distance between separators; D: measured diameter of the hematocrit layer deposit, S: calculated base surface, HLD: calculated depth of the hematocrit layer, $\overline{\%I}$: average observed parasitaemia and GR_{48}: calculated growth ratio at 48 h. W refers to cultures in 6-well plastic plates and B to cultures in 5 cm diameter flat-bottom glass bottles. W assays with volumes lower than 1.0 ml have not been presented because they did not develop well, due to the effects of surface tension. Reprinted from Ferrer et al. 2008, Malaria Journal* 7: 203.

3.2 ODD Description of INDISIM-RBC in 3D

3.2.1 Need for a 3D model

The results obtained with versions *2Dv.1* and *2Dv.*2 show that the model can be improved to better reproduce some features observed in real cultures.

The need to impose a maximum threshold on the number of parasites that can invade a single RBC (Section 2.3.3) is aimed to improving the model of the propagation of extracellular merozoites through the hematocrit layer.

The absence of diffusive limitations that may cause the local exhaustion of medium in closed systems (Section 2.3.5), and which could explain the differences found between static and agitated cultures (Section 2.3.6), demands revisiting the model of substrate diffusion.

The detailed study of different experimental static cultures also stimulates the development of INDISIM-RBC.*v3D.* The aim is to compare and understand the different performances observed in culture systems that are distinguished only by differences in their geometry.

A description of the model is outlined below, following the scheme of the ODD but detailing only the specific modifications of this version. The complete description of INDISIM-RBC is found in Section 2.2.

3.2.2 ODD description of the 3D model

Overview

Purpose:
The model aims to decipher how the spatial structure of the hematocrit layer affects the static *in vitro* cultivation of *Plasmodium falciparum* IRBCs.

Entities and state variables:
Two low level entities are those defined in the 2D versions, the RBC and the spatial cell (*sc*). Their characteristic variables remain unaltered. The sole modification pertains to the extracellular merozoites. In the current version merozoites are tackled one by one, yet they are not considered as individual entities because they do not have individual characteristics beyond their time spent in the extracellular medium and their position in the spatial grid (Grimm, 1999).

Characteristic scales:

The model is spatially explicit and represents a small patch of the hematocrit layer. In this version (*v3D*), the space is modeled as a regular 3D grid formed by $I \times J \times K$ spatial cells, typically, with $I = J = 20$ *sc* and $K = 40 - 1000$ *sc*. Each spatial cell is a square of $l_{sc} = 5$ μm side. Therefore, the whole grid represents a column with a base surface of 0.01 mm^2 and a depth slightly bigger than the thickness of the hematocrit layer (HLD). Processes are modeled discretely and events take place at finite time steps. The time step (*ts)* is set to 1 hour. The value of the simulation unit that represent substrate particles (*su*) is set to 10^6 molecules.

Boundary conditions:
The fraction of the culture system included in the model is a column of the hematocrit layer and a small fraction of the culturing medium that covers it, the extent of the Diffusion Boundary Layer (DBL). Therefore, the limiting surfaces of the model have different boundary conditions.

i) *Top boundary:* The upper layer of spatial cells represents the free culturing medium. Bulk medium can be considered as an infinite source of glucose and a reservoir for lactate during the spans between two successive subcultivations. For this reason, top cells are considered as a closed boundary, meaning that they exchange substances only with side and lower spatial cells, and still maintain a substrate concentration fixed to the initial values: $C_{gluc}(i, j, HLD + DBL; t) = C_0$ and $C_{lact}(i, j, HLD + DBL; t) = 0$.

ii) *Bottom boundary:* The lower layer of cells lies on the surface of the culturing device. Bottom cells are considered a closed boundary that may accumulate the substrate, waste and deposits that settle in the hematocrit layer

iii) *Side boundaries:* Typically, the modeled fraction of the culture system is placed in the middle of the hematocrit layer, so it is surrounded by near replicas of the model. For this reason, the cells of the side surfaces are characterized with Periodic Boundary Conditions (PBC).

Eventually, side boundaries are considered a closed boundary to represent the effect of the side walls of the culturing device.

A depiction of the characteristic scales, entities and spatial structure of the model is depicted in Figure 3.3. The list of variables, their values (in simulation units and

measured values) and the reference to the reference source to set them are presented in Table 2.1.

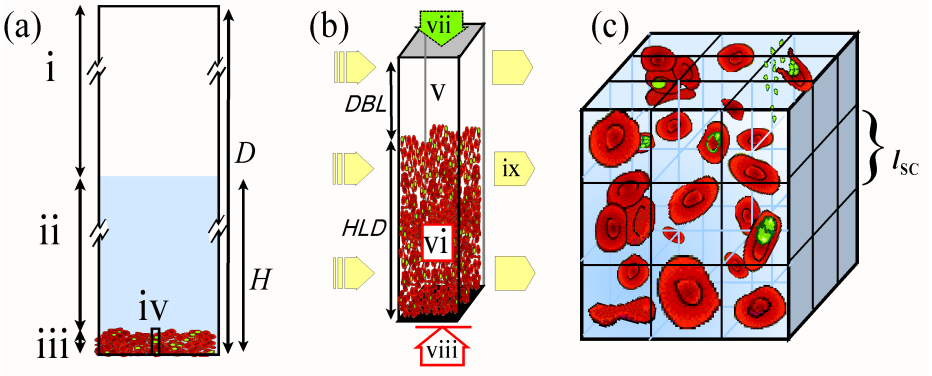

Figure 3.3: *Depiction of the entities, variables and spatial structure of the 3D versions of INDISIM-RBC. a) Side view of an experimental static culture, according to the MR4 protocol: i) controlled atmosphere, dimensions of the culture vials* $D \sim 3$ *cm ii) free culture medium, depth of the culture* $H \sim 5$ *mm, iii) hematocrit, iv) modeled fraction of the culture system; b) spatial model of the hematocrit in 3D, measures* $X \times Y \times Z$, with $X = Y = 0.1$ *mm* and $Z = HLD + DBL$, *v) boundary layer, measures* $DBL \sim 0.1$ *mm, vi) hematocrit layer, its depth ranges* $0.1\ mm < HLD < 2\ mm$, *vii) top layer: reservoir boundary conditions, viii) bottom layer: closed boundary conditions, ix) side walls: periodic boundary conditions; c) 3D view of a fraction of the spatial model, size of the spatial cell* $l_{sc} = 5\ \mu m$.

Process overview and scheduling:
The general procedure and scheduling of *v3D* are the same as in versions *2D* (see Section 2.2).

Design concepts

The design concepts stated in Section 2.2 remain unaltered, except for the observation of the model outcome.
Observation: The graphical interface of the model shows data collected at the end of each time step (label 7iii in Figure 2.1). Figure 3.4 shows a screen shot of the graphical interface of the *v3D* version of the model. It is comprised of four graphical windows and a numeric display showing system-level variables that are updated on the fly. The windows represent (1) the temporal evolution of RBCs, (2) merozoites and IRBCs, (3) the vertical cross-section of the hematocrit layer showing total number of RBCs, IRBCs, merozoites and packing factor for each stratum, and (4) the IRBC age structure. The

numeric display (5) shows the number of time steps and equivalent time in real cultures, and the number of RBCs, IRBCs and extracellular merozoites.

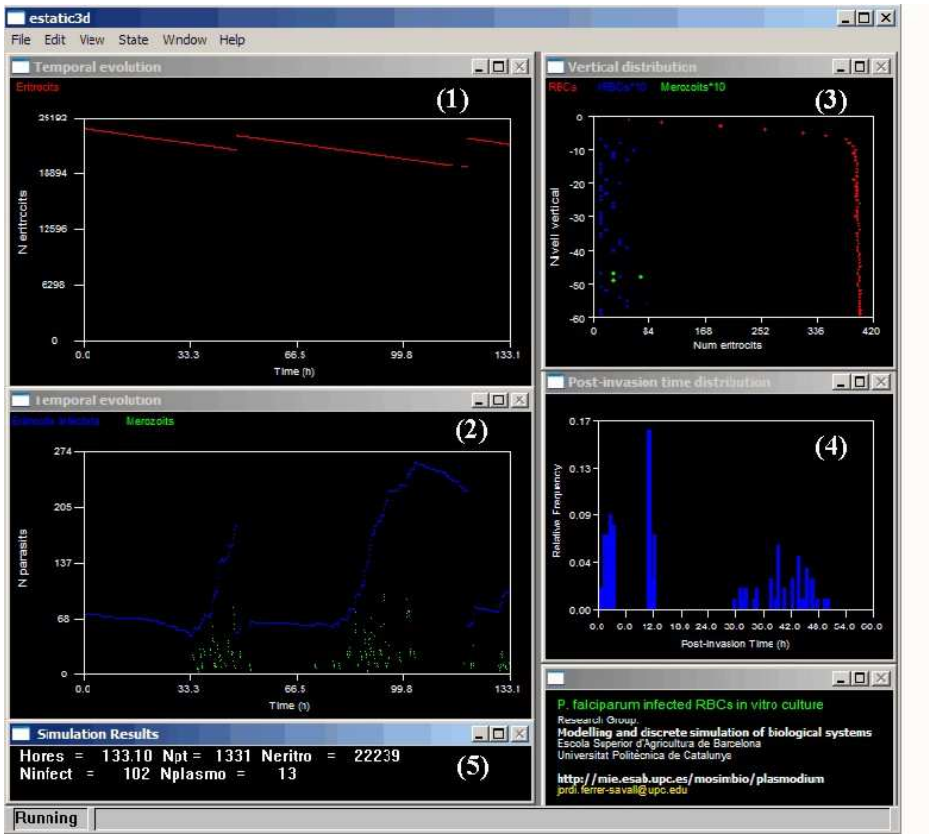

Figure 3.4: *Screen shot of the version v3D of the model. This version operates with a time step set to six minutes. The size of the spatial grid in this simulation is 20x20x60 spatial cells. The initial population is 24000 RBCs and the initial parasitaemia 1%. Windows show: 1) temporal evolution of RBC population (red line); 2) temporal evolution of parasitaemia (blue) and merozoites (grey); 3) vertical representation of the hematocrit layer: total number of RBCs (red), IRBCs (blue) and merozoites (green) per stratum; 4) age structure, histogram of the post invasion times among the IRBC sub-population; 5) Numerical display showing the time step, number of RBCs, IRBCs and merozoites.*

Details

The *Initialization* and *External inputs* in version *v3D* are like those defined in 2D versions. Most of the *Submodels* also remain unchanged, except for: 1a) RBC Motion, 1d)

RBC infection, 1e) RBC death, 2a) SC Diffusion and 2b) SC Propagation of extracellular merozoites. These modifications are the consequence of considering that the spatial cells have a finite volume, that can be filled with RBCs, merozoites and the solid RBC remains that do not dissolve into the culture medium after cell lysis.

1a) RBC Motion. Apart from the processes described in Section 2.2, RBCs can now move through the hematocrit layer as a consequence of the settling process. Motion of RBCs may occur when there is enough room in the 9 spatial cells that comprise the nearest neighbors immediately below them. This shift of positions occurs with a fixed probability P_{fall}. Each RBC may fall on one spatial cell, at most, per time step.

1d) RBC Infection. The infection process does not vary from the mechanism described in Section 2.2, except for three modifications:

i) merozoites are randomly distributed among one of the 27 nearest neighboring spatial cells, including the same cell. The probabilities of being dumped into each of the spatial cells are listed below. Same spatial cell (i, j, k): $P(1) = 0.33$, side-to-side cell (for instance (i+1, j, k)): $P(2) = 0.035$, immediate diagonal cell (for instance (i+1, j +1, k)): $P(3) = 0.025$, and double diagonal cell (for instance (i+1, j+1, k+1)): $P(4) = 0.02$ (See Figure 3.5).

ii) the extracellular merozoites can spread through the hematocrit layer during four time steps after merozoite egress ($\sim$ 30 minutes). Each merozoite has a probability of moving to one of the nearest neighboring spatial cells in the level immediately below the present cell, at each time step. As for RBC motion, given that there is enough room, the probability of falling to any of the spatial cells is given by P_{fall} (see Figure 3.5).

iii) No limitations are imposed on the maximum number of parasites that can invade a healthy RBC.

1e) RBC Death. After cell lysis the fraction of volume corresponding to solid unsolvable materials, which represent approximately $V_{remain} = 6\ \mu m^3$, are left in the same spatial cell. At the end of each time step, just after merozoite propagation, these remains may fall to any of the nine nearest neighboring cells in the stratum immediately below, just as RBCs do.

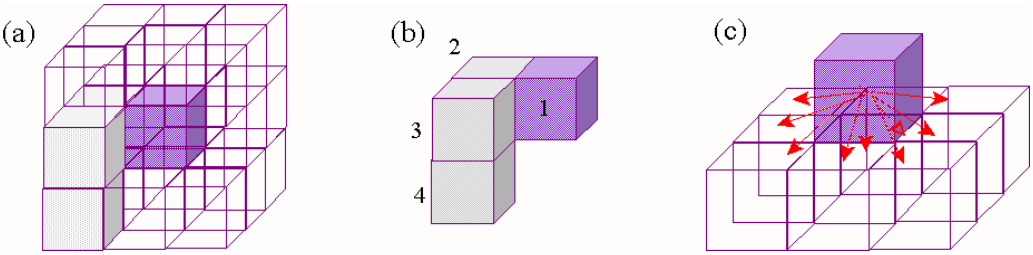

Figure 3.5: *Stencils of the 3D models for: a) the egress of merozoites at the end of the infection cycle; b) detail showing the four different types of neighboring cells (i.e. 1) same cell, 2) side-to-side neighbor, 3) immediate diagonal and 4) double diagonal); c) extracellular propagation through the hematocrit layer to spatial cells. Lysis occurring at the grey spatial cell.*

2a) SC Diffusion. Substrate diffusion is initially explicitly modeled with a FCTS as in Section 2.2.

$$C^{t+1}_{i,j,k} = C^{t}_{i,j,k} + \widetilde{D} \sum_{l,m,q}^{nn(i,j,k)} w_{l,m,q} C^{t}_{l,m,q} \tag{3.4}$$

l, m and q are the spatial coordinates of the 27 cells that constitute the immediate environment of cell (i,j,k). Equation 3.4 adds up the contributions of each cell to the diffusive transport of substrate. The set of weights $w_{l,m,q}$ are:

$$w_{l,m,q} = \begin{cases} \frac{1}{2(6\sqrt{\frac{3}{2}}+3\sqrt{3}+4)} & ;\ 8\ \textit{cubic diagonals}:\ (l \neq i) \cap (m \neq j) \cap (q \neq k) \\ -1 & ;\ \textit{same spatial cell.}\ (l,m,q = i,j,k) \\ \frac{1}{2(\frac{6}{\sqrt{2}}+3+\frac{4}{\sqrt{3}})} & ;\ 12\ \textit{square diagonals, i.e.}\ (l \neq i) \cap (m \neq j) \cap (q = k) \\ \frac{1}{2(6+3\sqrt{2}+4\sqrt{\frac{2}{3}})} & ;\ 6\ \textit{square sides, i.e.}\ (l \neq i) \cap (m = j) \cap (q = k) \end{cases} \tag{3.5}$$

Substrate diffusiveness is initially set to its maximum values $\widetilde{D} = 0.8$. The choice of this value is arbitrary, motivated by the best fit of the simulation outcome, but is not deduced from a description at a lower level. It is not known whether this value correctly accounts for real diffusion in the hematocrit. The submodel has been switched off in the applications of INDISIM-RBC in Section 3.4, which focuses on the extracellular propagation of the parasite. The problem regarding the model of diffusion is analyzed in Section 3.5.

3.3 Analysis of experimental observations with the 3D model

3.3.1 Calibration of version *v3D*

The experimental results on the evolution of the parasitaemia for the short-term preservation (Pavanand et al., 1974) and long-term cultivation (Trager and Jensen, 1976) of *P. falciparum* in static *in vitro* cultures are used to calibrate the set of parameters employed by the model in the current version *v3D*.

The values for $\%I_0$, $\bar{t}_{INF}(t=0)$, P_{death} and P_{inf} are maintained as the ones used in version *2Dv.2* (see Sections 2.3.1 and 2.3.2). One additional parameter is introduced to represent a specific three-dimensional phenomenon: the settlement of the hematocrit layer, which is modeled as the fall of individual RBCs and merozoites ($P_{fall} = 0.05$). The graphical outcome of the simulations that best fit the experimental results are presented in Figure 3.6 and extended in Ferrer et al. 2007.

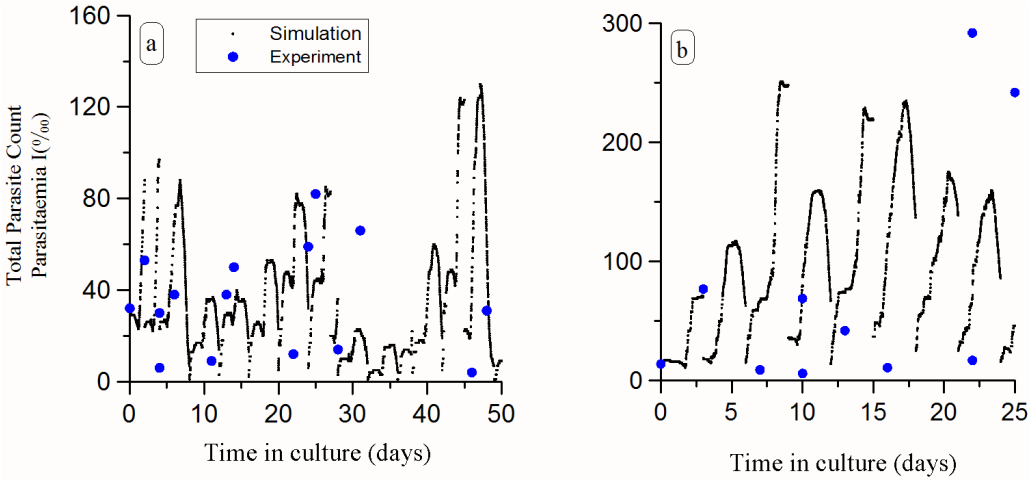

Figure 3.6: *Simulation results compared to experimental data of the overall infection course in: a) Continuous-flow static cultivation of P. falciparum; and b) static cultivation through the Candle-jar method (Jensen and Trager, 1978).*

3.3.2 Settling processes in the hematocrit layer

RBCs are slightly denser than the culturing medium. For this reason, they sink to form hematocrit layers. Once the hematocrit is formed, it can be assumed that the settling

process may continue wherever there is room available on inferior levels. The subsidence of RBCs and of the material that remains after cell lysis (cell membrane and organelles that are not dissolved in the medium) through the hematocrit layer may play a role in the local limitations to the propagation of the infection. Several mechanisms modeling this settling process have been trialled, namely:

1. No fall of RBCs, merozoites and remains. Voids are left after each RBC lysis and the hematocrit layer gradually turns into a porous structure with homogeneously distributed holes containing 1 spatial cell, and bigger cavities in the regions where the infection has been spreading. These cavities caused by the successive lysis of IRBCs are empty enclosed spaces that usually contain 16 spatial cells (this is the number of RBCs that get infected and burst during the span between two consecutive sub-cultivations).

   These cavities do not impede nor facilitate the spreading of individual parasites, but they indirectly slow down the propagation of the infection because extracellular merozoites may spend several time steps in non occupied spatial cells.

2. RBCs, merozoites and remains fall to a lower stratum with a fixed probability P_{fall}, only if there is enough room in the spatial cell immediately below them.

Infection growth ratios (GR) are higher when compactation methods (2 and 3) are introduced. However, the simulated GRs do not significantly vary from one mechanism to the other if the probability of falling is small enough ($P_{fall} \leq 0.05$). Simulations with the first mechanism eventually show unrealistic configurations of the hematocrit layer (for instance, a vertical profile where a stratum has nearly no RBCs while the strati above it are full).

The sedimentation rate of real RBCs in the free culturing medium depends on many physical constraints (for instance the terminal velocity, achieved when the downward force of gravity minus buoyancy equals the upward force of drag) and physiological factors (for instance, the presence of substances promoting or inhibiting RBC adhesiveness, or the membrane potential of RBCs). The order of magnitude of the RBC sedimentation rate is around several millimeters per hour (this means that the RBC may fall down several spatial cells per time step). However, the fall of RBCs through the hematocrit layer presumably involves much slower precipitation velocities due to the interaction among cells.

The averaged cell density of the simulated hematocrit ($\overline{pf}_{sim}$) depends on the settling mechanism and on the sedimentation rate. Simulation results that fit the experimental cell density observed in the real hematocrit layers are obtained with mechanism 2 and

$P_{fall} = 0.05$. Using these values, the simulated cell densities in the hematocrit layer range from $\overline{pf}_{sim} = 0.7$ to $\overline{pf}_{sim} = 0.9$.

3.3.3 Multiple invasion of RBCs

The propagation of the infection is modeled as a local invasion process that occurs when a merozoite contacts a healthy RBC with a probability described by $P_{inf}(t_{RBC})$ that varies with RBC age (t_{RBC}). Given that it is a stochastic process, the distribution of RBCs invaded by more than one parasite can be assumed to be a Poisson distribution with $\lambda = \overline{P_{inf}}$.

Following the discussion in Section 3.1.2, one can posit the inverse question: that is, to determine the value λ that best fits the observed corrected Poisson distribution. Once this value has been set, the fraction of RBCs that are effectively susceptible to invasion can be calculated.

The distributions of multiply parasitized IRBCs that best fit with the distributions observed in real cultures (see Figure 3.7) are obtained when the model uses average invasion probabilities that range from $\overline{P_{inf}} = 0.8$ to $\overline{P_{inf}} = 1$.

Of note, no significant differences are found when the values of the parameters defining P_{inf} are varied. Therefore, it can be deduced that the lack of availability of healthy RBCs is more related with local limitations on the process of spreading of merozoites, than with the values of P_{inf} . According to the Poisson distribution that best fits experimental data, the fraction of non-infected RBCs should be $f_P(0; \lambda = 0.85) = 40\%$. Instead, the observed fraction of healthy RBCs is $f_{obs}(0) = 1 - \%I \simeq 95\%$. Therefore, the estimated fraction of RBCs that are not effectively available to the infection is around half the population of RBCs (in fact, around $\sim 55\%$).

This prediction is consistent with the rules defining the local propagation of the parasite through the hematocrit with low parasitaemia and randomly distributed IRBCs: first, as long as the parasitaemia is below 4%, the hematocrit layer is composed of unit cells that consist on one IRBC and 26 healthy RBCs; second, only 16 (on average) of these healthy RBCs can eventually be invaded by an egressed merozoite. This makes the fraction of RBCs susceptible to invasion around $\sim 60\%$. In conclusion, the observed distribution of multiply infected parasite forms is consistent with the individual models of healthy RBC invasion and merozoite spreading.

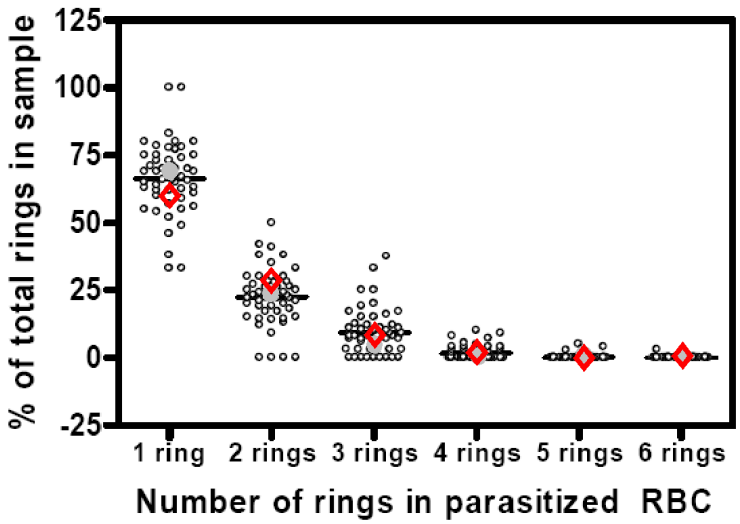

Figure 3.7: *Multiple infections expected by the model compared to experimental results observed in in vitro trials. White small circles (○): experimental counts; big solid grey circles (●): simulation outcome with the typical $P_{inf} = 0.8$; big empty diamonds (◊): simulation outcome with $P_{inf} = 1$ for all RBCs.*

3.4 Local limitations on the infection process

The experiments performed by EMG-GSK (see Section 3.1) reveal two system-level geometric characteristics of the hematocrit layer that affect the *in vitro* development of the parasite: depth (HLD) and separation between walls of the culturing device (L). Culture characteristics and observed results for W, B and P trials are presented in Table3.3.

The observed measurements throughout each trial show a strong time correlation, as a consequence of both the parasite infection cycle and the external manipulation of the culture system. In fact, each single trial can be regarded as the evolution of a single culture, thus characterized by the average parasitaemia throughout the whole culture trial ($\overline{\%I}$), or as a set of replicas of the evolution of the culture between two subsequent subcultivations, thus characterized by the average growth ratio (GR_{48}). The latter magnitude is used to compare the trialled geometries.

The data sets from the different trials are compared one with each other using non-parametric statistics, such as the Kruskal-Wallis (KW) test (see Appendix C) because GR_{48} data sets from any single trial do not come from a normal distribution. Experi-

mental observations are also compared both to the results obtained with a system-level model, and to the outcome of INDISIM-RBC. Comparison of theoretical predictions and experimental results is carried out using parametric statistics.

3.4.1 Effect of the walls of the culturing device

Firstly, the effect of base surface (S) of the hematocrit layer on *in vitro* development of the parasite is assessed. Heuristic knowledge of the experimental group indicates that the extension of the hematocrit turns out to be important when the separation between the walls (L) of the culturing device is small enough. The data obtained from different trials (see Table3.3) show that the development of the parasite is significantly hindered by small separations between walls. However, the functional dependence of this hindrance on L can not be unequivocally determined: the observed average growth ratio at 48 hours from the P series can be fitted to the same degree of confidence both to a linear regression and to an inversely linear regression (see Figure 3.8).

Best fits to the experimental observations are:

the linear fit $GR_{48}^{1}(L) = 0.93 + 0.46 \cdot L$, with $r^2 = 0.86$, and

the inversely linear fit $GR_{48}^{-1}(L) = 4.67 - \frac{3.29}{L}$, with $r^2 = 0.91$.

The chi-square (χ^2) test applied to the experimental data series shows that either dependence is supported by a statistical significance of only $\sim$ 90% (there is a 10% chance of obtaining equal or better fitting results assuming that experimental data do not come from the distribution proposed by the model).

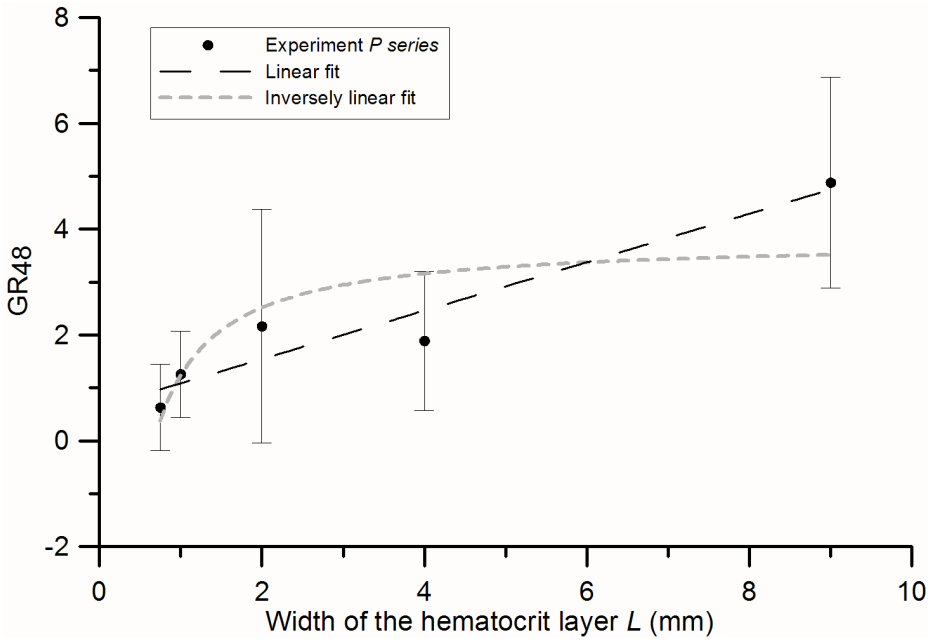

Figure 3.8: *Observed dependence of the average growth ratio at 48 h on the distance between separators (L). Dots with error bars represent the observed data. Dashed line (–) denotes the best linear fit. Dotted grey line (· · ·) denotes the best inversely linear fit (L-model).*

System-level model of the effect of the walls (L model)

The inversely linear dependence can be explained by means of a whole-system model that considers an exclusion region in the vicinity of the walls of the culturing device where the infection cannot progress (see Figure 3.10a). Under this assumption, the expected average growth ratio of a culture as a function of the separation between walls ($GR_{48}(L)$) is described by:

$$GR_{48}(L) = A_1 \frac{2L_{EXC}}{L} + A_2(1 - \frac{2L_{EXC}}{L}) \tag{3.6}$$

where A_2 is the growth ratio in the exclusion region (L_{EXC}), and A_1 is the bulk growth ratio ($A_1 > A_2$). The parameters that best fit this curve to the experimental data set are shown in Table 3.4. According to this fit, the exclusion region would be spread over approximately $L_{EXC} = 2.5\ mm$. Many microscopic mechanisms may be speculated as being responsible for creating this exclusion region (e. g., limitations on the spread of the metabolic waste products, hindered propagation of the parasite due to the existence of a meniscus) but it is not possible to discriminate among them with the current information.

The observed exclusion region is consistent with the expected exclusion region induced by the capillary length (L_C) of the culture medium, but not by capillarity of the hematocrit layer (see Section 3.1.2).

3.4.2 Effect of the hematocrit layer depth

The KW test showed that data obtained from W and B series come from the same distribution function with significance greater than 99.9%. For this reason, all the measurements have been grouped in a single data sample. Average values are given for the data corresponding to similar geometric conditions, for each of the subgroups: {W3+B1}, {W4+B2+B3} and {W5+B4+B5}.

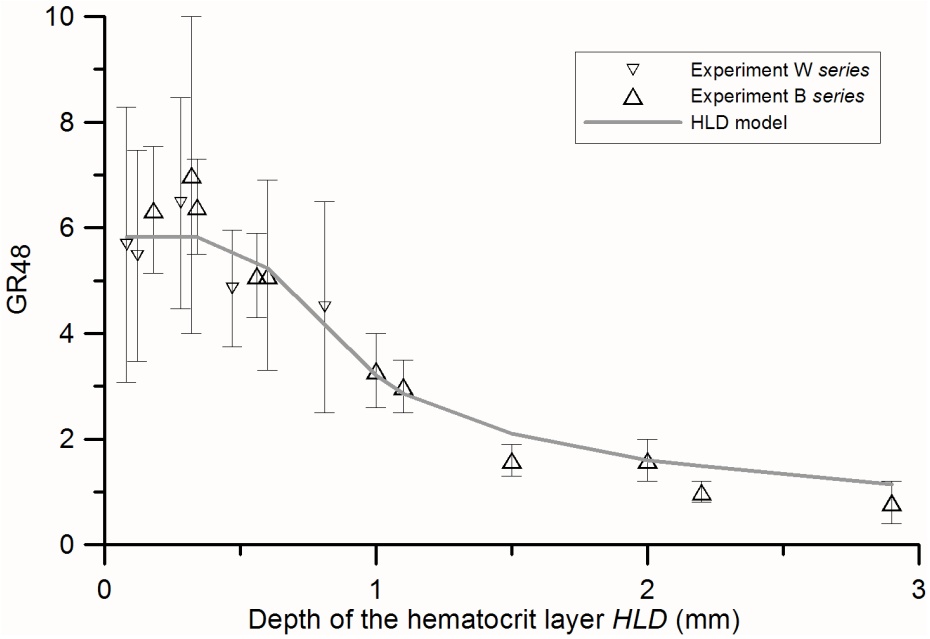

Figure 3.9: *Dependence of the parasitaemia on the hematocrit layer depth (HLD). Triangles with error bars represent the observed data of W (▽) and B (△) series. The solid line denotes the best fit for the system-level model HLD.*

Experimental data show asymptotic inversely linear behavior for large values of HLD (Figure 3.9). Optimum parasite development occurs in cultures with HLD between 0.18 mm and 0.34 mm.

System-level model of the effect of hematocrit layer depth (*HLD* model)

A simple system-level model that reproduces the observed behavior consists in splitting the hematocrit layer into discrete regions that have different behaviors. The inversely linear decay for deep cultures is reproduced when the hematocrit layer is split into two horizontal regions: $HLD = h_1 + h_2$. The first region has a fixed depth $h_1 = h$ and shows a high fixed infection multiplication ratio B_1, while the infection spreads at a lower rate ($B_2 < B_1$) in the remaining part of the hematocrit layer ($h_2 = HLD - h$, see Figure 3.10b). Let B_1 and B_2 be the multiplication ratio per infection cycle in each of the above mentioned sub-regions, and HLD the total depth of the hematocrit layer. The average infection growth ratio at 48 h, $GR_{48}(HLD)$ is then given by:

$$GR_{48}(HLD) = \begin{cases} B_2 + (B_1 - B_2) \cdot \frac{h}{HLD} & if\ HLD > h \\ B_2 & if\ HLD < h \end{cases} \tag{3.7}$$

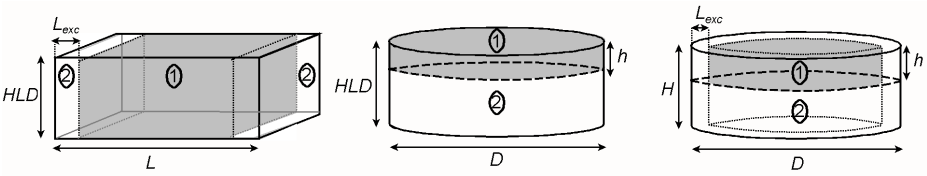

Figure 3.10: *Depiction of the whole system model of the hematocrit layer. a) Schema of the culture system according to the model used to tackle P trials. HLD indicates hematocrit layer depth, L stands for the separation between walls of the culturing device and L_{exc} represents the extent of the exclusion region, where the spread of the infection is hindered. The shaded sub-region (1) indicates the fraction of the hematocrit layer where the rate of infection spreading is high. b) Schema of the culture system according to the model used to tackle B and W trials. HLD indicates hematocrit layer depth, D stands for the diameter of the hematocrit layer and h represents the extent of the exclusion region, where the spread of the infection is hindered. The shaded sub-region (1) indicates the fraction of the hematocrit layer where the rate of infection spreading is high. c) Schema of the model to tackle all the geometric effects simultaneously. The propagation of the infection is hindered both by the walls of the culturing device and by diffusive limitations. The shaded sub-region (1) indicates the fraction of the hematocrit layer where the rate of infection spread is high.*

The parameters that define the curve that best fits with the experimental data set are shown in Table 3.4. The parameters have been estimated using a numerical approximation, and their likelihood given the observations has been checked using both the Kolmogorov-Smirnoff (KS) test and the chi-squared test (χ^2). The former statistic has

been taken as the reference for the statistical significance of likelihood, because it implies fewer restrictions on both data sets and provides a smaller value for the degree of confidence of the results.

3.4.3 Whole-system model that accounts for HLD and L together (WS model)

The models presented in sections 3.4.1 and 3.4.2 can be merged into a system-level representation of the whole culture (Figure 3.10c). The evolution of the culture as a function of the geometric variables L and HLD is shown in:

$$GR_{48}(L; HLD) = \begin{cases} K_2 + (K_1 - K_2) \cdot \frac{h}{HLD} \cdot (1 - \frac{2L_{EXC}}{L}) & if\ HLD > h \\ K_1 \frac{2L_{EXC}}{L} + K_2(1 - \frac{2L_{EXC}}{L}) & if\ HLD < h \end{cases} \tag{3.8}$$

The values for L_{EXC} and h have been adopted from the best fit values found in Sections 3.4.1 and 3.4.2. The two other parameters (K_1 and K_2) refer to the growth ratios in the sub-regions with high and low parasite propagation, respectively, and are calculated from the best fit to experimental data. The obtained values are presented in Table 3.4. A graphical comparison between the theoretical predictions and the experimental behavior is shown in Figure 3.11. The KS test applied to expected and observed data sets gives a confidence between 90% and 95% for the proposed model.

System-level model	Characteristic parameters			p-value
L-model	A_1	A_2	L_{EXC} (mm)	
	4.1	0.6	2.45	0.03
HLD-model	B_1	B_2	h (mm)	
	5.8	0.4	0.48	$<$ 0.001
WS-model	K_1	K_2	h, L_{EXC}	
	5.4	0.4	set	0.0012

Table 3.4: *Characteristic parameters of the PbM that best reproduces the experimental observations.* L-model *considers the effect of the distance between the walls of the culturing device.* L_{EXC} *is the region near the walls where the propagation of the parasite is hindered;* HLD-model *considers the effect of the depth of the hematocrit layer.* h *is the horizontal layer where the propagation of the parasite is not hindered;* WS-model *merges both models.* L_{EXC} *and* h *maintain their values from the best fits of the* L- *and the* HLD- *models. The values for* K_1 *and* K_2 *are set to optimize the likelihood of the model outcome to experimental data. Significance of the obtained results* ($p - value$) *has been assessed using the Kolmogorov-Smirnoff (*KS*) test.*

This merging allows checking of the consistency of the models proposed above both

with each other and with the experimental observations. It enables the specification of the most appropriate geometric conditions for the static in vitro cultures of *P. falciparum* infected erythrocytes.

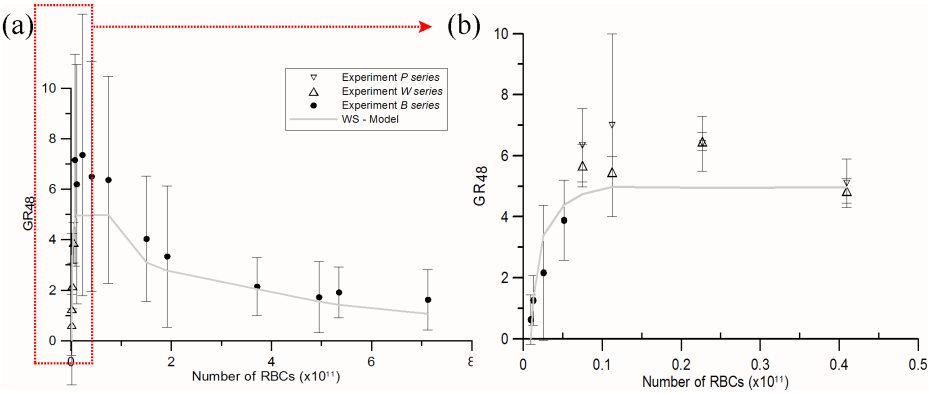

Figure 3.11: *Dependence of the parasite growth ratio (GR_{48}) both on the hematocrit layer depth (HLD) and width (L), represented together as the number of cultured RBCs. Solid dots (●) with error bars represent observed data from P trials. Empty triangles with error bars represent the averaged observed values from W (△) and B (▽) trials. The solid line (-) represents the best fit values provided by the WS model. a) Experimental data sets from W and B trials on the whole; b) detailed view of trials with small culturing volumes.*

3.4.4 Outcome of INDISIM-RBC

The experimental results have been compared with the corresponding outcome of the simulations carried out with a version of INDISIM-RBC.*v3D* that accounts for whole-system model. The WS assumptions were introduced as *ad hoc* constraints on the rules governing the individuals (parasites and RBCs) at a cellular and local level. The simulation space is split into two horizontal sub-regions (layers 1 and 2, respectively; see Figure 3.4.3b) with different probabilities of infection ($P_{inf}(1)$ and $P_{inf}(2)$, respectively). The simulations also allow assessing of the effect of the walls on real cultures. This is implemented through placing a vertical wall on one of the side boundaries of the simulation space, thereby changing one of the periodic boundary conditions for a pair of closed and open vertical boundaries facing each other.

INDISIM-RBC is used to tackle solely the geometric constraints on parasite proliferation at a cellular level, so just a few parameters from the general model are modified: P_{max} is set to $P_{inf}(1)$ and $P_{inf}(2)$ at each of the sub-regions. The probability of infection for mature RBCs ($P_{min} = 0$), the rate of local spreading of merozoites ($P_{fall} = 0.05$)

and all the remaining parameters have been held to the values fixed by the simulation results of the *candle-jar* method (see Section 3.3.1). The estimated values for $P_{inf}(1)$ and $P_{inf}(2)$ are those that provide the simulation outcomes that best fit with the macroscopic observations on the average growth rate at 48 hours (GR_{48}), and they are presented in Table 3.5.

Characteristic parameters of the IBM	$p_{inf}(1)$	$p_{inf}(2)$	p_{fall}	Significance
Best fit values	0.51	0.01	0.05	0.014

Table 3.5: *Characteristic parameters of theIbM that best reproduces the experimental observations. The values for the parameters representing the infection probability in the upper (1) and lower (2) regions of the hematocrit layer ($p_{inf}(1)$, $p_{inf}(2)$) as well as the sedimentation rate(P_{fall}) have been set through the optimization of the likelihood of the model outcome to experimental data. Significance of the obtained results has been assessed using the Kolmogorov-Smirnoff (*KS*) test; the presented values show the probability of obtaining better results with the null hypothesis.*

A graphical comparison between the outcome of the IbM and the whole-system models is shown in Figure 3.12. The estimation of the optimal values was carried out through a systematic, yet not exhaustive exploration of the space of parameters; better fits could be found by the recursive refinement of the optimization protocol.

Such an improvement of the values at a cellular level has not been carried out because parameters such as P_{inf} stand for very specific characteristics of the parasite strain, blood sample and culturing conditions. Therefore, the accurate estimation of the best-fit values alone does not provide general insight concerning the culture system. The likelihood of the simulation results, given the experimental observations, was checked using both the Kolmogorov-Smirnoff (KS) test and the chi-squared test (χ^2). Again, the former statistic has been taken as the reference for the statistical significance of the likelihood, because it implies fewer restrictions on both data sets and it provides a smaller value for the degree of confidence of the results.

3.4.5 Discussion

The systematic study of different macroscopic culturing conditions has allowed for the building of a quite simple whole-system model which is compatible with the experimental observations and which may account for some as yet poorly understood phenomena.

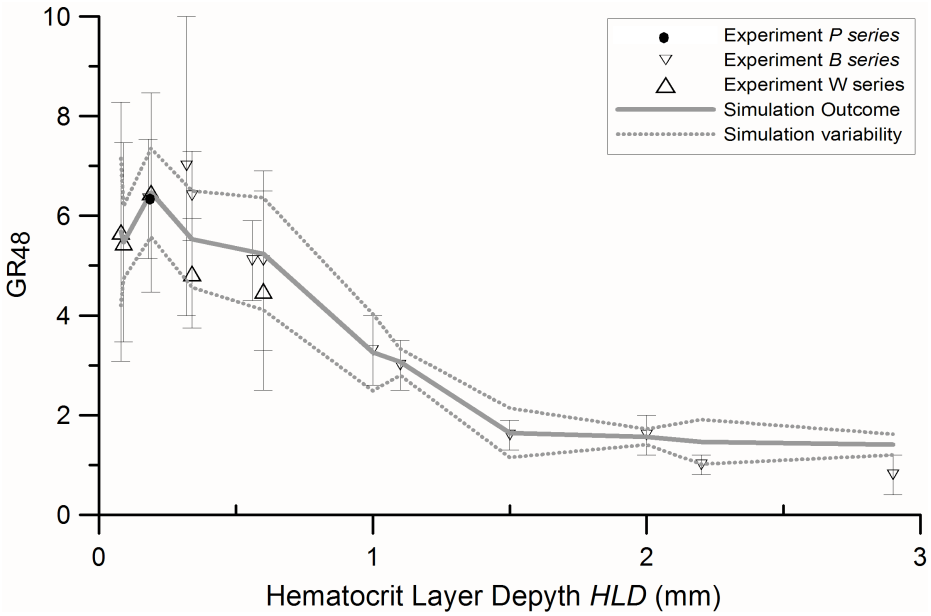

Figure 3.12: *Dependence of the growth ratio of the simulations on the hematocrit layer depth (HLD). Symbols with error bars represent the observed data, (●: P-series, △:W-series and ▽: B-series). Grey solid line denotes the best fit for the INDISIM-RBC simulation outcomes. Dashed lines represent the deviation in the outcome observed for 4 simulation runs of each of the observed points.*

Some conclusions may be drawn from the experimental results, assuming the whole-system model:

1. Cell-cell interactions such as erythrocyte aggregation and rosette formation around parasitized cells can be accounted for as average inter-cellular binding energy that determines the macroscopic shape of the hematocrit layer in the *in vitro* cultivation of *P. falciparum*-infected erythrocytes. Under custom culturing conditions, the hematocrit layer can be considered as a flat film, but such depiction is not valid when the hematocrit volume decreases. At small volumes of hematocrit the intracellular binding energy (which can be tackled as surface tension on the hematocrit boundaries) is comparable to gravitational energy, so the hematocrit must be regarded as a sessile drop at the macroscopic scale. Different shapes of the hematocrit are observed depending on the material of the culturing device.

2. According to the L-model, the spread of the infection is hampered by short distances between walls (L), and cultures are expected to be strongly hindered when $L <$

$L_{EXC} \sim 2.5\ mm$. By extrapolating Equation 3.6, it is deduced that the effect of the exclusion region can be overlooked in the WS-model when $L > 2\ cm$ with more than 95% confidence.

3. The geometric conditions of the culture systems at a macroscopic level of description play an important role in parasite development. According to the HLD-model, the spread of the infection is hampered by thick hematocrit layers. In particular the hindrance appears when the depth of the hematocrit layer exceeds a maximum threshold value $HLD > h \sim 0.478\ mm$.

4. A bottom-up approach can be used to check the validity and consistency of the system-level models. The simulation outcomes are consistent with observations when the hematocrit layer is split into two sub-regions. Such an IbM also enables specific study and treatment of the relevant processes occurring on the scale of the parasite. As a result, some additional conclusions may be drawn from analysis of the IbM:

   (a) The maximum threshold for the growth ratio at the zone of high parasite proliferation is geometrically fixed: the multiplication of the number parasite is not enough to ensure an appropriate GR_{48}, but a minimum spreading of the merozoites is also required. This is shown by the model because increasing the maximum probability of individual infection (P_{inf}) above a certain threshold value does not entail an increase in the infection growth ratio. In contrast, greater diffusion rate for the extracellular parasite through the hematocrit layer (P_{fall}), or the slight rift of RBCs (P_{fall}) does affect the global growth ratio.

   (b) The macroscopic sub-regions of different parasite proliferation (layers 1 and 2 in Figure 3.10b) are externally imposed on INDISIM-RBC as *ad hoc* modifications on the rules governing the parasite proliferation at a local level. The reduction of sub-region 1 in Figure 3.10c can be recreated with the IbM as a consequence of the perturbations caused by introducing a closed boundary into the simulation. The quantitative reproduction of the observed behaviors has not been achieved because simulations representing the whole hematocrit layer are unfeasible due to excessive computational demands.

   (c) The observed reduction in the average GR_{48} for very thin hematocrit layers (infeasibility of the harvest when $HLD < 0.04\ mm$, and noticeable hindrance on cultures with $HLD < 0.1\ mm$) is reproduced with INDISIM-RBC even

though it has not been introduced as an *ad hoc* external input on the rules governing the model. This is an example of an emergent behavior that arises from the local interactions among individuals. It is caused by the effect of the existence of a closed boundary of the system on the propagation of the parasite.

According to the obtained results as a whole, most appropriate dimensions of tested HLDs range from 0.18 mm up to 0.34 mm and most convenient distances must exceed $L > 2\ cm$.

To sum up, the systematic study of a wide range of geometric configurations for the hematocrit layer allows for predictive capacity formalized in a system-level phenomenological model. The resulting whole system model is consistent with INDISIM-RBC as soon as the hematocrit layer is split into two sub-regions with different infection proliferation rates.

The bottom-up approach can not provide justification for the *ad hoc* phenomenological laws. In particular, the mechanism responsible for the emergence of two phases in the hematocrit layer is not included (but just assumed) in INDISIM-RBC. Heuristic arguments support the idea that limitations are due to the local scarcity of substrate in deep regions of the hematocrit layer, as a consequence of a limited diffusion.

3.5 Local substrate limitations in the hematocrit layer

The differences observed between different geometries in the static *in vitro* cultures of *P. falciparum* infected erythrocytes have been reproduced in Section 3.4.4 by defining a model that splits the hematocrit layer (HL) into two subregions: one with plentiful proliferation of the parasite and the other where the infection is hindered.

This section analyses how diffusion of substances through the hematocrit layer can affect the viability of healthy and infected RBCs and lead to the formation of these two regions. *Grosso modo:* RBCs need a continuous supply of substrate and clean surroundings in order to maintain their viability (IRBCs are even more demanding and dump more metabolic waste). During the course of the cultivation, the activity of cells leads to the local exhaustion of the medium, which appears more markedly around the IRBCs (see Section 2.3.5). The degradation of the environment can harm or even kill the RBCs, thus hindering or even stopping the proliferation of the infection.

3.5.1 Diffusion limitations in the hematocrit layer

Diffusion limitation in static cultures arises from the joint effect of the individual uptake and the limited diffusiveness through the hematocrit layer. Diffusion in the hematocrit layer can be correctly described with the reaction-diffusion equation (Eq. 3.9), which relates the temporal variation of the concentration ($C(\overrightarrow{x},t)$) at a given point ($\overrightarrow{x}$) characterized by the fraction of diffusing medium ($\epsilon_w = 1 - pf$, the fraction of volume not occupied by the cells), with the mechanisms that cause it:

i) the diffusion term, which is proportional to the second derivative of the concentration ($\triangle C = \sum_{i=1}^{N} \frac{\partial^2 C(\overrightarrow{x},t)}{\partial x_i^2}$) and to the effective diffusivity (D_e), and

ii) the reaction term, which accounts for local consumption of substrate or average uptake rate ($U(C)$).

$$\epsilon_w \frac{\partial C(\overrightarrow{x},t)}{\partial t} = D_e \cdot \triangle C(\overrightarrow{x},t) - U(C(\overrightarrow{x},t)) \tag{3.9}$$

The values for the effective diffusiveness in the hematocrit layer are taken from the literature and presented in Table 3.6.

	Free Culture Medium	Hematocrit Layer
Glucose	$D_0 = 9.3 \cdot 10^{-6}\ cm^2/s$	$D_{eff} = 0.24 \cdot D_0 = 2.2 \cdot 10^{-6}\ cm^2/s$
Lactate	$D_0 = 1.5 \cdot 10^{-5}\ cm^2/s$	$D_{eff} = 0.19 \cdot D_0 = 2.9 \cdot 10^{-6}\ cm^2/s$

Table 3.6: *Effective diffusiveness of glucose and lactate. Diffusiveness in the Free Culture Medium are set to their respective values in water (Jou, 1985). The effective diffusiveness in the Hematocrit Layer is taken from measures in bio-films (Stewart, 1998).*

The hematocrit layer is modeled as a flat infinite film with an averaged homogeneous distribution of IRBCs. Only the limitations on the RBC viability resulting from the local scarcity of glucose are taken into account. In consequence, the diffusion problem is reduced to a problem in 1D: assessing the variations of the concentration of glucose in the direction normal to the hematocrit surface (z), $C(z,t) = C_{gluc}(\overrightarrow{x} = (\forall, \forall, z); t)$.

Two models of the averaged uptake have been considered i) a fixed uptake rate (zero-order kinetics), and ii) a linear dependence of the uptake with the concentration of glucose until reaching a maximum value (first-order kinetics).

Diffusion limitation for a 1D model with zero-order kinetics

Zero-order kinetics states that the uptake in Equation 3.9 is a constant value: $U(C(\overrightarrow{x},t)) = \overline{U}_{HL}$. The averaged uptake of a culture with $\%I \sim 0.5\%$ is $\overline{U}_{HL} =$

$4.7 \cdot 10^{-7} mol \cdot l_{RBCs}^{-1} \cdot s^{-1}$ (Rapoport et al., 1976). The effective volume available for diffusion is the fraction of volume not occupied by RBCs ($\epsilon_w = 1 - \overline{pf} = 0.15$). The relevant characteristics of the culture system are HLD and the bulk concentration of substrate in the culturing medium ($C_0 = 2.67\ mM/l$). Equation 3.9 can be rewritten in the following adimensional form:

$$\frac{\partial \widetilde{C}(\zeta,\tau)}{\partial \tau} = \frac{\partial^2 \widetilde{C}(\zeta,\tau)}{\partial \zeta^2} - \psi^2 \frac{\overline{U}_{HL}}{C_0} \tag{3.10}$$

where $\tau = \frac{t \cdot D_{eff}}{(1-pf) \cdot HLD^2}$ and $\zeta = \frac{z}{HLD}$ are adimensional time and length, respectively; $\widetilde{C}(\zeta,\tau) = \frac{C(\zeta,\tau)}{C_0}$ is the instantaneous substrate concentration profile, normalized to the concentration at the free culturing medium; and $\psi^2 = \frac{(1-pf) \cdot HLD^2}{D_{eff}}$ is a scaling factor.

The contour conditions for Equation 3.10 are:

i) $C(1) = C_0$, meaning that the concentration in the top of the HL is equal to the bulk concentration. This is actually a super-estimation of its real value, because a diffusion boundary layer (DBL) should be considered above the top of the HL.

ii) $C'(0) = 0$, meaning that there is no flux in the bottom of the HL because there is a wall at the bottom of the culturing system.

A maximum threshold to the concentration profile is obtained by considering diffusion alone. If the uptake is neglected ($\overline{U}_{HL} = 0$), Equation 3.10 with the contour conditions (i) and (ii) has an analytical solution, which is easily obtained with the Laplace transform ($\widetilde{c}(\zeta,s) = \mathcal{L}\left[\widetilde{C}(\zeta,t)\right]$). The solution of this algebraic equation in the frequency domain (s) is (Stewart, 1996):

$$\widetilde{c}(\zeta,s) = \frac{cosh(\sqrt{\frac{s}{D_{eff}}} \cdot \zeta)}{s \cdot cosh(\sqrt{\frac{s}{D_{eff}}})} \tag{3.11}$$

The corresponding solution in the temporal domain is given by an infinite series of trigonometric terms, with every term multiplied by the factor $e^{-\frac{t}{\psi^2}}$ that tends to zero when $t \to \infty$ (Crank, 1990). Such a solution corresponds to a substrate profile that graphically resembles a complementary error function with a variance that increases with t (see Figure 3.13). In the limit $t \to \infty$, the concentration profile is a homogeneous distribution $\widetilde{C}(\zeta,\infty) = 1,\ \forall\zeta$.

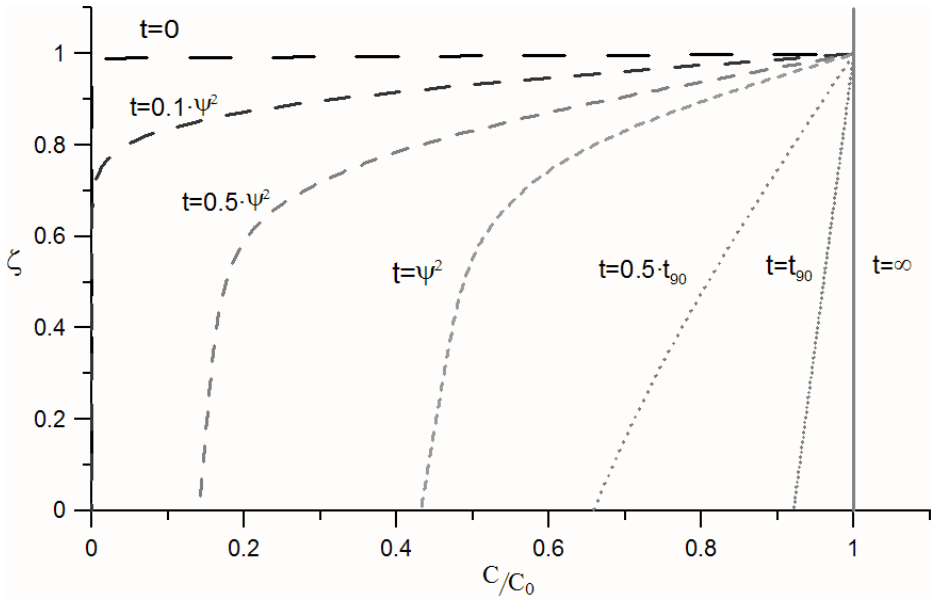

Figure 3.13: *Temporal succession of the normalized vertical profiles of glucose concentration ($C(\zeta,t)$) through the levels of the hematocrit layer (ζ). Initially ($t=0$), $C(\zeta,0)=C_0$ in the bulk culture medium ($\zeta>1$) and $C(\zeta,0)=0$ in the hematocrit layer ($\zeta<1$). The bulk concentration is maintained $C(1,t)=C_0$, and glucose accumulates at $\zeta=0$, $\frac{\partial C(0,t)}{\partial t}=0$. Diffusion makes the substrate penetrate the layer until reaching a uniform concentration $C(\zeta,\infty)=C_0$; ψ^2: characteristic temporal scale of the diffusion process; $t_{90}=7\psi^2$, when $C(0,t_{90})\simeq 0.9\cdot C_0$.*

The characteristic temporal scale of the diffusion process is defined by the scaling factorψ^2. A typical measurement of this characteristic scale is $t_{90}=7\psi^2$, which stands for the time when the concentration of a substances at the bottom of the cultured layer reaches $\sim$ 90% of its value at the bulk medium, as a consequence of diffusion alone, $\widetilde{C}(0,t_{90})=0.9$ (Stewart, 2003). In the limit $t\to\infty$, the concentration profile reaches a steady state that does not vary with time. The substrate concentration profile of the steady state ($C_{st}(\zeta)$) is reached when $t>t_{90}$, and is given by:

$$C_{st}(\zeta)=C_0-\frac{\overline{U}_{HL}\cdot HLD^2}{2\cdot D_{eff}}\cdot(1-\zeta^2) \tag{3.12}$$

The static cultures studied in this work (*P-, W-* and B-series) reach the steady state before the subcultivations take place (see Table 3.7). This means that the steady solution $C_{st}(\zeta)$ can be considered a good approximation of the concentration profiles in the hematocrit layer.

The characteristic spatial scale of the zero-order reaction-diffusion model (Eq. 3.10) in the steady state ($t > t_{90}$) is given by the depletion depth (Stewart, 2003):

$$d_0 = \sqrt{\frac{2 \cdot D_{eff} \cdot C}{\overline{U}_{HL}}} \tag{3.13}$$

Once the steady state has been reached, d_0 represents the level (position in the z direction) below which the concentration of substrate is zero. In the *P*-, *W*- and *B*-series: $C = C_0 = 2.67 \cdot 10^{-3} mols/l$, $D_e = 2.2 \cdot 10^{-6} cm^2/s$. According to the zero-order approximation, $\overline{U}_{HL} = 4.7 \cdot 10^{-7} mol \cdot l^{-1}_{RBCs} \cdot s^{-1}$. As a result: $d_0 = 1.6\ mm$.

Other characteristic scales may be useful for the current analysis: d_{RBC} and d_{IRBC}. They represent the levels of the *HL* that can hold RBCs and IRBCs, respectively. These levels are placed above the depletion depth and fulfill: $d_{RBC} > d_{IRBC}$.

These characteristic levels define three different sub-regions in the hematocrit layer:

1) $z > HLD - d_{IRBC}$, where both RBCs and IRBCs have enough substrate to fulfill their metabolic needs in the stationary state.

2) $HLD - d_{RBC} > z > HLD - d_{IRBC}$, where RBCs have enough substrate but IRBCs don't, in the steady state.

3) $HLD - d_{RBC} > z$, where there is not enough available nutrient to fulfill the RBCs metabolic needs in the stationary state.

The growth ratio measured for a culture system is the average of the contributions of sub-regions (1), (2) and (3). With regards to the overall observed performance of the culture system, these three sub-regions define three expected behaviors.

i) $GR(HLD)$ is maintained to a maximum value, for those cultures where both healthy and infected RBCs are viable. $GR(HLD) = GR_{max}$ as long as $HLD < d_{RBC}$.

ii) $GR(HLD)$ decreases with HLD in those cultures that allow two different sub-regions for the viability of the IRBCs, as long as the ratio of the volumes of both regions ($\frac{V(1)}{V(2)}$) is varied. $GR(HLD) = GR_{max} \cdot \frac{d_{IRBC}}{HLD}$ while $d_{RBC} > HLD > d_{IRBC}$.

iii) $GR(HLD)$ is maintained to a minimum value when the thickness of the culture exceeds the region of RBC viability. $GR(HLD) = GR_{min} = GR_{max} \cdot \frac{d_{IRBC}}{d_{RBC}}$ as long as $HLD > d_{RBC}$

The values $d_{RBC}$and d_{IRBC} can not be estimated with the current approximation, because the model does not consider different uptake regimes for RBCs and IRBCs, but rather an averaged overall uptake rate.

Diffusion limitations with first-order kinetics

First-order kinetics states that the uptake in Equation 3.9 is described with the following function:

$$U(C(z,t)) = \begin{cases} U_{max} & ;\ if\ C < C_{max} \\ \frac{K_{eff}}{pf} \cdot C(z,t) & ;\ if\ C < C_{max} \end{cases} \tag{3.14}$$

The values for the maximum uptake rate are $U_{max}(RBCs) = 4.7 \cdot 10^{-7} mol \cdot l_{RBCs}^{-1} \cdot s^{-1}$ and $U_{max}(IRBCs) = 4.7 \cdot 10^{-5} mol \cdot l_{RBCs}^{-1} \cdot s^{-1}$, respectively (see Equation 2.3).

Two measurements taken from literature are used to define the parameters in this model: the kinetic constant for the glucose slow uptake phase $K_{eff} = 0.016\ min^{-1}$ and the saturation concentration, at which RBCs uptake at their maximum rate $C_{max} = 40\ mM$ (Leitch and Carruthers, 2007).

Equation 3.14 allows defining the threshold glucose concentrations below which the uptake requisites of healthy RBCs and IRBCs are not fulfilled. These concentrations (C_V) determine the viability of healthy and infected cells. Their values are $C_V(RBC) = 2.4 \cdot 10^{-5} mols/l$ and $C_V(IRBC) = 2.4 \cdot 10^{-3} mols/l$, respectively.

The values $C_V(RBC)$ and $C_V(IRBC)$ can be introduced in Equation 3.13, assuming $C = C_0 - C_V$, to obtain a coarse estimation of the threshold depths for cell viability: $d_{RBC} = 1.6\ mm$ and $d_{IRBC} = 0.5\ mm$.

The assumption of constant depletion depths (independent from the HLD of the culture) constitutes the so called K-model. The decrease in the culture performance ($GR_{48}(HLD)$) with HLD predicted by K-model is depicted in Figure 3.15 and compared with the predictions of other diffusion models described below.

Optimal HL depth estimated with the Thiele modulus

Splitting the *HL* into three sub-regions with fixed thicknesses is indeed a coarse approximation. In real systems, the uptake rate (reaction term in Equation 3.9) explicitly depends on the concentration profile ($C(z,t)$), which even in the steady state is affected by HLD (see Equation 3.12). Consequently, the values $d_{RBC}$and d_{IRBC} must be affected by *HLD* and can not be taken for constants. An alternative approach is required.

An adimensional number usually employed in the study of diffusion-reaction processes is the Thiele modulus (ϕ), defined as the ratio between the reaction rate and the diffusion rate. The Thiele modulus allows for the distinction between two regimes in the reaction-diffusion models. If ϕ is small, diffusion is fast compared to reaction, but when it is large ($\phi \gg 1$), diffusion is slow (Stewart, 1996).

Assuming zero-order kinetics in a flat thin layer, the Thiele modulus ϕ_0 and the normalized substrate concentration profile achieved once the reaction diffusion reaches the steady state $\widetilde{C^0_{st}}(\zeta)$ are defined as:

$$\phi_0 = \sqrt{\frac{pf \cdot \overline{U_{HL}} \cdot HLD^2}{D_{eff} \cdot C_0}} \quad ; \quad \widetilde{C^0_{st}}(\zeta) = \begin{cases} 0 & ; \; if \; 0 \leq \zeta \leq (1 - \frac{1}{\phi_0}) \\ (1 - \phi_0 - \phi_0 \zeta)^2 & ; \; if \; (1 - \frac{1}{\phi_0}) \leq \zeta \leq 1 \end{cases} \tag{3.15}$$

For the first-order kinetics, ϕ_1 and the corresponding $\widetilde{C^1_{st}}(\zeta)$ are defined as:

$$\phi_1 = \sqrt{\frac{(1-pf) \cdot K_{eff} \cdot HLD^2}{D_{eff}}} \quad ; \quad \widetilde{C^1_{st}}(\zeta) = \frac{cosh(\phi_1 \cdot \zeta)}{cosh(\phi_1)} \tag{3.16}$$

The calculated values of ϕ_0 and ϕ_1 for the different HLD are listed in Table 3.7. The vertical profiles obtained for different values of the Thiele modulus are depicted in Figure 3.14.

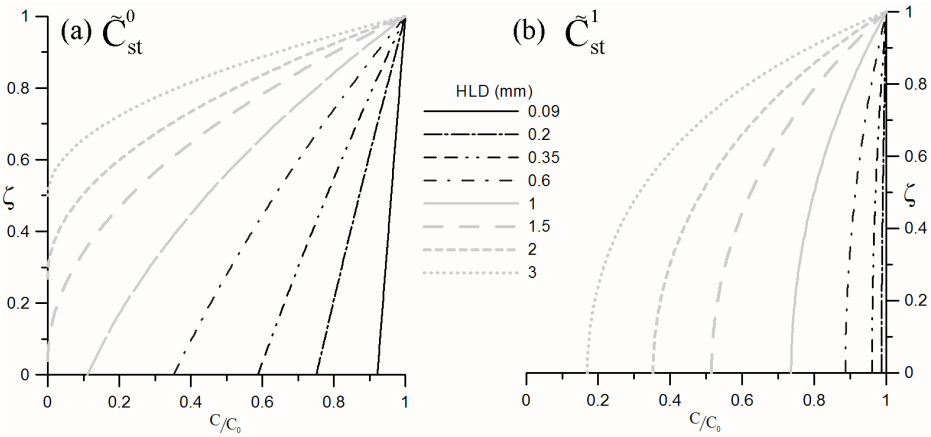

Figure 3.14: *Expected normalized glucose concentration profiles at the steady state* $\widetilde{C}_{st}(\zeta) = \frac{C}{C_0}$ *for different hematocrit layer depths (HLD). a) Predictions of the zero-order kinetics model (see Equation 3.15); b) predictions of the first-order kinetics model (see Equation 3.16)*

The depletion depths d_{RBC} and d_{IRBC} can be calculated using these profiles and the estimated threshold concentrations of cell viability ($C_V(RBC)$ and $C_V(IRBC)$). They are listed in Table 3.7. As expected, the values for d_0, d_{RBC} and d_{IRBC} decrease with HLD: more cells entail a more severe exhaustion of the medium.

The concentration gradients obtained assuming zero-order kinetics (0-model) are greater than those obtained assuming first-order kinetics (1-model) because the former fixes a constant consumption of nutrient within the viability region, while the uptake continually decreases with nutrient concentration under first-order kinetics.

No measurements of the concentration profiles have been carried out for the experimental culture systems, so it is not possible to determine which model provides a better approximation to reality. However, the soundness of these models can be indirectly determined by comparing the expected decrease in the culture performance ($GR_{48}(HLD)$) with HLD predicted by each model.

Expected infection growth rates

The threshold depths for RBC and IRBC viability (d_{RBC} and d_{IRBC}) obtained with the K-model, 0-model and 1-model can be used to determine the values of the key parameter h of the HLD-model and the WS-model (see Equations 3.7 and 3.8):

$$h = \begin{cases} HLD & ;\ if\ HLD < d_{IRBC} \\ d_{IRBC} & ;\ if\ d_{IRBC} < HLD < d_{RBC} \\ HLD \cdot \frac{d_{IRBC}}{d_{RBC}} & ;\ if\ d_{RBC} < HLD \end{cases} \tag{3.17}$$

The expected growth ratios $GR_{48}(HLD)$ obtained when the value h is defined using Equation 3.17 are depicted in Figure 3.15, together with the experimental values and the best-fit model.

The obtained results show that models considering just one dimensional diffusion limitations qualitatively follow the trends observed in the experimental hematocrit layers, but still fail to reproduce the behavior of real systems. Apart from the decay for very thin HLs ($HLD < 0.1\ mm$) (which can be explained as an emergent behavior of merozoite propagation, see Section 3.4), there are two main differences between the predicted and observed $GR_{48}(HLD)$: firstly, the decrease with HLD is greater in the diffusion models than in real systems. Secondly, for thick HLs ($HLD > d_{RBC}$), the slope of GR_{48} is flat, while in the experimental systems and in the best-fit WS-model, $\frac{dGR_{48}}{dHLD} < 0$ in all the explored domain.

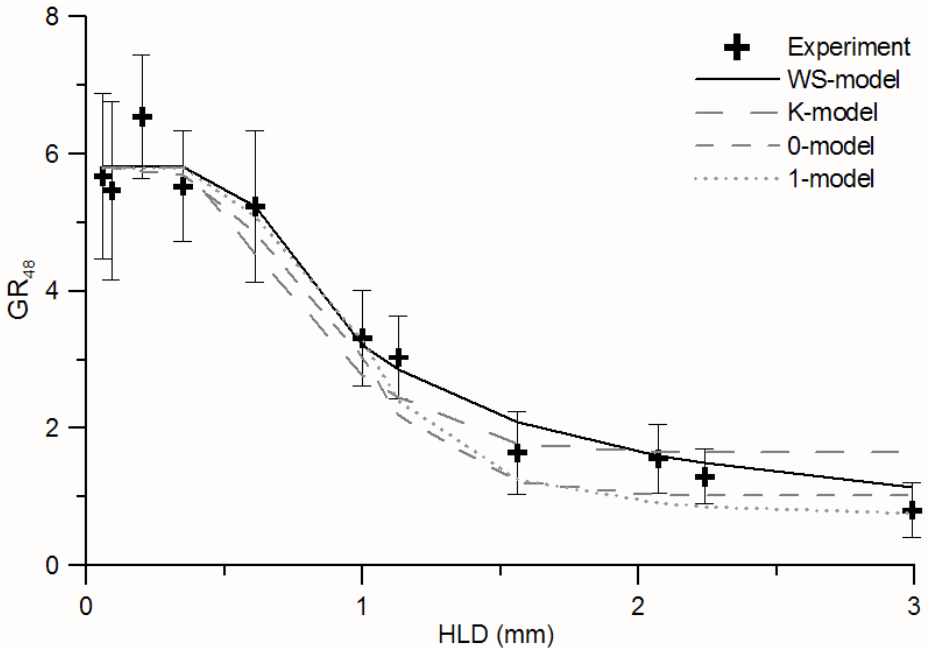

Figure 3.15: *Average infection growth ratio (GR_{48}) as a function of the hematocrit layer depth (HLD). Crosses (+): experimental results; black solid line: best-fit HLD-model (see Section 3.4); long-dashed line: results obtained with the K-model; short-dashed line: results obtained with the 0-model, dotted line: results obtained with the 1-model.*

The first difference can be explained as the result of considering a constant uptake for a RBC population that is not affected by nutrient availability. The models here presented can not take into account mortality due to the scarcity of substrate. In real systems the average uptake will be drastically reduced below the depletion boundary. As a consequence, the thickness of the layer with no substrate limitations (h) will be increased.

The second difference, the asymptotic $GR_{48}(HLD)$ is a consequence of the definition of h for thick HLs (see Equation 3.17). It is observed in the K-model and 0-model alone, while the 1-model does not show this behavior because all the trialled HL fulfill $HLD < d_{RBC}$. Nevertheless, there is a slight difference in how the K-model and the 0-model define the extension of the sub-region with high infection propagation: in the K-model $d_{RBC}$and d_{IRBC} are fixed values while in the 0-model both distances vary with HLD ($d = d(HLD)$) and what remains almost fixed is the ratio $\frac{d_{IRBC}}{d_{RBC}}$.

Another key difference between the WS-model and the diffusion models presented in

the current section lies in the fact that the former defines two sub-regions in the HL with different infection propagation rates, but parasites can still spread in both regions. In contrast, the sub-regions defined by the diffusion models here presented consist of a domain where infection spreads and another where it does not.

		Zero order kinetics				**First order kinetics**			
HLD (mm)	t_{90} (h)	ϕ_0	d_0	d_{RBC} (mm)	d_{IRBC}	ϕ_1	d_0	d_{RBC} (mm)	d_{IRBC}
0.06	18 s	0.04	NA	NA	NA	0.05	NA	NA	NA
0.09	45 s	0.06	NA	NA	NA	0.07	NA	NA	NA
0.2	4 min	0.13	NA	NA	NA	0.16	NA	NA	NA
0.4	10 min	0.23	NA	NA	NA	0.29	NA	NA	NA
0.6	1 h	0.41	NA	NA	0.5	0.5	NA	NA	NA
1.0	1.5	0.67	NA	NA	0.48	0.82	NA	NA	0.48
1.1	1.9	0.75	NA	NA	0.21	0.93	NA	NA	0.21
1.5	3.6	1.04	NA	NA	0.18	1.28	NA	NA	0.19
2.0	6.4	1.38	1.2	1.3	0.13	1.7	NA	NA	0.14
2.2	7.4	1.49	1.1	0.9	0.09	1.84	NA	NA	0.14
2.9	13.2	1.99	0.7	0.6	0.06	2.46	NA	NA	0.07

Table 3.7: *Characteristic parameters of the reaction-diffusion balance in the hematocrit layer as a function of its depth (HLD). Characteristic diffusion times ($t_{90} \simeq 7\psi^2$). Thiele modulus (ϕ) and characteristic lengths for zero order and first order kinetics: depletion depth (d_0) with substrate concentration $C(d_0) = 0$; threshold depth for RBC viability (d_{RBC}) with $C(d_{RBC}) = 2.4 \cdot 10^{-5} mols/l$; and threshold depth for IRBC viability (d_{IRBC}) with $C(d_{IRBC}) = 2.4 \cdot 10^{-3} mols/l$. NA stands for non-applicable, meaning that the threshold concentration values are not reached.*

3.5.2 Explicit model of the diffusion process

The individual-based approach deals with the reaction-diffusion problem stated in Equation 3.9 accounting for $U(C(\overrightarrow{x}, t)$ with the model of the individual cells and solving diffusion through the hematocrit layer with the model of the local environment.

The simplest numerical method to model diffusion through a discrete spatial grid is the explicit approach Forward-Time- Centered-Space (FTCS, see Appendix C). Diffusion is solved iteratively at each time step by converting the spatial derivatives at each point of the grid ($\overrightarrow{x}$) to finite differences, and by converting the temporal derivatives to finite increments to be transferred at the end of the time step ($t \rightarrow t + 1$).

This method is used by INDISIM (see Equations 2.5 and 3.4 in 2 and 3 dimensions,

respectively), and has the general form

$$C_{\overrightarrow{x}}^{t+1} = C_{\overrightarrow{x}}^{t} + \widetilde{D} \sum_{i \in nn(\overrightarrow{x})} w_i C_i^t \tag{3.18}$$

where i is any of the spatial cells in the surroundings of $\overrightarrow{x}$, $nn(\overrightarrow{x})$, w_i is the weight of each contribution, and $\widetilde{D}$ is the numerical diffusion coefficient, which can be related to the real experimental coefficient (D_{eff}) with a correction factor (cf) that depends on the discretization characteristic scales (spatial $l_{sc}$and temporal ts) of the model:

$$\widetilde{D} = cf \cdot D_{eff} = \frac{ts}{l_{sc}^2} \cdot D_{eff}$$

The numerical analysis of the method (Eq. 3.18) in one dimension reveals that the characteristic scales must fulfill $\widetilde{D} < \frac{1}{2}$ to maintain the numerical stability of the solution (Hoffmann and Chiang, 2004). This means that $ts < \frac{l_{sc}^2}{2 \cdot D_{eff}}$, otherwise numerical errors propagate and multiply through the spatial grid, completely altering the simulation outcome. For models in 2D and 3D, this constraint is $ts \leq \frac{l_{sc}^2}{D_{eff}} \cdot (1 - max\,(w)) < \frac{l_{sc}^2}{D_{eff}}$, where $max\,(w)$ is the maximum value of the diffusion weights as they appear in Equations 2.6 and 3.5 respectively.

When the size of the spatial cell is set to the scale of the RBC ($l_{sc} \sim 5\ \mu m$), the time step must fulfill $ts < \frac{l_{sc}^2}{D_{eff}} = \frac{10^{-7} cm^2}{10^{-6} cm^2/s} = 0.1\,s$. This time step is too small to simulate the long-term evolution of cultures (i.e. $48\ h \sim 10^7 ts$). Choosing this time step would also entail that some of the cellular processes that are now considered with simple time-averaged models (such as cellular metabolism and uptake) would require more elaborate time-dependent descriptions.

Alternatively, the time step can be set to the current value ($ts = 6$ minutes, which is small enough to correctly account for merozoite propagation through the HL and big enough to use simple average models for the RBC metabolism), and the size of the spatial cell can be fixed to fulfill numerical stability. In this case $l_{sc} > \sqrt{\frac{ts}{1 - max\,(w)} \cdot D_{eff}} = 0.02\ cm$. Adopting this scale would entail that each spatial cell contains $\sim 10^4$ RBCs. Consequently, many of the current models (such as the release and spreading of merozoites) wouldn't be correctly implemented.

The FTCS approach forces a compromise regarding the characteristic scales of the model, between representing the spreading of the parasite (and a coarse model of the RBC) and the substrate diffusive limitations (and a detailed model of the RBC). The solution employed in INDISIM-RBC so far ($l_{sc} \sim 10\ \mu m$, $ts \sim 6$ minutes) opts for the former, and banks on the spreading of the parasite as the limiting factor to the

propagation of the infection. The numerical diffusion coefficient corresponding to this choice ($\widetilde{D} = \frac{ts}{l_{sc}^2} \cdot D_{eff} \sim \frac{10^2 s}{10^{-6} cm^2} \cdot 10^{-6} cm^2/s \sim 10^2$) leads to the numerical instability of the diffusion model.

So, the values of $\widetilde{D}$ and the algorithm employed in the INDISIM-RBC simulations either overestimate the restrictions to diffusion whenever $\widetilde{D} < 1$, or do not consider diffusion limitations at all when $\widetilde{D} = 1$. The diffusion limitations observed in the simulations with arbitrary values of $\widetilde{D}$ (for instance, $\widetilde{D} = 0.8$ in Figure 2.15b, in Section 2.3.5) correspond to excessively small values of the real D_{eff}.

Yet, the analysis of the reaction diffusion problem obtained in section 3.5.1 states that diffusion is indeed a limiting factor. FTCS simulations of diffusion alone carried out in an empty spatial grid using appropriate scales ($l_{sc} \sim 1\ \mu m$ and $ts \sim 1\ ms$, $\widetilde{D} \approx 0.3$) show that there are differences in the concentration of glucose in the scales employed by INDISIM-RBC (after 6 minutes, the relative difference in concentration observed at an RBC distance is $\delta = \frac{\triangle C(L=10\ \mu m; t=6\ minutes)}{\triangle C(L=10\ \mu m; 0)} \sim 10^{-3}$). This is the amount of nutrient that is not transported during a time step, so it can be assumed that $\widetilde{D} = 1 - \delta$.

This result is consistent with the application of values of $\widetilde{D} < 1$ above mentioned. The problem when trying to set the value for $\widetilde{D}$ in INDISIM-RBC is that the value calculated using smaller fractions of the spatial cell and time steps is not a scale invariant. In other words, that $\widetilde{D}$ varies with the resolution of the explicit simulation. The discussion above indicates that diffusion plays a limiting role in the characteristic scales of the modeled system, even though our explicit models cannot account for it satisfactorily.

3.5.3 Implicit model of the diffusion process

An alternative to the explicit resolution of any numerical problem is to use an implicit method. Implicit methods give the future configuration of the system by solving an equation that involves both the present and future states of the system.

The implicit Crank-Nicholson (C-N) method (Hoffmann and Chiang, 2004) can be employed to numerically solve diffusion within the hematocrit layer with a 1D model and the discretization used by INDISIM-RBC.

The reaction-diffusion process described by Equation 3.10 can be expressed with a finite difference approximation. Discretization of $\widetilde{C}(\zeta, \tau)$ leads to a grid containing N sites ($6 < N < 290$ to cover the range of observed HLDs using the scales $l_{sc} \sim 10\ \mu m$ and $ts \sim 6$ minutes), each one characterized by its instantaneous concentration $\widetilde{C}_i^{\tau}$. Equation 3.10 results in a linear system containing N coupled equations, one for each spatial cell i, to be solved at each time step (τ).

The explicit method described in the previous section (Equation 3.18) takes the form

of Equation 3.19. For each equation i, the FTCS scheme assumes one unknown variable ($\widetilde{C}_i^{\tau+1}$) to be determined through a linear combination of known values at neighboring sites $\widetilde{C}_i^{\tau}$, $\widetilde{C}_{i+1}^{\tau}$ and $\widetilde{C}_{i-1}^{\tau}$.

$$\frac{\widetilde{C}_i^{\tau+1} - \widetilde{C}_i^{\tau}}{\triangle\tau} = \widetilde{D}\frac{\widetilde{C}_{i+1}^{\tau} - 2\cdot\widetilde{C}_i^{\tau} + \widetilde{C}_{i-1}^{\tau}}{(\triangle\zeta)^2} - U_i^{\tau} \qquad ; \quad i = 1:N \tag{3.19}$$

Alternatively, Equation 3.10 can be represented by the set of N linear equations that constitute the implicit C-N approach. This results in the following system of equations, one for each site i:

$$\frac{\widetilde{C}_i^{\tau+1} - \widetilde{C}_i^{\tau}}{\triangle\tau} = \widetilde{D}\frac{\widetilde{C}_{i+1}^{\tau+1} - 2\cdot\widetilde{C}_i^{\tau+1} + \widetilde{C}_{i-1}^{\tau+1}}{(\triangle\zeta)^2} - U_i^{\tau} \qquad ; \quad i = 1:N \tag{3.20}$$

In this case, each equation contains three unknown values ($\widetilde{C}_i^{\tau+1}$, $\widetilde{C}_{i-1}^{\tau+1}$ and $\widetilde{C}_{i+1}^{\tau+1}$) regarding the future concentration at the present and neighboring sites and only two known values ($\widetilde{C}_i^{\tau}$ and U_i) that include the information regarding the present state of site i. The C-N scheme can be expressed in the generalized form:

$$a_i\cdot\widetilde{C}_{i-1}^{\tau+1} + b_i\cdot\widetilde{C}_i^{\tau+1} + c_i\cdot\widetilde{C}_{i+1}^{\tau+1} = D_i^{\tau}$$

where D_i^{τ} is a linear combination of $\widetilde{C}_i^{\tau}$ and U_i. This set of equations, together with boundary conditions (i) $\widetilde{C}_N^{\tau} = 1$, and (ii) $\frac{\partial\widetilde{C}_0^{\tau}}{\partial t} = 0$, stated in Section 3.5.1, can be expressed in the matrix form:

$$\begin{pmatrix} b_2 & c_2 & & & \\ a_3 & b_3 & c_3 & & \\ & & \ddots & & \\ & & a_{N-3} & b_{N-2} & c_{N-1} \\ & & & a_{N-2} & b_{N-1} \end{pmatrix} \begin{pmatrix} C_2^{\tau+1} \\ C_3^{\tau+1} \\ \vdots \\ \\ C_2^{\tau+1} \end{pmatrix} = \begin{pmatrix} D_2^{\tau} \\ D_3^{\tau} \\ \vdots \\ \\ D_2^{\tau} \end{pmatrix} \qquad \Rightarrow \qquad C^{\tau+1}\cdot A = D^{\tau} \tag{3.21}$$

Matrix A in Equation 3.21 is determined and tridiagonal. It can be inverted to determine $C_i^{\tau+1}$, $\forall i$. The results obtained with such an approach are consistent with the analytical results presented in Section 3.5.1.

This implicit formulation of the reaction-diffusion problem can be worked out through different numerical methods. The main advantage of such type of algorithms is that they are numerically stable regardless of the size of the spatial and temporal scales. These

methods can also be implemented in INDISIM-RBC in order to solve the diffusion of substances. Their most important drawbacks are i) that the inversion and solution of matrix A requires a lot of computational effort, and ii) that the differences in the individual uptake (and the consequent heterogeneity in substrate concentration appearing in the hematocrit layer) can lead to ill conditioned matrices, not amenable to to computational solving.

Implementing C-N in INDISIM-RBC entails extending the presented formulation to 2D and 3D, and dealing with band diagonal matrices nn number of terms per row (the total number of spatial cells that constitute the nearest neighborhood of cell i). Matrix A in 2D contains $nn = 5$ terms per row when only side neighbors are taken into account, and $nn = 9$ when diagonal nearest neighbors are also included. Matrix A in 3D can contain up to $nn = 27$ terms when all the first neighboring cells are considered. The number of operations to be carried out at each time step linearly increases with N and the computational time also increases with the number nn.

Built-in libraries and numerical packages prepared to solve diffusion problems using optimized implicit methods are available from several academic sources (BLAS: Basic-Linear-Algebra-Subprograms and LAPACK, among others). However, their inclusion in INDISIM-RBC resulted in excessively time-consuming simulations. The optimization of the computational methods to solve specific problems like the one here presented constitutes a research field on its own (Morton and Mayers, 1996) and its application to INDISIM-RBC is regarded as an option for further work.

3.6 Discussion and open questions

INDISIM-RBC.*v3D* improves the set of rules governing individuals (RBCs and IRBCs) and the environment (spatial structure, extracellular parasite and substrate concentration) to better account for local interactions and transport phenomena (Section 3.2). Three representational goals have been achieved with the 3D version of the model (Section 3.3):

1. *Population dynamics and structure:* The temporal evolution of parasitaemia ($\%I$), infection growth ratio (GR) and distribution of post-invasion times among IRBCs ($q\left(t_{INF}\right)$) observed in short-term and long-term culture systems are still quantitatively reproduced by the model.

2. *Spatial structure of the hematocrit layer:* the RBC packing factor within the hematocrit layer (pf) in the model is consistent with the cell density measured for the

settled hematocrit in experimental systems.

3. *Infection dynamics:* the distribution of IRBCs with (k) multiple invasions predicted by the model ($f_{sim}(k; P_{inf})$) is in accordance with the experimental observations ($f_{obs}(k)$).

Yet, the foremost advantage of this version is that it allows tackling static *in vitro* culture systems with different macroscopic geometries and comparing them with each other (Section 3.4). The study of different experimental setups shows that optimal harvests are obtained with the following ranges: depth of the hematocrit layer $0.2\ mm < HLD < 0.5\ mm$ and distance between walls of the culturing vial $L > 2\ cm$.

INDISIM-RBC.*v3D* alone does not provide a justification for these ranges. However, when additional top-down phenomenological laws are assumed, the behavior of real systems is accurately reproduced. These whole-system laws basically consist in splitting the hematocrit layer into different sub-regions with high/low infection proliferation rates. This leads to the decrease in GR for high values of the hematocrit layer depth (when $HLD > 0.48\ mm$) and defines a minimum threshold for the separation between walls. The reduction in GR observed for smaller cultures ($HLD \sim 0.1\ mm$) can be explained through individual-based limitations in the spreading of extracellular merozoites.

A continuous model for diffusion of substrate and waste products has been developed to support the phenomenological laws (Section 3.5). Further work in this line should include a better model for diffusion at the individual level, which can give rise to the limitations observed in real culture systems. Nevertheless, it must be kept in mind that the restrictions observed in experiments may be caused by other factors besides diffusion limitations, such as an unaccounted influences of the spatial structure of the hematocrit layer, merozoite spreading or RBC invasion.

Turning to another topic, INDISIM-RBC can be modified to account for cultures in suspension. In such models, the invasion of healthy RBCs differs from the one in static cultures. The development of a model that can correctly describe both static and agitated culture systems would be of great interest in order to design an operative and reliable bio-reactor-like culturing device.

Chapter 4

Individual-based Models of microbial communities

The research in modeling of *in vitro* cultures of *Plasmodium falciparum*-infected erythrocytes has been carried out by MOSIMBIO in parallel with the individual-based modeling of other microbial communities. The diversity in the scopes of application of INDISIM offers an opportunity to analyze the general characteristics of the methodology and to review and formalize the procedures followed when trying to develop new models, alongside with the specific results obtained for each system under study.

This chapter outlines some reflections on the work carried out by the group MOSIMBIO that result from applying the same methodology to tackle different microbial systems, in parallel. This general discussion follows the more specific results previously shown, following a bottom-up strategy, because the tools and results presented in previous chapters are required for building a more general framework.

Firstly, some practical considerations are presented regarding the use of IbMs to tackle microbial systems (Section 4.1). Secondly, the approach based on individuals is compared to the models defined at a system level using an analysis of the strengths and weaknesses of each kind of approach (Section 4.2). Finally, some general principles underlying the characterization of microbial systems are proposed following the argumentation of bottom-up approaches. Special attention is placed on the contribution of IbM to the study of variable and stationary states in microbial population dynamics, in general (Section 4.3).

The application of the model to different kinds of microbial systems (bacteria, yeast and protozoa) is put together to illustrate this theoretical discussion, and its particular application to malaria in *in vitro* cultures is presented as an exemplifying closing.

4.1 Pattern-oriented modeling

Modeling, the act of representing and interpreting reality in order to anticipate phenomena, is one of the most universal (if often unconscious) human activities. Models are also the principal instruments in science. *Grosso modo,* research either consists in building, testing, comparing and revising models, or in applying models and interpreting their predictions.

Models can carry out two representational functions. On the one hand, a model can describe a selected part of the world. The emphasis of the model is on the reproduction of observed patterns, the best fit to measured data, and the sensitivity of the model outcome to variations in the model parameters. This type of model is called empirical, or data-driven, model. The purpose of such models is to gain predictive capability concerning the modeled system.

On the other hand, a model can represent a theory (or part of it), in expressing the application of the laws and axioms of that theory to specific systems. The accent is on the structure, mechanisms and rules proposed by the model, and on assessing to what extent these are consistent with the real-world observations. This type of model is called mechanistic, or theoretical, models and its purpose is to increase understanding of the modeled system. These two aspects of modeling are not mutually exclusive, and scientific models usually are used in both senses (Silvert, 2001).

4.1.1 Pattern-oriented strategy to build models in microbiology

The central concern of the predictive function of models is to determine the values of the parameters that best fit with the experimental observations (Bernaerts et al 2000) and to statistically verify and validate the applicability of models in particular situations (Giffel and Zweitering, 1999). These problems have recently been addressed in detail by the group MOSIMBIO (Gras, 2004; Prats, 2008).

As for understanding purposes, the accent in models is placed on quite different issues. The keystone of mechanistic modeling is to select which features of reality are considered as objects (entities and variables) and as laws (equations, algorithms and constraints on the variables) to optimally represent a set of patterns. In other words, to set the structure and degree of complexity of the model that unravels the underlying causes of the behavior observed in real systems.

The factors that influence population dynamics of a microbial system operate at various spatial and temporal scales, and it might not be straightforward to aggregate all them into a single model. The coupling of multiple scales plays a significant role in microbial

dynamics, in particular, and the account of a factor at any scaled is usually critical to the model outcome.

Deciding which factors must be included in a model is no trivial issue and implementing the decisions as a set of rules is not an easy task. A strategy to grapple with this problem is to set the structure of the model by proposing explanations to multiple specific patterns observed in the real system. This strategy is called **Pattern-oriented modeling** (PoM) and is commonly employed in the context of ecological modeling (Grimm et al., 1996).

PoM entails using observations of the real system at multiple levels of description as the guideline for building a model. The key point is that the *only* information to be included in the model is that which serves to reproduce and interpret a given pattern. Thus, the objective of the model is no longer to integrate all the information regarding the real system, but to explain as many observed patterns as possible with the least number of *ad hoc* assumptions. By assuming this strategy, one should be able to avoid building either excessively comprehensive and detailed models that are difficult to analyze, or models that are too naive and fail to capture essential features of the system.

The idea underlying this strategy is that when the system under study is complex enough, any model will "inevitably either leave out relevant information or become over-parametrized and lose predictive power" (Grimm et al., 2005). PoM is a systematic plan to optimize the payoff of modeling: the model must be complex enough to reproduce and predict many of the patterns observed in reality, but simple enough to be grasped and applied. It is in this Medawar zone (not too naive, nor too complex) where models produce fruitful results.

4.1.2 Simplicity and specification in PoM

The law of parsimony (the simplest is the best) is as valid when modeling microbial systems as in any other field. Yet, the simplest model shouldn't be mistaken for the model containing the fewest variables, parameters or rules. The rules included in the model should be theoretically sound (and more complicated) rather than phenomenological (simpler but opaque). The strategy for guaranteeing the mechanistic nature of the proposed rules is to address multiple patterns at different levels of description, simultaneously, with the same set of rules.

Each pattern observed in the real system means an additional constraint on the model. A rule that explains many unrelated patterns is more likely to be a real mechanism of the system under study, or at least it will act as if it was so. When a great number of patterns are reproduced with a minimal assumption of rules (without increasing the complexity of the model), the confidence in the model is reinforced. In this sense, PoM provides a

criterion to account for the degree of mechanisticity of a model.

Mechanistically rich models offer several advantages over more empirical models. First, they are well-grounded and more transparent than simpler models, in which the information is put together into parameters or variables that can not easily be related to measurable data. Second, they are more complete: they can be applied to the study of different aspects of the same problem. And last but not least, they are more versatile: they can easily be modified to introduce new features of reality, or to address problems similar to those modeled. In this sense, PoM provides robustness and credibility to the model.

An advantage of models built using a pattern-oriented strategy is their robustness: the simultaneous reproduction of multiple patterns deters the variation of the model. Any eventual introduction or modification of a rule defined by the model should at least preserve, if not increase, the number of patterns reproduced. An additional advantage of PoM is that it requires the explicit description of reality at multiple levels. The model is forced to consider and include the characteristic scales of the system under study. Simulation units are then unequivocally related to physical magnitudes, and unrealistic outcomes are easily detected. In this sense, PoM facilitates the development of models and helps detect inconsistencies.

In conclusion, multiple patterns help in building realistic models that are not too simple or too complex. Understanding of the real system is gradually acquired when the model is tested, modified and corrected to interpret multiple patterns at different hierarchical levels.

4.1.3 Application of PoM to INDISIM-RBC

Three versions of INDISIM-RBC have been presented in this manuscript, but many more intermediate or alternative versions have been tested (and improved or discarded) in the development of the current schemes. The application of PoM strategy in the development of INDISIM-RBC is presented next.

The development of INDISIM-RBC is the response to the specific demands of EMG-GSK. It is the result of struggling to apply a single model to reproduce different patterns and answer the specific questions posed by the experimentalists (see the independently observed phenomena included in Section 2.3.3), and performing the pertinent modifications when they were unavoidable. The development of each of the different versions of the program has followed from the procedure below, which is in accordance with the PoM strategy:

i) the information in literature and the heuristic knowledge of the experts are used to define the characteristic parameters and variables of the model (Sections 1.2.1, 1.2.2 and 1.3),

ii) the temporal evolution of the infection for different specific cultures is used as the main pattern to calibrate the parameters and to check the validity of the model (Sections 2.3.1, 2.3.2 and 3.3.1), and

iii) several independently observable patterns are used to check the consistency of the proposed rules and model structure, and to detect when the rules must be modified, and steps *i* and *ii* must be repeated (Sections 2.3.3, 2.3.4, 2.3.5, 2.3.6 and 3.2.1).

The list below shows some factors governing the infection dynamics of *P. falciparum* in erythrocytic *in vitro* cultures that are considered/omitted by INDISIM-RBC. They are sorted into categories that refer to model configuration, spatial/temporal structure and randomness.

1. **Model configuration** refers to the entities, variables and rules considered by the model and to how they are implemented into an algorithm to perform simulations.

   The model considers individual RBCs, merozoites and two sample metabolites, but leaves out most of the composites of the culturing medium. Individual traits and variables are related to motion, infection, viability and consumption of glucose, but the model does not account for energy trade-offs, cell growth or detailed metabolisms. Processes are implemented sequentially in the algorithm although data are updated in parallel when required.

2. **Spatial/temporal structure** refers to the spatial dimensions and temporal scales that are considered by the model.

   The model considers short-range (L$\sim$ micrometers) of biological interactions (e.g. uptake of nutrient, egress and spreading of extracellular merozoites) and physical interactions (e.g. diffusion of glucose and lactate). The model considers only a small fraction of the system (L$\sim$millimeters) and assumes it to be representative of the whole (L$\sim$ centimeters). The model has a temporal resolution of minutes.

3. **Randomness** refers to the diversity among the entities and to the stochasticity in the rules that are considered by the model.

   Diversity in the population, spatial heterogeneity and uncertainty in the processes occurring at an individual level (e.g. invasion of a healthy RBC by an extracellu-

lar merozoite, or accidental death) or at a collective level (e.g. variations in the subcultivation or medium renewal protocols) are key factors in the model outcome.

Many other important factors and processes are not included in the model because they are not essential to explain the patterns (e.g. temperature and heat dissipation, concentration and transport of other metabolites, genetic variability and phenotypic changes through the life course) that have not been taken into account.

Modifications of the program arise from observations of new patterns or when a version is found to be incapable of correctly representing a pattern that was considered relevant by the EMG-GSK. For instance, the explicit representation of the motion of extracellular merozoites forced the increase of the temporal resolution of the model from hours to minutes, and the lack of success in representing merozoite spreading resulted in the 3D version.

Finally, the current version of INDISIM-RBC is not a closed issue but rather a suitable platform open to improvement with further versions. For instance, the calibrated study of diffusion within the hematocrit layer revealed that the explicit formulation of the transport of glucose implemented in INDISIM-RBC was not appropriate to the operating scales and should be corrected in subsequent work.

4.2 IbM and PbM approaches in microbiology

SWOT (Strengths-Weaknesses-Opportunities-Threats) analysis is a strategic planning method used to evaluate a project prior to its implementation. It consists in identifying the internal factors (characteristics of the project) and external factors that are favorable and unfavorable to achieving a predefined objective. This results in a list of Strengths (favorable internal factors), Weaknesses (unfavorable internal factors), Opportunities (favorable external factors) and Threats (unfavorable external factors).

This analysis is adapted here to evaluate the perspectives of two current approaches: models built at an individual cell level (IbM) and the more usual continuous approach at a population level (PbM). In the analysis below, the internal factors are the intrinsic characteristics of each approach; Opportunities refer to the current and potential applications of each methodology to specific real systems and Threats describes the risks and limitations of each method. This consideration aims to be a tool for deciding which kind of approach is more appropriate to tackle a given problem. Tables 4.1 and 4.2 summarize the discussion below. Each point listed in the tables refers to symbols in the text, for instance PS1, PW1 stand for Strengths and Weaknesses of the population models, while IS1, IW1, etc. refer to Individual-based models. The disquisition below constitutes the

core of a manuscript published in the International Journal of Food Microbiology (Ferrer et al., 2009) that is the extension of a communication presented in the 21^{st} symposium of the International Committee of Food Microbiology and Hygiene (august 2008), for this reason, citations and examples mostly refer to topics in the ambit of food microbiology. A complementary discussion on this topic, together with a much broad list of references, has been recently published (Hellweger and Bucci, 2009).

Population-based Models (PbMs)		
S	PS1	Prevalent and widely accepted formalism
	PS2	Standardized methods to build, implement and test models
	PS3	Models are easily shared and communicated
	PS4	Rapid and effective simulations
	PS5	Feasible experimental measurements at the population level
W	PW1	Disregard individual variability and the local nature of interactions
	PW2	Unspecific translation between levels of description.
	PW3	Aggregated parameters might lack biological basis
	PW4	Extra parameters increase likelihood but not soundness
	PW5	Hardly integrate measurements at a cell level
O	PO1	Canonical models can be used in aggregated approaches
	PO2	Non-specific models can be exported to different real situations
	PO3	Synthetically incorporate new experimental information
	PO4	Provide criteria for analyzing and tracking real systems
	PO5	Can use information available in huge databases (i.e. ComBase)
T	PT1	Phenomenological models might be based on "wrong" biological basis
	PT2	Cell diversity and process uncertainty are joined as general randomness
	PT3	The scope of validity of the mechanistic assumptions is hardly determined
	PT4	Validation of the model does not entail recognition of the real mechanism

Table 4.1: *SWOT (Strengths-Weaknesses-Opportunities and Threats) analysis of the population-based approach.*

4.2.1 Population-based models

Most of the current models in quantitative microbial ecology are continuous models that describe population dynamics using differential equations or the relations found when solving them. The equation-based approach is the traditional way to quantify the functional relations between the characteristic variables of an observed system in science (PS1). This formalism simplifies and guides the raising, development and communication of the models because it provides a familiar and structured framework and because it restricts the complexity of models to the complexity of the equations. Continuous models

Individual-based Models (IbMs)		
S	IS1	Explicit and transparent use of randomness
	IS2	Explicitation of the biological basis of the defined rules
	IS3	Adjustable degree of generalization and complexity of the models
	IS4	Capacity to disentangle mechanisms
	IS5	Easily deal with small populations and measurements at a cell level
W	IW1	Lack of standard methodologies to guide model building and analysis
	IW2	Slow to implement and requiring higher computing capacity
	IW3	Data hungry models that require hardly available information
	IW4	Seldom use unavailable information to define individual parameters
	IW5	Multiscale systems are covered with difficulty
O	IO1	Check the consistency of the hypothesis of PbMs
	IO2	Estimate cell parameters that are often indirectly determined
	IO3	Provide new "statistical" indicators of the state of a microbial community
	IO4	Integrate information obtained from -omics research and bioinformatics
T	IT1	Synthetic understanding hardly extracted from the models
	IT2	Lack of standards hinders the use and development of the approach
	IT3	Difficulty in analysis of the model outcome (e.g. sensitivity)
	IT4	Validation of the model does not entail recognition of the real mechanism

Table 4.2: *SWOT (Strengths-Weaknesses-Opportunities and Threats) analysis of the individual-based approach.*

are a good tool for synthesizing experimental observations (PO3): they are rapid to implement, analyze and evaluate (López et al., 2004) (PS4) and they provide criteria and indicators for assessing the quality of a given product and estimating the safety of a proposed process (PO4).

Continuous models are easily communicated and shared among researchers (PS3): they are comprehensible and provide analytical tools that do not require expert mathematical skills or too much calculus. The wide experience and general use of continuous models in predictive microbiology have given rise to standardized methodologies in the process of building, implementing and testing a model (PS2), which facilitate and improve the efficiency of the research process (Geeraerd et al., 2004). They also provide widely accepted models that can be used as building blocks for more sophisticated applications (PO1) (Li et al., 2007).

Continuous models usually employ parameters and variables that may be directly measured by observing the population and which are common to other models, so it is quite easy to compare population models with experimental observations (PS5) and with each other (PS3). The comparison is feasible even between models designed to deal with different situations or systems (PO2). Besides, there is a great amount of experimental

information stored in databases such as ComBase (Baranyi and Tamplin, 2004) and the Pathogen Modeling Program (Buchanan, 1991) in a format compatible with standardized population models (PO5). These databases also provide modeling tools that may be easily and effectively used to address specific systems (McMekin et al., 2006).

The basic drawback of the continuous approach lies in the plain fact that the object under study (the microbial community as a whole) is not an entity *per se*, but the aggregation of many agents (each of the microorganisms). Continuous models intrinsically ignore how the collective behavior of a population specifically comes after the activity and interactions of individual microbes and may overlook important information in situations where the discrete nature of the system is relevant (PW1).

PbMs can hardly deal with population heterogeneity (e.g. age distribution), local and cell-to-cell interactions (e.g. propagation of a disease within the population), competition (e.g. uptake of a finite amount of substrate), collaboration (e.g. use of extracellular polymeric substances in bio-films) and individual adaptation (i.e. phenotypic modifications). PbMs are often unable to link the properties of the system with the behavior at a cellular level (i.e. emergence), can not account for discrete events (e.g sudden changes in the medium, genetic mutations) and do not represent properly those systems where the continuum hypothesis is untenable (e.g. small populations). Indeed, For a long time, there has been a call for innovative approaches that could better tackle these questions (Dens and Van Impe, 2001), and many concerns are still a challenge to PbMs, for instance: how to deal with the increasing heterogeneity of microbial populations in axenic cultures (Vives-Rego et al., 2003), and how to describe and predict transitional states (e.g. lag phase or decay phase) with a mechanistic, rather than phenomenological, approach (McMekin et al., 2006).

The problem arises from the fact that when building a mechanistic continuous model, a big abstraction must be made in order to translate the hypothesis regarding the individual bacteria to a community level. In doing this, PbMs gain access to dealing with macroscopic observations, but are restricted to an averaged and plain description of the population. The results of the model must be translated again to the individual level to assess the validity of the model and thus gain insight into the microbes (PW2, PT1). As there is no systematic procedure to perform such a translation process, which is usually implicitly unspecified, it has been found that sometimes individual properties cannot be easily retrieved from the information at a system level (Kutalik et al., 2005). This difficulty often strangles the provision of a biological basis for the interpretation of the parameters appearing in the system-level model (PW3). Examples of this are the controversy surrounding the definition of the lag parameter λ in food microbiology (Prats et al.,

2008), or the interpretation of the dependence of the growth rate on the concentration of nutrient, according to the Monod model (Yu, 2007). These limitations entail that it is sometimes arbitrary (and therefore difficult) to decide which model to use in order to describe a particular phenomenon. Another consequence of the problem with the "translation between levels of description" arises from the meaning associated with probabilities in the model. The interpretation of the nature of probability (whether it stands for uncertainty in the measurements or for intrinsic variability) modifies the model outcome (PT2). This hinders decision making, for example when dealing with risk assessment (Nauta, 2000).

As a consequence of the above mentioned limitations, it is difficult to discriminate empirical information from mechanistic rationales in population models (see Section 4.1). Indeed, adding empirical terms to a model is a commonly used resort to improve the goodness of fit, yet it does not necessarily make its biological basis any better (PW4). This makes it very difficult to assess the appropriateness of a model beyond its accuracy (PT3)

One last challenge of continuous population models is integrating the information available at a molecular and cell level. The increasing amount of experimental information obtained at a cellular level (Brehm-Stecher and Johnson, 2004) is not easily covered with average PbMs. Moreover, the huge amount of data provided by he -omics research, which is still hardly assembled to build single cell models, is out of range of the PbM approaches, because the difficulties imposed by the jump from the molecular to the cell level must be added to those found in the translation from the cell to the population level (PW5).

4.2.2 Individual-based Models

Individual-based models have their basis in the model of the microbial cell and its immediate surroundings. The main entities of the model are individuals, the natural constituents of real populations (IS2). The provision of a biological basis is more direct for the parameters defined at an individual level than for those defined at a population level. As a result, IbMs are an appropriate tool for compiling the experimental information available at a molecular and cellular level, therefore linking the descriptions provided by systems biology (-omics data and bio-informatics) to the description of an entire community (IO4). The individual-based approach can also integrate different population-based points of view; for instance, the distinction between 'Survival' and 'Growth' models (a criterion to classify PbMs in the ambit of food microbiology) is not necessary, as IbMs account for both collective phenomena with the same model for the cellular dynamics (which copies cell maintenance vs. death, biomass growth and reproduction).

IbMs often use parameters that cannot be directly measured, or that are much more difficult to measure than macroscopic variables. Often such parameters must be deduced from system-level observations (IW4), and other must be fixed *ad hoc*. This apparent drawback may be an opportunity to indirectly estimate the values of variables unobservable in real systems (IO2). The value for the unmeasurable parameters can be determined (or at least delimited) by fitting the model outcome to the observed behavior of a real system. This has been used, for instance, to estimate the average minimum mass to start the individual reproduction process for a particular strain (Prats et al., 2008).

IbMs are particularly useful for dealing with small populations, since the smaller the number of individuals, the more the outcome of the model depends on the distribution individual traits. Moreover, IbMs have a more transparent and direct use of probabilities and statistics, which is critical when numbers are not large. In IbM simulations, the individual traits of the cells that constitute the inoculum are set according to a probability distribution function defined by the model, which can be contrasted to measurements. Then, many simulations are run with the same set of parameters and different random seeds to assess the effect of uncertainty in the processes. Finally, the set of parameters and/or the initial configuration of the model is varied to check the sensibility of the simulation. All these processes allow the staggered testing of the contributions to the variance of the simulations. In conclusion, IbMs can more clearly relate the variability of the model outcome with the diversity of individual traits, or with the uncertainty associated to the measurements (IS1, IO1).

Yet, it must be stressed that the problem of understanding and predicting microbial responses near the survival/death interface (the study of systems with small populations and/or with high uncertainty in the collective behavior) still persists. First, the results provided by IbM must still be handled with care so that no misuse of statistics is carried out (López et al., 2004). Secondly, even when the results seem statistically conclusive for a particular phenomena, the credibility of the model must be reinforced with prediction/reproduction of independently observable phenomena (see Section 4.1), which might be difficult for less predictable real systems. A critical factor to predictability is the cell damage that commonly occurs in the survival/death interface. This potential damage is phenotypically highly variable and consequently very difficult to assess, still, it strongly drives the population dynamics. IbMs seem an appropriate tool to deal with this problem, as they permit taking profit of the advantages typical of probabilistic models (McKellar et al., 2002).

The main strength of IbMs probably resides in their capacity to disentangle the different microscopic (individual-level) factors that cause particular macroscopic (population-

level) behavior (IS4). IbMs allow a staggered inclusion of complexity in the model. For example, a given model can initially consider identical bacteria in a homogeneous medium, and then introduce spatial heterogeneity or individual variability in the population. Individual metabolism, interactions among cells, and any other factor can be taken into account by the model with increasing detail. With IbMs, each degree of elaboration of the model can be introduced and studied on its own without modifying the structure of the model. This implies that the effects of any individual biological hypothesis can be studied either alone or in combination with other mechanisms. This makes IbM a good complementary tool for the development of continuous mechanistic models (Ginovart et al., 2006):

i) firstly, because they allow more thorough characterization and better study of the underlying factors of empirical terms used in PbMs. For instance, IbMs can tackle the transitory states of population growth, such as the lag phase and the stationary and decay states of bacterial growth;

ii) secondly, each IbM can provide a more general applicability and may constitute a common framework for different population models. Once the microbes are well characterized, the model can explain their behavior in many situations or systems and with different detail (IS3).

Besides, IbM can inspire the development of novel experimental applications, for instance, new indicators and monitoring protocols to assess the quality and safety of a product (IO3). Such indicators could use measurements on distributions among the population to diagnose the state of a community during the transient states. An example of such indicator can be used to characterize the progress in the lag phase: it entails measuring the *mathematical distance* ($\mathcal{D}$) between the instantaneous normalized biomass distribution for a sample of an observed population and comparing it to the biomass distribution at the exponential growth phase, to predict how long will the lag last (Prats et al., 2006). This method has already been applied to the analysis of experimental measurements (Prats et al., 2010a).

This freedom from the constraints of analytical models comes at a price: it allows IbMs to become more complex and thus more difficult to develop, understand and communicate (IT2). Such complication, together with the relative youth of computer-aided modeling in the field of microbiology and with the unfamiliarity of the researchers with numerical (discrete) mathematics, hinders the widespread use of IbM (IT1). Such a limitation is being faced since recently with the promotion of standardized methodologies to develop, analyze and communicate IbMs (Grimm et al., 2006).

The foremost hindrance found when building an IbM is the definition of the essential individual traits (parameters such as the minimum mass to start the reproduction process or values for the individual variables) to be included in the model for the micro-organism (IW4). It is also difficult to choose the structure of the model that will correctly account for all the important interactions. The tricky task is deciding which of the numerous combinations of mechanisms and global structures of the model best represents a real situation. This becomes critical for the assessment of the degree of confidence associated with a simulation outcome. Assessing the error propagation through the model may not be enough, because sources of error may reside in the model structure (IW3). There is a lack of standard methods for guiding model building and analysis, and there is no conventional way to assess the soundness of a model (IW1) (Grimm and Railsback, 2005).

The lack of standard IbM frameworks makes it difficult to analyze a given model, and so may hinder the extraction of synthetic understanding (IT1). It also means there are few established models that can be used as building blocks for other models, which generally hinders the development of a field. A deeper concern is related to this lack of standards: models addressed to similar situations often have different parameters and variables and may show patent differences in their structure. This impedes the comparison between different IbM models (IT2).

When compared with continuous models that are used to explain the same problems, IbMs are slower to implement, require more time- and memory-intensive simulations (IW2), and are more difficult to understand and to communicate (IT2). IbMs also have to deal with a specific problem that arises from the use of different spatial scales in the model: the basic unit is the micro-organism (microscopic scale), but some important environmental processes may take place at the population level (macroscopic scale). The spatial and temporal scales of the microbial processes and the environmental processes may also be quite different from each other (IW5). As a result, IbMs are often "difficult to up-scale to larger spaces when there is heterogeneity at a larger scale" (Hellweger and Bucci, 2009), and turn to be inadequate to cover wide multiscaled systems. This makes many problems very difficult to tackle and leaves some of them beyond the scope of IbMs (see Section 3.5).

One final remark is valid for both kinds of approach. It must be stressed that, even when a model has been verified and validated, this does not ensure that the proposed mechanisms are the real underlying causes of the studied phenomenon (PT4, IT4).

4.2.3 Applicability of each approach

We will now discuss the applicability of PbM and IbM approaches. The discussion is summarized in Table 4.2.3.

	Mathematical model	Sample	Characteristics
Top-down	Mechanistic Model: *focus on providing understanding*	Baranyi and Roberts, 1994	Relates duration of the lag phase and the growth rate of a culture
	Empirical Model: *focus on describing, reproducing and forecasting specific systems*	Buchanan, 1991	Predicts growth of microbial population using information in databases
	Probabilistic Model: *focus on the effect of randomness*	McKellar et al, 2002	Studies the growth/no-growth interface
	Individual-oriented Model: *include a description of population structure or individual traits*	Doona et al, 2004	Introduces distributions of characteristic properties among the population
Bottom-up	Cellular automata: *focus on the local nature of interactions*	Barker and Gimson, 1993	Studies microbial growth and pattern formation in colonies
	Individual-based model: *focus on the individual diversity and on the emergence of collective patterns*	Kreft et al, 1998	Bacsim: versatile IbM bacterial simulator
Hybrid model	*Multiscalar models that combine the strength of both approaches*	Picioreanu et al 2004	IDynomics: IbM simulator with continuous modeling of transport phenomena

Table 4.3: *Samples of the different types of models described in Section 4.2.*

The 'Opportunities' and 'Threats' presented in the sections above provide a guide for the development of models in predictive microbial ecology. It is just a matter of taking advantage of the strong points of each approach and skipping over its limitations. Continuous models are appropriate to tackle simple systems as a whole (see Section 1.4.2), even at transient states with macroscopic heterogeneity (Gujer et al., 1999; Sole-Mauri et al., 2007). Individual-based Models are useful in the study of transient processes related to the evolution of individual properties and heterogeneity at the cell scale (Ginovart et al., 2005), whenever the local nature of interactions among individuals may be relevant, and also in dealing with small populations (Prats et al., 2006).

IbMs and continuous models can also be used together to describe the microscopic and local scale with an individual-based approach and the coarser scales by means of a continuous model. Such hybrid approaches are currently being used with notable success, (Picioreanu et al., 2004; Alpkvist et al., 2006), and may be a keystone in improving the general holistic knowledge of static microbial systems.

Other examples of the complementary use of these approaches are:

i) IbMs can be used as a source of virtual experiments in order to test the assumptions of a given mechanistic PbM (Prats et al., 2008), and also to discriminate between different hypothesis that can not be singled out experimentally (Ginovart et al., 2006).

ii) IbMs can provide a generic scaffold for the study of a given microbial system through different empiric PbMs. The rules are describing the cells are potentially valid under any environmental conditions. Then, the specific response of a microbial system under a defined situation (as described by an empiric PbM) is just a particular solution of the corresponding general IbM. The compatibility between the particular solutions with the general framework theoretically provides a criterion for the assessment of PbMs (Grimm and Railsback, 2005).

iii) On the other hand, empirical or phenomenological models can be used to synthesize experimental information and provide a framework for developing IbMs proposing the mechanisms that may be operating at a cell level (see, for instance, the use of the in Section 2.3.6 and the use of the WS-model in Section 3.4).

Continuous modeling is a well-established methodology that requires little improvement. However, further efforts must be made in the rationalized provision of biological bases

to mechanistic models. It would be particularly interesting to revise the cases where the PbMs and IbMs suggest different biological basis for the same system level observation.

On the other hand, specific improvements are required for IbMs so that they can be used comfortably in the field of predictive microbiology and become more widespread. Firstly, it is important to build a standardization of the IbM formalism to make the models easily communicable between modelers, and extensively available to any researcher as user-friendly tools. In this sense, it is a good idea to take advantage of the efforts being made in theoretical ecology toward reaching a consensus on a standard communication protocol; i.e. the ODD protocol, see Section 2.2 and the reference (Grimm et al., 2006). Secondly, it is crucial to standardize a methodology for model validation and for parameter estimation in IbM. And thirdly, it is important to develop quantitative analytic methods to treat IbMs and population models jointly and to guide the research process (Grimm et al., 2005).

When dealing with a well-known system or with already understood phenomena and looking for accuracy in the predictive capacity of the model, the best choice is an empirical continuous model optimized by its fit to experimental data. Thus, for most industrial applications under controlled conditions, these models (usually presented as packages in tertiary models) are the most suitable. However, IbMs may be recommended as a suitable tool to guide population models? building process, to analyze model coherence and outcomes and, in particular, to understand population dynamics better. They may be particularly useful when tackling systems in which the expected behavior predicted by continuous models is not consistent with the behavior observed in real systems. Once the IbM has been used as a tool for understanding, the improvement and application of the population approach may proceed more fluently, and the continuous model can be used as a more effective predictive tool. Hybrid models that combine the two approaches allow for multi-scale representations.

4.3 IbM, microbiology and basics of thermodynamics

The IbM approach to microbial communities explicitly tackles the connection between the microscopic and macroscopic worlds, covering several temporal scales (from the cell cycle $\sim minutes$, to the course of cultures $\sim days$ or of ecosystems $\sim years$) and distances (from the individual cell $\sim \mu m$, to the aggregate culture $\sim cm$ or even to ecological communities $\sim km$). Three levels of description are covered by this methodology: a) the microscopic level, b) the macroscopic level, and c) the mesoscopic connection between the two levels (see Figure 4.1).

This aim, to connect the microscopic and macroscopic worlds, evokes the joint application of probability theory and deterministic mechanics that eggs on statistical thermodynamics in physics. In this section we take advance of this analogy. Methods from statistical thermodynamics are employed here to gain holistic insight into the general framework of theoretical microbial ecology, and in malaria cultivation, in particular.

ODD is the standard formalism for communicating IbMs in general (Grimm et al., 2006), but the specific terminology of each model depends of the context -i.e. bacterial cultures and cell cultures employ different terms. The following analysis uses the nomenclature of physics. In order to avoid *linguistic misunderstandings*, some clarifications are outlined first.

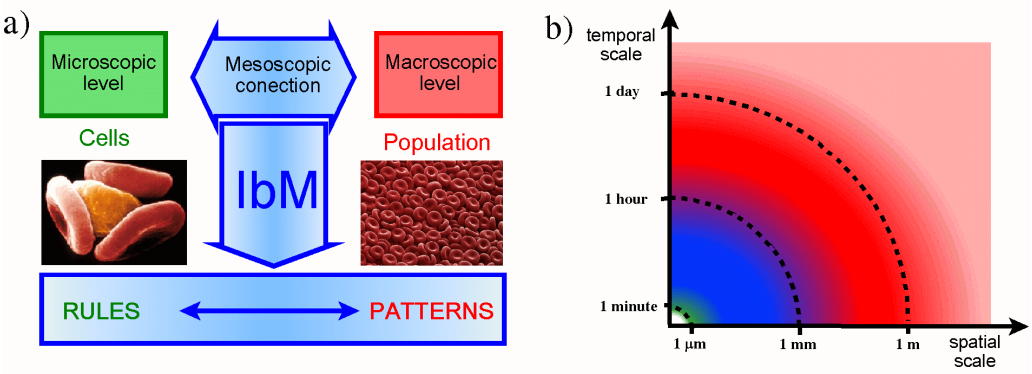

Figure 4.1: *(a) Levels of description covered by Individual-based modeling of microbial systems and (b) their characteristic scales.*

4.3.1 Basic concepts of thermodynamics in biology

Preliminary considerations

Systems that comprise a large number of elements can be tackled at two different levels of description: microscopic and macroscopic. Specifying their microscopic configuration requires detailed and exhaustive description of each component element, and this is not feasible in practice. The observation of any system on the whole takes a finite time, during which several microscopic configurations are adopted. The correspondence between the microscopic and macroscopic descriptions is not one-to-one: many microscopic configurations are compatible with a given macroscopic observation.

The notion of macroscopic state (*a.k.a* thermodynamic state, or just "state") of a system is central to the discussion below. The state of a given system is defined by a

set of observed macroscopic properties (state variables) that completely characterize its momentary condition. The definition of the state of a system does not depend on its previous history.

Thermodynamic systems fulfill the continuum hypothesis, and thermodynamic variables (e.g. temperature, pressure and volume, among others) are real numbers that represent statistics of mechanical properties of the elements composing the system (e.g. velocity, number of collisions). State variables are classified into intensive properties (which do not depend on the size of the system, e.g. temperature, density) and extensive properties (which depend on the size of the system, e.g. volume, mass). The ratio of two extensive properties can be used to define an intensive property (e.g. density is the volume of a system divided by its mass).

From the thermodynamic perspective, the state of a system is either an equilibrium state, a stationary state (*a.k.a* steady state) or a variable state (*a.k.a* transient state). A system in the thermodynamic equilibrium exhibit no net fluxes of energy or matter and its properties (intensive and extensive) are constant. All of its state variables are uniquely determined once some of them have been specified. Systems that exhibit net fluxes of energy or matter and whose intensive properties are constant are in a stationary state. Those systems whose thermodynamic properties vary with time are in a variable state (see Table 4.4).

From the thermodynamic scope, and strictly speaking, biological systems are open dissipative systems far from the thermodynamic equilibrium. They can be in a steady state (e.g.: homeostasis of a cell and balanced growth of a population, among others) when they can continuously exchange energy and mass with their local environment and maintain their (dynamic) internal organization through the active export of heat to their surrounding environment (Prigogine, 1955; Margalef, 1991; Toussaint and Schneider, 1998).

Conservation of mass and energy

Biological systems (each individual cell and the population as a whole) continuously exchange mass and energy with their surroundings. In doing so, they must obey the fundamental laws of conservation of mass and energy.

Conservation of mass imposes a balance on metabolic reactions: substrate uptake is spent in the increase of biomass and (eventually) in the excreted waste products. It also imposes the return to the medium of biomass after cellular death and lysis, as waste products or as available substrate.

Conservation of energy imposes additional constraints on cellular metabolism: energy obtained from substrate uptake is spent on cell maintenance and cell growth (if possible)

with a certain yield. The remaining fraction of energy is dissipated as heat. Metabolic heat results in a local increase in temperature, which is transported through the system and finally exported to the surroundings.

Although models may not explicitly account for every mass and energy exchange, an overall balance is always taken into account. Any violation of the conservation laws indicates that the model is not operating correctly.

Entropy and thermodynamic stability

The second principle of thermodynamics states that any isolated system tends towards a state of equilibrium after a transient state. Steady states can be indefinitely maintained in non-isolated systems, but this requires a continuous exchange of energy (at least, if often, also matter) with the surroundings of the system.

The tendency of thermodynamic systems towards equilibrium and towards steady states can be statistically interpreted as the trend towards those macrostates that are compatible with the maximum number of microscopic configurations (Kittle, 1958). The quantity used to measure the number of microscopic configurations that are compatible with an observed macroscopic state is called entropy (S). In statistical mechanics, the Boltzmann entropy (S_B) is an approximation to this measure (see Equation 4.1, where N is the number of accessible microstates, p_i is the probability of the system being in one of them, and $k_B = 1.380{\cdot}10^{-23}\frac{J}{K}$ is Botlzmann's constant). In classical thermodynamics, entropy represents the amount of energy spent as heat dissipation per unit temperature. In information theory, entropy is a measure of the variability of an outcome (Jaynes, 1957). In ecology, the term exergy (amount of energy available to be used) is commonly employed to account for the increase in entropy (Margalef, 1963).

$$\mathcal{S}_B = -Nk_B \sum_i p_i log(p_i) \tag{4.1}$$

Thermodynamic equilibrium is characterized by the maximization of entropy under the given macroscopic constraints (see Table 4.4). It shows a lack of heat dissipation and the entropy of the universe (the sum of the entropy of the system and the entropy of the environment) is constant ($\triangle S_{univ} = 0$). Steady states continually produce entropy ($\triangle S_{univ} > 0$), but the increase in entropy is the least possibly compatible with the macroscopic observations ($\frac{d(\triangle S_{syst})}{dt} = 0$). In any variable state, entropy is smaller than in the equilibrium, and entropy production bigger than in a steady state ($\triangle S > 0$, $\frac{d(\triangle S_{syst})}{dt} < 0$).

In far-from-equilibrium systems entropy cannot always be defined, so the characteri-

	Extensive Variables	**Intensive variables**	$\triangle S_{univ}$	$\frac{d(\triangle S_{syst})}{dt}$	Energy fluxes
Equilibrium	*constant*	*constant*	0	0	0
Stationary	*constant*	$\exists$ *gradients*	> 0	0	*constant*
Variable	*vary*	*vary*	> 0	< 0	*vary*

Table 4.4: *Apparent thermodynamic states and their characteristic features.* $\triangle S_{univ}$*: increase in entropy of the universe;* $\frac{d(\triangle S_{syst})}{dt}$*: entropy production of the system.*

zation appearing in Table 4.4 is not exhaustive but rather orientating. Still, generalized thermodynamic terms and laws can be applied to biological systems, assuming that they lack appropriate formalization.

The fluxes of energy and entropy can be quantitatively assessed in biological systems (Odum, 1969). Any system resists being removed from its equilibrium state by countering the applied gradients; as applied gradients increase so does the system's ability to oppose further movement from the equilibrium, eventually leading to self-organized dissipative structures (Toussaint and Schneider, 1998). Computer simulations indicate that any ensemble of agents in a complex interaction may evolve towards temporal or spatial stable patterns, under the appropriate constraints (Solé and Valls, 1992; Solé et al., 1992). Understanding the dynamics of open dissipative systems (and biological systems, in particular) does not require the strict definition of S, but the identification of principles and mechanisms that can lead towards the evolution of stable patterns (Karsenti, 2009).

Maximum entropy and biomass distribution function

Experimental observations show that biotic communities exhibit stable states during their evolution; for example, empirical data for the biomass distribution function among a bacterial colony growing unrestricted fit well with a log-normal distribution function (Koch, 1966; Akerlund et al., 1995). Such an observed shape is also theoretically obtained from the optimization of the entropy of the system (MAXENT, Equation 4.1): the log-normal distribution results from imposing the maximization of biological diversity among the population, once physical energetic constraints are taken into account (Wagensberg et al., 1988a; Wagensberg et al., 1988b; Margalef, 1991).

Population diversity is defined using the Shannon entropy ($\mathcal{S}$), with the different classes of organisms (i) being different bins in a histogram of the biomass distribution (m_i), and the relative frequency (q_i), the fraction of individuals with $m = m_i$:

$$\mathcal{S} = -\sum_i q_i log(q_i) \tag{4.2}$$

The biomass distribution $\widetilde{q_S}(m)$ that maximizes entropy depends on the constraints imposed on the system. If there is no additional constraint, then $\widetilde{q_S}(m) = k$ is a uniform distribution, with $k = \frac{1}{N}$ being the normalization constant and N the total number of i classes.

If the diversity is maximized assuming the following conditions:

i) normalization of the biomass distribution,

$$\sum_i q_i = 1 \tag{4.3}$$

ii) and finite observed mean mass of the culture $\overline{m}$ (or limited total biomass).

$$\sum_i m_i \cdot q_i = \overline{m} \tag{4.4}$$

The resulting distribution is given by the exponential decay (Equation 4.5):

$$\widetilde{q_s}(m_i) = \frac{e^{-\beta m_i}}{Z} \tag{4.5}$$

Where $Z = \frac{\triangle m}{\overline{m}}$ is the normalization factor, determined by the choice of the histogram bins ($\triangle m$) and the normalization of the probability of the sample space (Equation 4.3), and $\beta = \frac{1}{\overline{m}}$ is the Lagrange multiplier determined by the mean mass observed in the exponential growth (Equation 4.4). The log-normal distribution is obtained by introducing additional constraints into the biomass distribution (Equation 4.6):

iii) uncertainty (L) in the reproduction mass (m_{Ri}) and minimum threshold for bacterial mass (m_0):

$$L = \sum_i q(m_{Ri}) log \left[\frac{m_{Ri} - m_0}{\triangle m} \right] \tag{4.6}$$

IbM provides a framework to connect the theoretical approach with experimental observations. It reproduces the described behaviors (Giró et al., 1985; Bermúdez et al., 1989; Ginovart et al., 2002a) through the contribution of two simple mechanisms: a) individual variability on the uptake of nutrients and b) minimum threshold to start the reproduction cycle (see Figure 4.2).

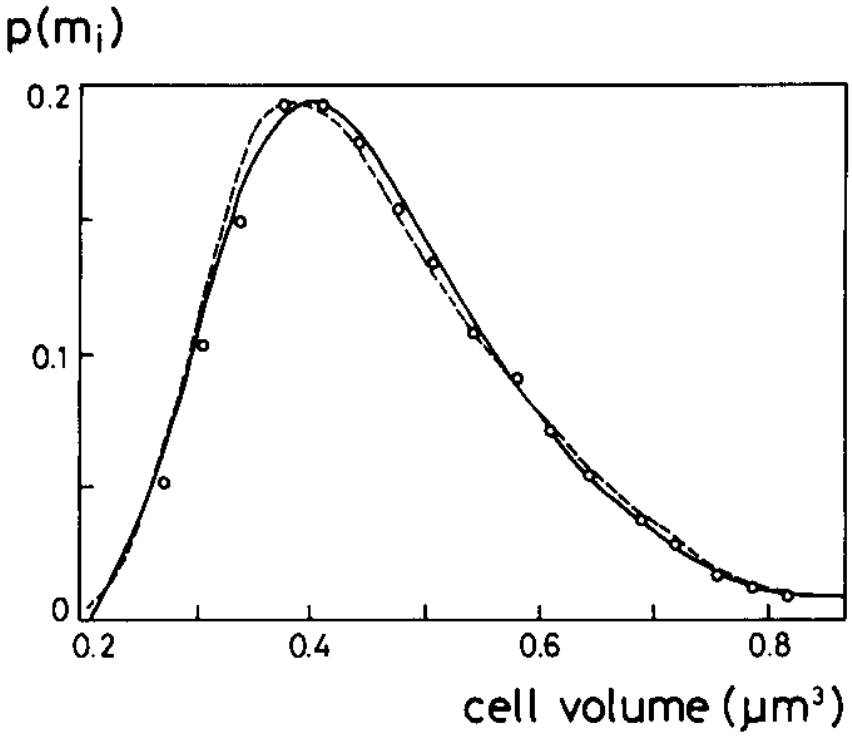

Figure 4.2: *Comparison of the observed biomass distribution function for a bacterial culture (dashed line; after Koch and Higgins, 1984), the theoretical prediction from the maximization of entropy (solid line) and the simulation results obtained with Barcelonagrama, a preliminary version of INDISIM. Reprinted from Wagensberg et al (1998a).*

Mathematical distances and population dynamics

Population diversity is a valuable tool for studying the dynamics of populations. For instance, in the study of the microscopic causes of the lag phase in bacterial cultures (see Section 4.3.2), Prats *et al.* (2006) defined a measure, the "product distance" (D_p), to assess the similarity of a distribution at a given moment with the one corresponding to stationary state under given conditions. The product distance consists of two factors,

$$D_p = D_{\overline{m}} \cdot D_q \tag{4.7}$$

They are two independent measurements, the *mean mass distance ($D_{\overline{m}}$):*

$$D_{\overline{m}} = \frac{\overline{m} - \overline{m}_{exp}}{\overline{m}_{exp}} \tag{4.8}$$

and the *mass distribution distance (D_q):*

$$D_q = \sum_i \left| q(m_i) - \overline{q}(m_i)_{exp} \right| \tag{4.9}$$

Where $\bar{m}$ is the mean mass of the population, $\overline{m}_{exp}$ is the mean mass during the exponential growth, $q(m_i)$ is the frequency of bacteria with mass m_i and $\overline{q}_{exp}(m_i)$ is the expected frequency of class i during the exponential growth. Unlike other measures of the similarity between probability distribution functions, such as the chi-square test or the K-S test (see Appendix C), the product distance has a biological meaning and provides an increase of understanding that can be used to better study bacterial growth (Prats et al., 2008; Prats et al., 2010a).

Three major conclusions can be drawn from the arguments above. First, thermodynamic laws can be applied to microbial populations. Second, IbM constitutes a consistent method for understanding how cellular mechanisms give rise to the observed macroscopic (thermodynamic) laws. And third, the application of thermodynamic perspective to microbial systems allows a quantitative definition of the state of the microbial community. It must be stressed that assessing the state of a system is crucial to ensure a common reference to compare different systems.

4.3.2 Characterization of bacterial and cell cultures

Microbial communities prevail only as open dissipative systems: they must continuously exchange heat and mass with their environment. Experimental models of real systems try to recreate some aspects of the natural world: they can be classified according to their mass and energy transfer regimes (see Table 4.5). *Batch cultures* are closed systems with finite substrate supply. *Fed-batch cultures* are open batch cultures where the substrate is permanently added to supply the microbial community with mass and exergy. *Continuous cultures* entail the renewal of both the culture medium and the bacterial population. The entries and outputs of substrate and cells can be performed continuously or at at discrete events.

Bacterial growth in axenic cultures

Different stages or phases can be observed during bacterial growth in a closed system. After inoculation, cells must adapt to new environmental conditions and the population undergoes the *lag phase*. Gradually, each bacterium reaches its optimal uptake rate, in which its increase in biomass is the fastest, its reproduction cycle is the shortest, its offspring increases geometrically and the population enters the *exponential growth phase*. During the *exponential growth phase* bacteria may undergo *balanced growth*, which entails constant population structure and distribution of properties. This is maintained as long as the environmental conditions remain favorable and there is enough nutrient for every cell. When the culture medium runs out of nutrient (or when the medium conditions become inauspicious, due to a harmful secondary metabolite, for instance), bacteria start their metabolic adaptation to the new adverse conditions: the culture enters the *stationary phase* and, if circumstances remain adverse, subsequently they enter the *death phase* (or *decay phase*). The *stationary phase* is anything but stationary from the thermodynamical perspective since the population structure and the distribution of cellular properties continually vary.

Bioreactors operate under regimes with continuous input of nutrients and output of exhausted medium and cells. In an ideal bioreactor, a microbial population grows unrestricted and the exchange regime is set to maintain a fixed cell density inside the reactor tank. A thermodynamically stationary state of the bioreactor is reached during the balanced growth phase and when microbial growth rate equals the output (or dilution) rate. Then, the number of cells and the distribution of properties remain constant.

Intraerythrocytic malaria cultivation

Cultures of *Plasmodium falciparum*-infected RBCs differ slightly from other microbial cultures. Firstly, RBCs use the substrate to maintain their structure and metabolism but do not increase their biomass, whereas the extracellular parasite has nearly no metabolic requirements. The most important expenses occur during the infection cycle, when the parasite reproduces itself. Maintenance requisites of each IRBC strongly depend on the infection stage. Secondly, RBCs act as a 'substrate' for parasite proliferation, and the maintenance of the infection requires a continuous input of fresh blood cells. Therefore, limitations on the propagation of the parasite may arise from the unavailability of RBCs and from the exhaustion of the medium (scarcity of substrate or excess of metabolic waste).

Several culture systems are distinguished among the erythrocytic cultivation of malaria.

Short-term cultivation of IRBCs lasts for three days at most, until healthy RBCs are no longer viable due to overall substrate limitations (see Sections 2.3.1 and 2.3.5). Short-term cultures are typically performed under closed conditions (in analogy with batch cultures), but medium renewal alone (in analogy with fed-batch cultures) does not improve their continuance (see Section 2.3.5). *Continuous long-term cultivation* of the parasite can be carried out either in *static cultures* or in *suspended cultures*. Both of these require the discrete renovation of RBC population every four days, at a minimum (see Section 2.3.6). Continuous long-term cultivation of the parasite with continuous RBC renovation has been reported in the literature (see Section 1.3.1). However, no bioreactor-like apparatus has been successfully developed for continuous malaria cultivation.

	Accessible thermodynamic states	**Examples in bacterial cultivation**	**Examples in malaria cultivation**
Conservation	*EQ., ST., VAR.*	*Spores, quiescent cells*	*Frozen RBCs, parasite strain*
Closed culture	*VAR.*	*Batch culture*	*Preservation of the infection*
Culture open to substrate exchange	*VAR.*	*Fed-batch culture*	*Preservation of the infection*
Continuous culture	*ST., VAR.*	*Bioreactor*	*Continuous flow, candle-jar method*
Natural system	*VAR.*	*Microbial community*	*Murine model, host infection*

Table 4.5: *Types of microbial systems and their thermodynamic characterization. EQ.: equilibrium state, ST.: stationary state, VAR.: transient state.*

4.3.3 Thermodynamic description of microbial populations

The previous general classification of thermodynamic states (see Table 4.4) is applied to different microbial systems (see Table 4.5) at different levels of description (see Figure 4.1). It is important to note that the definition of the state of a system depends on the scale of observation. For instance, temporal constancy apparent when observing the system at a certain level may turn into fluctuations when further detail can be appreciated.

A) Observation at a cellular level

The description of a single cell during temporal scales of the order of the minutes (*i.e.* shorter than the reproduction cycle for bacterial cells and much smaller than the life span of RBCs in malaria cultivation) allows defining the following thermodynamic states:

i) *Equilibrium state. Cell inaction:* the cell does not exchange mass or heat with its surrounding. This occurs when the cell has no activity; strictly, that is only when the cell is dead. It can also be considered to be the state of dormant cells with very low metabolic upkeep (*e.g.* quiescent bacterium, spores, or frozen RBC).

ii) *Stationary state. Upkeep:* the cell uptakes a constant amount of nutrient and excretes a fixed amount of heat and metabolic products. The metabolic output is spent in the cellular upkeep: maintaining homeostasis, cell structure and activity (*e.g.* ideal sustaining of a bacterium or healthy RBCs).

iii) *Variable states. Bacterium growth:* cellular uptake exceeds the upkeep requirements and the energy surplus is spent in cellular increase of biomass. If maintained, this stage eventually leads to cells big enough to start a reproduction cycle. *Bacterium shrinkage:* cellular uptake is not enough to cover the upkeep requirements and the energy deficiency is fulfilled through self-digestion. *Infection cycle:* parasite reproduction entails the continuous alteration of RBC structure, and the consequent variation of the properties of the IRBC.

B) Observation at a population level

Description of the system at a population level and consideration of a coarser temporal scale on the order of several days (i.e. greater than bacterial cellular cycle and malaria infection cycle) allows us to distinguish the following states:

i) *Equilibrium state;* macroscopic fixed states appear only if there is (approximately) no metabolic activity, and require environmental conditions that limit cellular activity (*e.g.* reposition of spores, freeze storage of cell cultures).

ii) *Stationary state; - balanced growth in ideal bioreactors:* any biological activity requires the consumption of substrate and therefore needs a continuous regime of substrate exchange. For bacterial cells, ideal sustainment is not

viable for every cell in a community at once, as some cells will grow while others will fail in the competition for nutrients. Constant intensive properties, characteristic of stationary states, appear only if all bacteria are in optimal growth conditions, while constant extensive properties require the continuous extraction of a fraction of the population. The same argumentation holds for the stationary states in *in vitro* maintenance of *Plasmodium* infection in bioreactor-like cultures.

Discrete renewal of the population can lead to close-to-stationary states (stationary states that vary periodically). These states can be maintained in the long run, and they exhibit constant but show periodic variations on their intensive variables related to the short-term discrete renewals (*e.g.* continuous long term cultivation of *Plasmodium falciparum,* see Sections 2.3.6 and 4.3.4).

States that appear to be stationary at this level of description can also turn into variable at larger scales. For instance, the long-term degradation of individual cells becomes apparent in the effect of the storage period of the RBC sources in the continuous cultivation of *Plasmodium falciparum* (see Section 2.3.3). Such an effect is also observed in the long-term evolution of yeast populations, and how it affects industrial fermentation processes (Prats et al., 2010b, in press).

iii) *Variable states. Systems with no population renewal:* any other macroscopic system that maintains cellular activity in a system closed to population renewal is unavoidably in a transient state. This is true even for systems in an apparent steady state, such as the lag phase and the stationary phase of bacterial growth. In such cases, despite there being no variation in cell density is observed, other intensive variables (*e.g.* cellular composition and structure, or biomass distribution among the population) are continuously changing due to cellular adaptation to the new conditions (Prats, 2008; Prats et al., 2008; Prats et al., 2010a).

C) Mesoscopic connection

Knowing the state of any real system is crucial to interpreting the measurements and having predictive capability. The structure of the population is an indicator of the thermodynamic state of microbial communities. The quantitative assessment of the distribution of individual traits among the population is a valuable tool for assessing such states. The study of distributions among the population and of emerging collective

patterns constitutes the mesoscopic connection between the microscopic and macroscopic levels of description (Ferrer et al 2008a).

Recent experimental techniques provide a huge amount of data at the cellular level (Hellweger and Bucci, 2009). The interpretation and assimilation of this information requires models that can connect individual variability and measurements of the statistical properties of the population with macroscopic patterns and quantitative measurements at a system level. This is another area of application of the mesoscopic analysis of populations.

A distinctive mesoscopic tool is the statistical distribution of individual properties. The use of the biomass distribution function for bacterial populations was introduced in Section 4.3.1. The role of the distribution of post-invasion times (t_{INF}) among IRBCs will be examined below.

4.3.4 Population structure in *Plasmodium falciparum* in *in vitro* cultures

An apparent population structure of *P. falciparum*-infected RBC in static cultures is the distribution of infected forms. In this section, the evolution of the distribution function of post-invasion times in the simulated IRBCs ($q\left(t_{INF}\right)$, see Section 2.3.4) is examined through a deeper revision of experimental observations and through the analysis of INDISIM-RBC outcome.

Real cultures usually start with young populations of IRBCs mainly comprised of ring stages. This preponderance of young forms is observed all through the cultivation, and is maintained after multiple subcultures (for instance, note the differences in the average observed parasitaemias in Figure 2.6 b, c, d and e in Section 2.3.2). The evolution of the distribution of IRBCs is governed by many synchronizing and desynchronizing factors: individual diversity, differential mortality for young and mature IRBCs, and external manipulations, among others.

INDISIM-RBC simulations permit the definition and assessment of these factors one by one. The qualitative analysis of $q\left(t_{INF}\right)$ shows that the population of IRBCs desynchronises after each infection cycle due to individual variability. The introduction of an increased accidental death probability for the mature stages of the infection, together with the periodic subcultivation of the parasite, compensates this disordering trend and lead to stagnation of desynchronization (see Figure 2.12 in Section 2.3.4).

A more thorough analysis of the population structure requires a minute treatment of the post-invasion distribution function. Firstly, it must be stressed that the distributions shown in Figure 2.12 are selected snapshots that resemble Gaussian distributions (*a.k.a.*

normal distribution). Such a shape is not the most common one during the course of a simulation. Indeed, the Gaussian distribution reflects a particular instant state found in cultures with rather synchronized inocula at the middle of the infection cycle. In other snapshots, the post-invasion-time distribution may appear truncated or may present much more disordered populations (see Figure 4.3). Withal, the qualitatively observed trend, common to all IRBC populations, is a tendency towards increasingly homogeneous $q\left(t_{INF}\right)$ (although highly fluctuating and with a limited bandwidth).

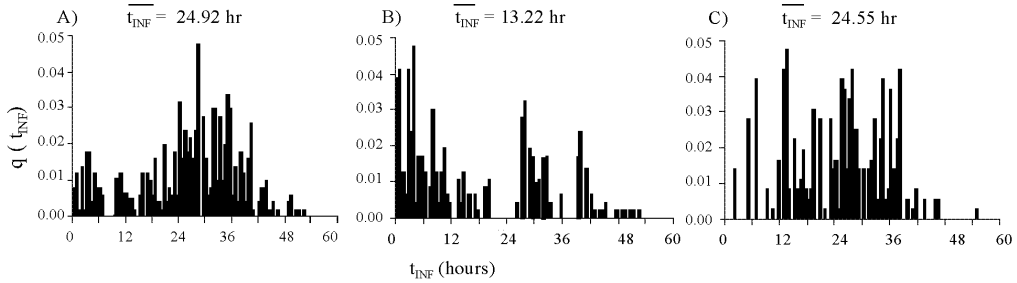

Figure 4.3: *Snapshots of the post-invasion-time distributions ($q\left(t_{INF}\right)$) obtained at different moments ($\overline{t_{INF}}$) of INDISIM.*v3D *simulations; (a) population structure after 3 generations showing a normal distribution at the middle of the infection cycle and the next generation of IRBCs at $0-6$ h; (b) population structure after 5 generations showing a truncated normal distribution of a rising generation between 0 and 18 hours and the remains of the previous one; (c) Disordered distribution of IRBCs after 12 generations (completed infection cycles).*

The quantitative analysis of $q\left(t_{INF}\right)$ requires defining its relevant statistics: the instantaneous mean value ($\mu = \overline{t_{INF}}$) and the standard deviation ($\sigma(t_{INF})$), calculated over the whole population at each time step. Figure 4.4 shows their temporal evolution in a static *in vitro* culture with periodic sub-cultivations under *candle-jar* culturing conditions. The distribution $q\left(t_{INF}\right)$ does not tend to a stationary state but rather its shape varies with cycles of 48 *hours*. The mean value $\overline{t_{INF}}$ oscillates in regular cycles with decreasing amplitude, gradually tending to a the mean value that can be estimated to be half the duration of the infection cycle. Meanwhile, the variance also oscillates with a periodic 48 *hour* cycle and 180^{o} out of phase. The standard deviation is maintained until the first invasion takes place, and reaches its maximum values near the moments of majority lysis, when most of the IRBCs are in the end of the infection cycle, while the next generation of parasites is already rising. The variance decreases as the IRBC population becomes more and more homogeneous and the contrast between young and mature forms fades.

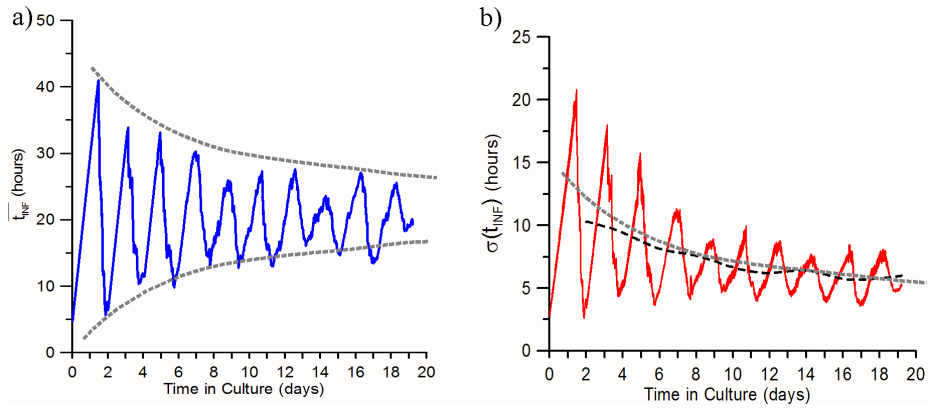

Figure 4.4: *Evolution of (a) the mean value ($\mu = \overline{t_{INF}}$), and (b) the standard deviation ($\sigma = StdDev(q\,(t_{INF}))$ of the distribution of post-invasion times among a simulated population of IRBCs under candle-jar culturing conditions, with a synchronous inoculum: $\overline{t_{INF}}_0 = 5\ hrs$, $\sigma_0 = 3\,hrs$. Solid lines show simulation outcome. Dotted lines show the envelope of $\mu(t_{INF})$, and is qualitatively compared with the average evolution of $\sigma(t_{INF})$. Dashed line is the observed average evolution of $\sigma(t_{INF})$, smoothed each 2 days.*

The variability of the distribution function, $\sigma = StdDev(q\,(t_{INF})$, which periodically oscillates in counter-phase with $q\,(t_{INF})$, tends towards oscillations of fixed amplitude around a fixed mean value value, $< \sigma >_t$. So there is a stationary degree of homogeneity of $q\,(t_{INF})$ that is maintained through successive cycles and which results from the contributions of the distribution among the present (f_{ts}), past (p_{ts}), and next (n_{ts}) generations of IRBCs (see Figures 2.12 and 2.13 in Section 2.3.4). This tend towards increasingly homogeneous distributions, until reaching an optimal shape that periodically varies through time, is consistent with the principle of maximum diversity -that is, with the optimization of $\mathcal{S}$.

The absence of a steady probability distribution function in the long-term cultivation of the parasite does not impede the application of variational principles similar to those presented in Section 4.3.1. In the present analysis, the probability distribution function $q\,(t_{INF})$ continually oscillates, and consequently, no fixed observable can be used to define and optimize the entropy of the system, $\mathcal{S}$.

However, if the evolution of the population is averaged over the period between two successive subcultivations, a mean entropy of the system can be defined $\mathcal{S}_{<48>}$ (see Equation 4.10). This entropy monotonically increases until reaching a steady state (see Figure 4.5).

$$\mathcal{S}_{<48>} = -\sum_i < q_i >_{48} log(< q_i >_{48}) \tag{4.10}$$

Where $< q_i >_{48} = < q(t_{INF}) >_{48}$ is the distribution of post-invasion times averaged over 48 hours. Averages are calculated from the first day and each of them includes a subcultivation.

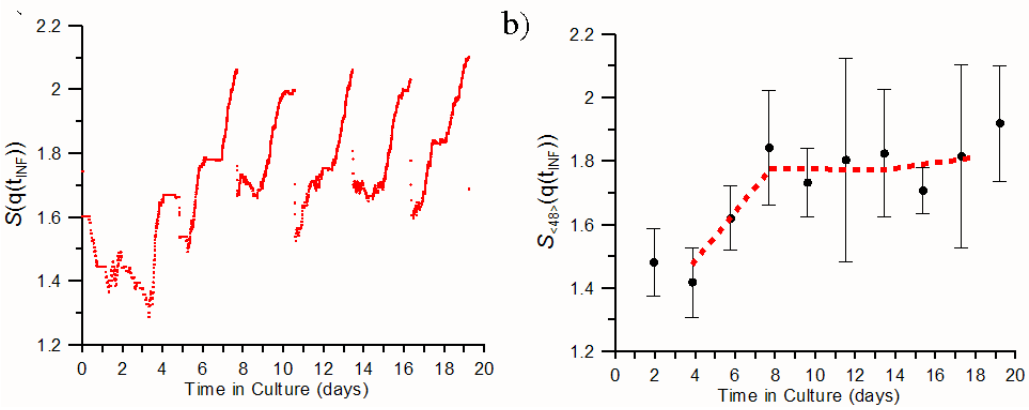

Figure 4.5: *Evolution of the entropy of the population of IRBCs in continuous static cultivation; (a) entropy calculated with the distribution of post-invasion times; (b) entropy calculated with the distribution of post-invasion times averaged over a subcultivation period of* 48 *h. Dashed line corresponds to the entropy averaged over* 8 *days.*

The definition of this entropy has an academic interest, but it is hardly applicable to real culture systems because the post-invasion time of IRBCs is not easy to measure.

Mathematical distance to the stationary state

The distribution of post-invasion stages (q_{ICL}) can be used as a measurement of the entropy of the system (see Figures 4.6 and 4.7). The first thing to note from Figure 4.6.a is that the simulated stationary distribution of infection stages ($\widetilde{q_S}(ICL)$) agrees with the experimental measurements averaged over time $\langle q(ICL) \rangle$. The prevalence of young forms of IRBCs observed in real systems is reproduced by the simulation outcome. But does this distribution really arise from the optimization of a variational principle?

According to the ergodic hypothesis, if the system is in a stationary process, the distribution of t_{INF} averaged through time ($\langle q(t_{INF}) \rangle$) should be similar to $\widetilde{q_S}(t_{INF})$ at the stationary state, a distribution that maximizes population diversity under the observed constraints. If the only constraint is the observed mean value $\overline{t_{INF}}$, the expected

distribution function is an exponential decay similar to that presented in Equation 4.5. Figure 4.6 shows $\langle q\,(t_{INF})\rangle$ covering multiple infection cycles and sub-cultivations.

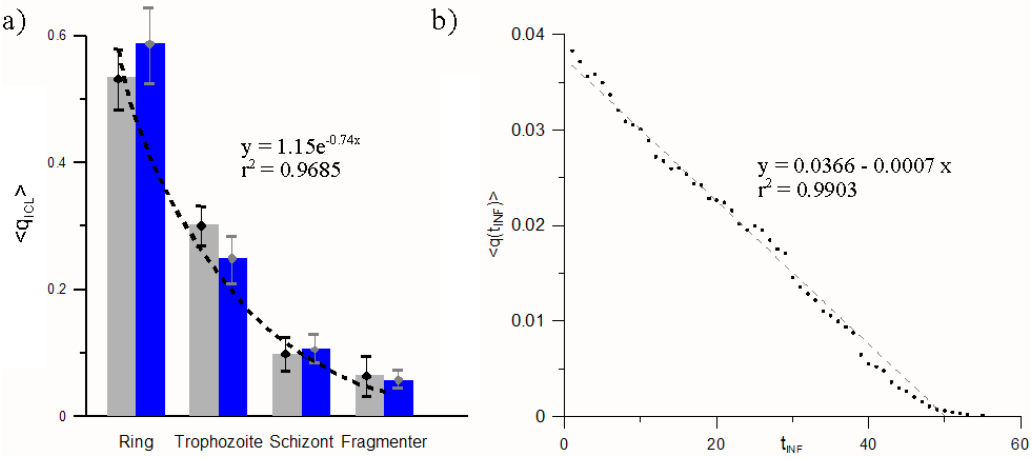

Figure 4.6: *Distribution of post-invasion times among IRBCs, averaged through a complete simulation of a static culture. (a) Histogram representing the cumulative prevalence of the different infection stages* $< q_{ICL} >$*. Simulation outcome (grey bars) are compared to the experimental data (blue bars) of the candle-jar and continuous flow cultures (Trager and Jensen, 1976). Dashed line shows that simulation results best fit an exponential decay. (b) Averaged distribution of post-invasion times* $< q(t_{INF}) >$ *in a static culture with* $P_{death}(\forall) = 0$*. Simulation outcome best fits a linear decay (solid line).*

The distribution $\left\langle q(\overline{t_{INF}})\right\rangle$ does not resemble the inverse exponential function found in Equation 4.5, but better fits a linear decrease (see Figure 4.6b). The decrease in $\left\langle q(\overline{t_{INF}})\right\rangle$ for increasing t_{INF} does not result from the optimization of any analytical function, but is a consequence of the particular course of the IRBCs' life.

IRBCs have a mortality rate that results from the joint effect of $P_{death}(RBC)$ and $P_{death}(t_{INF})$. The distribution $\left\langle q(\overline{t_{INF}})\right\rangle$ also reflects the repeated removal of IRBCs in subcultivation (see Section 2.2). If there were no mortality of RBCs and no subcultivations, Figure 4.6b would show a flat line: every IRBC would undergo its complete infection cycle. The slope of the linear fit is only slightly steeper than the slope corresponding to the expected removal rate due to subcultivation (~ -0.0005). The result is consistent with the optimization of diversity, but the constraints on $\left\langle q(\overline{t_{INF}})\right\rangle$ are stricter than the observable mean value.

In any case, the distribution of post-invasion times allows for the definition of an entropy function ($S(ICL)$, see Equation 4.2). This entropy can be compared to the entropy

corresponding to the stationary system ($< S_{exp}(ICL) >$), which can be calculated from the measurements on real systems averaged over time, to define a mathematical distance between the two ($D(ICL)$).

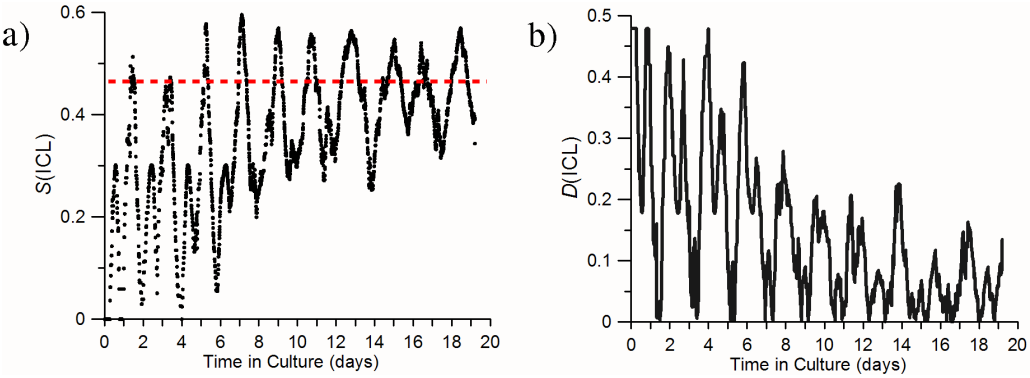

Figure 4.7: *Evolution of the (a) entropy and (b) mathematical distance calculated with the distribution of infection stages (ICL). a) Black dots indicate the evolution of $S(ICL)$ according to the simulation outcome. Red dashed line indicates the measured entropy averaged over a continuous cultivation using the candle-jar method $< S_{exp}(ICL) >$; b) calculated mathematical distance between $S(ICL)$ and $< S_{exp}(ICL) >$.*

The mathematical distance $D(ICL)$ can be used to assess the degree of closeness to the stationary state of a given culture. In fact, as $S(ICL)$ continuously oscillates, only high values of $D(ICL)$ indicate that the culture is in a state far from the stationary state, but low values of $D(ICL)$ are not enough to evaluate whether the culture is near stationarity or in a local minimum of the oscillation. All in all, $D(ICL)$ can be a useful tool for measuring to what extent two cultures are comparable, or at least for ruling out their equivalence.

4.4 Closing remarks

In the recent years, Individual-based Modeling has become a fully incorporated part of the methodologies used in predictive microbial ecology. Beyond the specific contributions of IbM to particular problems, this bottom-up approach involves a shift in the representation of microbial populations. Models based on individuals allow us to check the validity fundamental mechanistic assumptions regarding cellular behavior.

The rules proposed by different models can be used to evaluate the consistency and biological basis of the empirical rules defined at a system level. They can also be experi-

mentally contrasted through recently available measurements of population heterogeneity (e.g. fluorescence, cytometry and scanning microscopy, among others).

The joint use of IbM and the perspective of statistical thermodynamics, provides with tools that explicitly tackle mesoscopic connection between theoretical fundamentals (set at a cellular level) and observations on the real systems (usually obtained at a collective level). The application of such a view to INDISIM-RBC has been put forward as an example of the potentiality of combining the theoretical, experimental and computational approaches in predictive microbiology.

Chapter 5

Conclusions and perspectives

This chapter outlines the main results and consequent specific conclusions of this manuscript. Section 5.1 summarizes the general conclusions drawn from the thesis. Section 5.2 presents, in table form, a synthesis of the results in accordance with the structure of the specific objectives presented in Section 1.5. Finally, Section 5.3 compiles the pending issues and proposes lines for further work.

5.1 General conclusions

Malaria is still a major burden that causes approximately one million deaths annually worldwide. It is also a complex challenge to scientific research that covers many areas of work: biology of malaria parasites and their vectors, clinical disease management and treatment, biochemistry and molecular biology, epidemiology, and transmission dynamics, among many others. *In vitro* cultivation of the parasite is essential for the development of new drugs, but the control of the cultures and the interpretation of the observed behavior is not straightforward.

Progress in malaria research requires multidisciplinary collaboration. The contributions of experts in each specific field must be combined with those of experts in other disciplines, who offer a different and eventually revealing perspective. The constitution of a common framework for development analysis and communication among experts in each field is an imperative to this end. In this thesis, PoM has been adopted as the most appropriate way for approaching the raising of models in predictive microbiology and the ODD protocol has been proposed as the standard tool for communicating them.

Modeling can be more oriented towards increasing predictive capability or towards

understanding. In this thesis, the main objective of the model has been the interpretation of the results of pre-existing experiments. But the model has served to propose novel experimental procedures, such as the discrete and gentle agitation of static cultures for increasing and stabilizing the harvests. The model has also served to explore different experimental setups as model-generated hypotheses were tested, such as the packing of RBCs in the hematocrit layer. Finally, the model has been used to estimate the values of some magnitudes that are difficult to assess, such as the depths of RBC and IRBC viability.

PbM and IbM are two complementary strategies to tackle microbial populations; the former is more oriented towards empirical prediction and the latter is more oriented towards comprehension. In this thesis, the suitability of each approach to tackle specific problems in predictive microbiology has been evaluated.

IbMs define a set of rules governing each cell and its interactions with others and with its immediate surroundings. From these sets of rules, and taking into account diversity within the population and a certain degree of randomness in the individual processes, IbMs explicitly show the emerging behavior of the system as a whole. Methods from statistical thermodynamics can be applied to microbial communities to analyze the *mesoscopic connection* between the distribution of individual properties among the collectivity and the macroscopically observed features. In this thesis, IbM and the thermodynamic approach have been used together to better understand the emergence of macroscopic patterns from the characteristics of the population structure.

The specific application of the IbM methodology to cell cultures of *Plasmodium falciparum*-infected erythrocytes in *in vitro* cultures has been carried out through the development of the model and simulator INDISIM-RBC. The general conclusions to be extracted from the work here presented can be outlined as follows.

We have shown that INDISIM-RBC is a good tool to improve understanding of *Plasmodium falciparum*-infected RBCs in *in vitro* cultures. It is a mechanistically rich model, because it quantitatively reproduces and predicts several patterns observed in real cultures at different levels of description.

We demonstrated that INDISIM-RBC can be used to study in detail aspects of malaria cultivation that remain unclear, to perform virtual experiments, and to prompt novel lines of research and examine potential experimental techniques.

In particular, INDISIM-RBC has been used to predict the behavior of the current experimental culturing protocols in static cultivation, to set the subcultivatuion periods and renovation rates that lead to a stable and well performing continuous culture system, and to define the optimal sizes and shapes of the hematocrit layer in cultivations that

less harm the cells in culture.

The specific results obtained with this simulator are outlined in the following section, in accordance with the specific objectives listed in Section 1.5.

5.2 Synthesis of specific results

A) Conceptual approach

A1) Key experimental measures that provide information relevant to the study of malaria-infected RBCs in *in vitro cultures.*

The usual indicators of the performance of the *in vitro* cultivation of *Plasmodium falciparum* are the growth ratio GR, and the percentage of infected RBCs, or parasitaemia $\%I$. These observations are taken at discrete events, usually prior to subcultivation, every 48 hours.

- INDISIM-RBC used these indicators for the calibration of the model (Sections 2.3.1, 2.3.2 and 3.3.1). Increasing the resolution of the samples to periods smaller than 24 h did not lead to a significant improvement in the information obtained (Section 2.3.6).

The *characterization of the population structure* gives additional information. The distribution of individual properties among the population is often measured as an indicator of the degree of synchrony, which is considered to influence the growth ratio.

- INDISIM-RBC used the distribution of infection stages among IRBCs as an indicator to calibrate and validate the model (Sections 2.3.1, 2.3.2 and 3.3.1).
- The characterization of the population proved to be essential for assessing the state of the culture and for gaining predictive capability (Sections 2.3.4 and 4.3.3).

Finally, the conditions of the culturing medium are often evaluated by measuring the concentrations of different metabolites in the supernatant.

- INDISIM-RBC reproduced the overall consumption of glucose and production of lactate (Section 2.3.5). The local distribution of substrate was pointed as one of the mechanisms that limit the proliferation of the parasite in static cultivation (Sections 2.3.5 and 3.5.1).

A1) - *extension*- Characterization of the population structure

Bacterial populations and cell cultures can be thermodynamically characterized as open dissipative systems that continuously incorporate mass from their environment and release heat. Different culture systems can be classified according to their mass transfer regimes. Different states are accessible to each kind of culture (Section 4.3.2).

- We showed that the definition of the state (i.e. *equilibrium, stationary or variable*) of a microbial system depends on the scale of observation (i.e. *micro-/macro-scopic*; *short/long-term*) and that it can be assessed through the measurement of the appropriate characteristics (Section 4.3.3).

A thorough characterization of the state of populations allows for drawing a mesoscopic connection between microscopic and macroscopic levels of description (Section 4.3).

- The statistical distribution of individual features among the population (*a.k.a.* population structure) proved to be a good tool for assessing the state of microbial communities.
- The distribution of post-invasion times $q(t_{INF})$ and of infection stages $q(ICL)$ among IRBCs proved to be good indicators of the state of the infection in *Plasmodium falciparum* in *in vitro* cultures (Section 4.3.4).
- The analysis of *IbM* data suggested that the state with maximal population heterogeneity under the continuous cultivation of *P. falciparum* can lead to a stationary state characterized by a certain $q(t_{INF})$ which is periodically oscillating (Section 4.3.4).

A2) Individual features and local mechanisms that must be considered in order to describe the proliferation of the malaria infection in *in vitro* cultures.

INDISIM-RBC comprises RBCs, IRBCs, merozoites and medium. The local interactions that govern the infection dynamics were the competition for nutrients among RBCs (Sections 2.3.5 and 3.5) and the local transmission of the parasite (Sections 2.3.3 and 3.4). The most relevant features of RBCs are:

- Packing factor of RBCs; this controls the spatial structure of the hematocrit in static cultures.
- Age and time in cultivation; these govern the evolution of the susceptibility to invasion by the parasite.
- Metabolic requisites; these set the suitability of the local environment, which determines the viability of RBCs and IRBCs
- Susceptibility to infection of the RBC; this varies with the blood source and affects the infection growth rate and parasitaemia.

In addition to these features, IRBCs are characterized by:

- Duration of the infection cycle; this affects infection growth rate.
- Fragility of the infection stages; this affects infection growth rate.

Merozoites are not described in such detail, but are characterized by:

- Egress; this describes their expulsion during the lysis of IRBCs at the end of an infection cycle. The mechanism that performs best in INDISIM-RBC is consistent with recent observations at a cellular level.
- Mobility; this depends on the packing factor of RBCs and controls the spreading rate within the hematocrit layer in static cultures.
- Prevalence in the extracellular medium and capability of infection of the parasite strain; these control the infection growth rate.

A3) External manipulations of the system on the whole influence the dynamics of the infection.

External manipulations on the culture system consist of agitation, medium renewal and subcultivation. INDISIM-RBC tested different cultivation protocols and simulation outcomes were compared with experimental observations.

- Simulations matched real systems in that:

  i) period and extent of subcultivation affects the population structure, modifying the infection dynamics. The GR_{48} is slightly affected by this modification (Sections 2.3.4 and 4.3.3).

  ii) medium renewal is required every 72 hours at most to avoid the local degradation of RBC surroundings resulting from diffusive limitations (Sections 2.3.5 and 3.5).

  iii) Periodic agitation does not have an effect on the short-term evolution of closed systems (Section 2.3.6).

  iv) Cultures in suspension perform better than static cultures under similar periodic subcultivation regimes (Section 2.3.6).

- According to the model, subcultivation is required every 5 days at most to avoid infection of the whole culture, while real cultures require subcultivation every 3 or 4 days (Section 2.3.6).

- The restriction of maintaining $\%I < 10\%$, which is observed in static real cultures, was not reproduced by the model (Sections 2.3).

- Limitations on static cultures appeared in the model as a consequence of geometrical limitations on the system (Section 3.3).

A3) - *extension*- Macroscopic constraints and geometrical limitations on static cultures.

INDISIM-RBC showed that several factors may influence the static *in vitro* cultivation of *P. falciparum*

- The shape of the hematocrit layer (HL) is determined by cell-to-cell interactions and by the shape and size of the culturing vial (Sections 3.1, 3.2 and 3.3).
- Cell density in the HL is not affected by the shape and size of the culture vial (Sections 3.1.2 and 3.3).
- Shape of the HL affects the performance of cultures under similar culturing conditions (Sections 3.1.2 and 3.3).
- The HL can be split into two sub-layers, each of them with a different parasite proliferation rate (Section 3.4). This allows for the reproduction and prediction of the performance of different static cultivation regimes:

  i) The depth of the upper sublayer is fixed by cellular features, by the spreading of the parasite and by substrate diffusion. The model cannot completely disentangle the contribution of each factor to the performance of real cultures.

  ii) The upper sublayer of the HL shows unrestricted propagation of the infection, while in the lower, a reduction of the number of IRBCs emerges from local diffusive limitations.

  iii) A region of fixed depth in the lower sublayer has reduced proliferation of the parasite due to the mechanism governing merozoite spreading.

Withal, the first general objective, to gain understanding and predictive capability of a specific kind of microbial system, the long-term *in vitro* culture of human RBCs infected with *Plasmodium falciparum* under different controlled conditions, has been achieved.

B) Methodological approach

B1) The pattern-oriented (PoM) strategy was followed to build the model.

PoM is a protocol for building models in the context of theoretical ecology described by Grimm et al (2006), which consists in proposing a basic model for the individuals that includes *only* those features and mechanisms relevant to the behavior observed in real *in vitro* culture systems; including *only* those constraints observed in the experimental systems at different levels of description and refining and rebuilding the model repeatedly to obtain a structurally realistic representation of a culture system. Its application of PoM in microbiology requires:

- The identification of the specific questions to pose. This means the detection and observation of patterns to reproduce, predict and explain, in collaboration with the field experts that are familiar with the specific features of the system to be modeled.
- The definition of the scales and level of description of the model. This requires determining the ultimate agents of the system (the microbial cell), which are the ones that actively govern the ongoing processes. Their characteristic scales will probably define the appropriate level of description of the model.
- The identification of the relevant features and variables in real systems. Microbes are very complex autonomous organisms, and they have multiple features that have no influence over a given process. It is essential to neglect these irrelevant features, and to stick to those explaining a pattern. This must be carried out even at the risk of loosing realism.

In following these steps, modeling became a continual process, and the development of models with increasing complexity, INDISIM-RBC versions *2Dv.1 2Dv.2* and *3D,* favored the use of assumptions with a biological basis over those with only empirical efficacy. As a consequence, we obtained a mechanistically rich model (Section 4.1).

B1) -*extension*- PoM favored the development of mechanistically rich models.

Mechanistically rich models in microbiology offer (Section 4.1):

- Versatility: models built following PoM are less dependent on the values of the parameters and rely more on the structure of the model. A single model can reproduce and explain a variety of complex behaviors.
- Simplicity: rules defined following PoM are the simplest found to reproduce a group of selected patterns. In this sense, models are transparent and intuitive because they include only those assumptions that must be considered to account for the observations
- Specificity: models are built to reproduce specific properties of real systems and they have predictive capability over similar systems in a wide range of situations. This can be used to explore technically unachievable experiments or situations.
- Generality: in order to correctly account for microbial ecosystems we must consider the local range of interactions, the temporal nature of entities and processes, the diversity among the population and noise in the deterministic rules describing the system.

B2) SWOT analysis of modeling strategies

Tackling microbial populations entails choosing between two complementary approaches: IbMs and PbMs.

- The SWOT analysis was used to evaluate the suitability of each approach (Section 4.2). It comprised the description of the advantages and drawbacks of IbMs and PbMs, and the definition of the optimal application and complementarity of these approaches.
- The joint application of IbMs and PbMs allowed a better approach to the diffusion process through the hematocrit layer in static cultures.

B3) Basic concepts of thermodynamics in biology.

The terms employed in physics and in biology differ. Stipulating the language is crucial in a multidisciplinary research to avoid misunderstandings (Section 4.3).

- A central concept is the (thermodynamic) state of a system: *equilibrium, stationary* or *variable*. The characterization of the state of a population is essential for comparing different systems and for increasing the predictive capability of models.
- At a system-level, one-time observations are usually not sufficient to assess the state of the population. This state strongly depends on the constraints on the system (e.g.: open/closed boundaries), and on the features we focus on.
- The application of the first law of thermodynamics and of the law of conservation of mass imposes some basic constraints on microbial systems (i.e. equations of balance). Application of the second law of thermodynamics imposes additional limitations (i.e. optimization of entropy S and minimization of entropy production).
- These laws can be used to study, analyze and predict the structure of the population regarding a set of individual characteristics. The obtained results can be compared to measurements of real systems and to simulation outcomes of IbMs.

The use of the theoretical, experimental and computational approaches allowed for the characterization of the state of microbial populations in general, and of malaria cultures, in particular.

- The distributions of post-invasion times among IRBCs ($q(t_{INF})$) of infection stages ($q(ICL)$) proved to be good indicators of the state of the infection in *Plasmodium falciparum* in *in vitro* cultures.
- The entropy related to population diversity $S(q(ICL))$ was measured and compared to the one corresponding to the stationary state $< S_{exp}(ICL) >$.
- The mathematical distance between the two entropies, $D(ICL)$, was used to assess the state of infected cultures.

B4) The ODD (Objective- Design Concepts-Details) protocol (Grimm et al., 2006) used to describe, analyze and communicate the model.

The ODD protocol is a standard method for presenting Individual-based Models (Sections 2.2 and 3.2). It requires a detailed and thorough description of many aspects of the model. It comprises the following tasks: i) identification of the questions to pose; ii) definition of the scales and level of description of the model; iii) identification of the relevant features and variables and iv) the critical review and repeated application of points (i) to (iii) as the model is developed. Some conclusions may be drawn from the application of this method.

- Sticking to this protocol when communicating INDISIM-RBC revealed weak points and inconsistencies in the model structure and in the defined rules.
- The use of a standardized convention and terminology favored the communication with the experts in malaria research and with other groups that develop IbMs, thus stimulating the improvement of INDISIM-RBC.
- The joint application of ODD and PoM in the process recursively rebuilding the model helped in the development of a mechanistically rich model.

Withal, IbMs have been used to tackle multiple spatial and temporal scales of microbial communities using fundamental mechanistic assumptions. However they have proved to be more difficult to develop, communicate and analyze than PbMs, and standardized protocols have been adapted to facilitate their sharing and application.

Techniques from statistical thermodynamics have been used to thoroughly characterize the population structure of the systems under study, and its relation with observations at a system level. The joint application of IbM and of statistical thermodynamics has provided results that can be contrasted with real systems through the recently available measurements of population heterogeneity (e.g. fluorescence, cytometry and scanning microscopy, among others).

The combination of measurements at a cellular level and the IbM perspective has lead to an increase in the understanding and in the predictive capability regarding the dynamics of specific microbial communities and cell cultures.

C) Application approach

C1) Quantitative criteria for the comparison of different RBC storage periods, parasite strains and commercial sources of culture medium.

In addition to the parameters usually employed to assess the performance of *P. falciparum* infected RBCS in i*n vitro* cultures ($\%I$ and GR), our group proposed the quantitative measurement of the population structure of IRBCs ($q(t_{INF})$, $q(ICL)$ and $D(ICL)$). The joint use of $\%I$, GR and $q(t_{INF})$ allowed us to draw some specific conclusions regarding the variation of the culturing conditions:

- The growth ratio is measured every 72 h or 96 h in the first of the long-term cultures found in the literature. GR_{72} takes values ranging over $3 < GR_{72} < 15$ and exhibits great variability (Section 2.3.3).
- No significant differences are found in GR_{72} when comparing different parasite strains, RBC sources or commercial culture mediums (Section 2.3.3).
- A significant decrease in the GR is observed for high RBC storage periods, which entails increasing RBC ages (t_{RBC}). When $t_{RBC} < 16\ days$, the parasite can grow unhindered. After 20 $days$ of RBC storage, the parasite can be maintained in culture with $GR \sim 1$. After 60 $days$, cultures are not viable, $\%I = 0$ (Section 2.3.3).
- Synchronization of the population does not significantly affect GR_{48} in short-term cultivation. However, continuous cultures show different performances depending on whether they are synchronous or asynchronous (Section 2.3.4).
- The distribution of post-invasion times $q(t_{INF})$ and infection stages $q(ICL)$ among IRBCs evolve through time and can be used to assess the state of the infection (Section 4.3.3).
- The latter can be used to define a mathematical distance between an observed population structure and a supposed stationary distribution $D(ICL)$, which provides criteria for the comparison of different culture systems.

C2) Mechanisms at an individual or local level and macroscopically observed patterns.

- The mechanism of merozoite egress that leads to best-fit simulations is membrane breakdown. This is consistent with recent microscopy observations and reproduces several patterns: GR, $\%I$, and the distribution of multiply-infected RBCs (Section 3.3.3). The system-level limitations on the spreading of merozoites is tackled in Table C4.
- The best-fit model of the infection susceptibility of healthy RBCs ($p_{INF}(t)$) is a monotonically time-decreasing function with an abrupt fall after three weeks. Appropriate storage of RBCs (control of storage times and protocols) is relevant to optimization of the performance of the cultures (Section 2.3.3).
- The simplest model of RBC metabolism that reproduces local need for nutrients considers the local concentration of nutrient in a range $\sim 10\ \mu m$ (Section 2.2). The amount of glucose does not significantly affect the culture on the whole (Section 2.3.5). Local limitations on nutrient availability may lead to reduced performance of static cultures (Section 3.5).
- *In vitro* cultures of IRBCs tend towards increasingly asynchronous populations. This can be attributed to the inherent diversity among IRBCs. Synchronization of the IRBC population may result from appropriate subcultivation protocols, combined with stage-specific treatments (Section 2.3.4).
- No significant differences are observed between the culture performance of continuous cultures with subcultivations every 48 h and 72 h. No significant differences are observed between static and suspended short-term cultures, but they do appear in the long-term continuous cultivation (Sections 2.3.3 and 2.3.6).
- Discrete agitation significantly affects the GR in long-term continuous cultivation. The effect depends on the parasite strain: some strains are favored by agitation, while others are hindered by the same process (Sections 2.3.3 and 2.3.6).

C3) Benefits of different culturing protocols.

- We trialled various methods for discrete and continuous medium renewal, but they led to no significant improvements in the harvests (Section 2.3.5).
- We assisted in the design of semi-automated cultures with periodic subcultivation programmed at regular time spans of 48 h and 72 h. These cultures were either discretely agitated or continuously agitated with magnetic stirrers to create suspended cultures. Suspended cultures performed better, in the long-term, than those discretely agitated or set still (Section 2.3.6).

C4) Geometry and dimensions of the culture vials that lead to optimal static cultivation of the parasite.

Results obtained from the detailed analysis of the macroscopic geometrical constraints on static cultures are outlined below:

- The shape of the *HL* depends on the macroscopic dimensions and composition of the culture vial. The *HL* can take the form of a sessile drop or it may fill all the bottom of the culturing vial. It may form convex or concave menisci.
- The dimensions and shape of the *HL* affect the culture performance. Optimal depths of the HLs range $0.18\ mm < HLD < 0.34\ mm$ and optimal distance between walls of the culturing vials must exceed $L > 2\ cm$.
- These geometric constraints can be attributed to the stratification of the HL. According to this model, the optimal development of the infection occurs solely in the upper surface of the hematocrit, in an interface with the culturing medium that spans around half a millimeter. Very small HLs are hindered because they impede the optimal spreading of the extracellular merozoite.
- The hypothesis of a stratified HL is consistent with the models for diffusive limitations in nutrient availability (Section 3.5).

Withal, the application of INDISIM-RBC has proved to be a more than adequate representation of the cultivation of *Plasmodium falciparum* and has provided with a better characterization of its dynamics. It has been used for improving current culture protocols and for designing novel strategies.

5.3 Perspectives and further work

The results presented in this thesis attempted to answer some questions regarding the *in vitro* cultivation of *Plasmodium falciparum*-infected RBCs, but they have also suggested new ones and opened new avenues for exploration. Below are a few of the many possibilities for further developments on this field.

To begin, we present some limitations of the current model and the work to be done in order to get over their hurdle. First, diffusion of nutrient in the hematocrit layer is crucial for the static cultivation of the parasite. INDISIM-RBC operates at spatial and temporal scales that do not allow an explicit model of the diffusion process that connects the information at a molecular level (random walk of particles) with the macroscopic observations. Greater efforts should be undertaken to obtain a fully mechanistic model of diffusion within the hematocrit layer. Second, resource-based limitations may not be caused by glucose or lactate, but by other key metabolites or harmful substances. A better model of RBC and IRBC metabolism should be built to tackle substrate limitations *in vitro.* Third, the current model defines an individual degree of stress, and an individual fragility, which do not correspond to any measurement, nor are they directly linked to any measure of the physiological state of the cell. A better description of cell damage and the possibilities and costs of cell repair should be developed. And finally, descriptions of static and suspended cultures require different sets of parameters in the model. The connection between the two sets could improve our understanding of the process of invasion of merozoites in healthy RBCs.

An improved version of INDISIM-RBC would potentially allow for the expansion of its application to real systems. Some of the experimental situations that could be tackled when the limitations of the current model have been addressed are described below. First, the model could be used to better understand and predict the results of the evolution of the infection after drug treatment. The relation between the effectiveness of a drug and the lag in the outbreak of the disease observed in treated culture is a central concern of *in vitro* cultivation. A model that provided better understanding of the relation between the effect of the drug at a cellular level and the behavior observed at a system level would be a boon to the pharmaceutical industry. Second, the use of the model could help in the establishment and design of automated bio-reactor-like culture systems. These systems would not require daily human supervision and would imply a great saving in resources and effort. Moreover, the continuous input and output fluxes of medium and treatments would be a better recreation of the conditions *in vivo.* Finally, *in vivo* cultivation of the parasite in murine models is a recent achievement of the pharmaceutical industry. An application of a mechanistically rich model to these culture systems would be a great

contribution to the understanding of the observed results.

To conclude, the development of the platform INDISIM-RBC paves the way to the potential application of the same methodology in similar systems. For instance, the model could be used to tackle the cultivation of other species of *Plasmodium*, such as *P. vivax,* which is to date non-culturable for reasons that remain unknown. In addition, a model similar to INDISIM-RBC but with major changes could be used to support the design of cultures of other forms of the parasite, such as the *gametocytes* and the *hepatocytes.* The same IbM methodology could be used to model and simulate other problems related to malaria such as the macroscopic propagation of the disease within the human population. However, such an epidemiological approach would require a completely different model with a different set of rules and agents.

Shifting to another topic, the consolidation of standard protocols for developing, communicating and analyzing IbMs in the context of theoretical and applied microbiology is of great interest for the short- and middle-term development of the research in this field. The work carried out by MOSIMBIO consists of individual-based modeling of various microbial communities and cell cultures, in parallel. The diversity in the areas of application of INDISIM offers an opportunity to analyze the general characteristics of the methodology and to review and formalize the procedures followed when trying to develop, analyze and communicate new IbM models. To date, such a strategy has provided good results, as it allows for the sharing problems and solutions from different fields of application, and it provides a holistic view of IbM in microbial systems, in general. Further research will also require the periodic review of the advances and results achieved in each field of application.

Appendix A

Glossary

List of acronyms

AbM	Agent-based Model (computational science)
ANOVA	ANalysis Of VAriance (statistical analysis)
ATP	Adenosine Triphosphate
C-N	Crank-Nicholson Method (computational science)
DBL	Diffusion Boundary Layer
DDT	Dichloro DiphenilTrichloroetane
DDW	Diseases from the Developing World
EMG	Experimental Microbiology Group
FCM	Free Culture Medium
FTCS	Forward Time Centered Space Method (computational science)
HL	Hematocrit Layer
HLD	Hematocrit Layer Depth
IbM	Individual-based Modeling (computational science)
IbMs	Individual-based Models (computational science)
INDISIM	INDividual DIScrete SIMulation

INDISIM-RBC	INDividual DIScrete SIMulation of Red Blood cells
IRBC	Infected Red Blood Cell
IRBCM	Plasma Membrane of the Infected Red Blood Cell
MC	Monte Carlo Method (computational science)
MD	Molecular Dynamics Method (computational science)
MMV	Medicines for Malaria Venture
MOSIMBIO	MOdeling and computer SIMulation of BIOlogical systems group
MSE	Mean Squared Error
NMR	Nuclear Magnetic Ressonance
ODD	Objectives-Design concepts-Details (protocol to communicate models)
PBC	Periodic Boundary Conditions
PbM	Population-base Model (computational science)
PHIL	Public Health Image Library
PoM	Patter-oriented Modeling (computational science)
PVM	Parasitophorous Vacuolar Membrane
RBC	Red Blood Cells
RBM	Roll Back Malaria
SbM	Systems-based Model (computational science)
SEIR	Susceptible-Exposed-Infectious-Resistant (epidemiological model)
SEM	Scanning Electron Microscopy
SIR	Susceptible-Infectious-Resistant (epidemiological model)
SIRS	Susceptible-Infectious-Resistant-Susceptible (epidemiological model)
SSE	Sum of Squared Errors
SWOT	Strengths-Weaknesses-Opportunities-Threats
UPC	Universitat Politècnica de Catalunya
WCW	White-Chapmann-Watt model (PbM on IRBC synchronization)
WHO	World Health Organization

List of abbreviations

2Dv.1	First version of INDISIM-RBC in two dimensions
2Dv.2	Second version of INDISIM-RBC in two dimensions
mero	extracellular merozoite
gluc	glucose
GR_h	Growth Ratio at h hours
HgB	Hemoglobin
INF	infected, infection
lac	lactate
max	maximum
min	minimum
obs	observed
sim	simulated
sc	spatial cell
su	simulation units
susp	suspended culture
stat	static culture
ts	time step
X_0	initial value of *'X'*
X_{My}	maximum value of *'X_y'*
$\overline{X}$	mean value of *'X'* averaged over the population
$\langle X \rangle$	mean value of *'X'* averaged over time
σ_X	standard deviation of the variable X, assuming a normal distribution

List of symbols

Chapter 1

R_0	reproduction number in epidemiological models
β	infection rate in epidemiological models
γ	recovery rate in epidemiological models
S, E, I, R	Susceptible, Exposed, Infectious and Recovered sub-populations in epidemiological models
$\%H$	hematocrit
$\%$	parasitaemia
f	fraction of daily medium renewal in *continuous-flow* cultivation

Chapter 2

ICL: 0, R, T,	Healthy, Ring, Trophozoite, Schizont and Fragmenter stages
S, F	of the RBC infection cycle
$P_{inf}(t)$,	parameters to set the susceptibility of individual RBCs to
P_{max}, P_{min}	invasion by extracellular merozoites
IM	metabolic stress index
P_{death}	death probability of RBCs
t_{RBC}, t_{CULT}	age and time in culture of RBCs
t_{MRBC},	maximum age and time in culture of RBCs
t_{MSUS}	
t_{INF}, t_{MINF}	post-invasion time and maximum duration of IRBCs
t_{ICL}, τ_{ICL}	duration and post-invasion times of each infection stage
t_{mero}	duration of merozoites in the extracellular medium
u, U_{in}, U_{eff}	uptake needs, instant uptake requisite and effective uptake
$f(t_{INF})$	factor governing the uptake needs
$\widetilde{D}_s$	diffusion coefficient of a substance s
w_{ijk}	weights of the neighboring cell (i, j, k) to local diffusion
nn	nearest neighboring spatial cells
C_s	local concentration of substance s
(x, y, z)	coordinates of a spatial cell
ρ_J, n_J	cell density, and number of cells of species J
v_{sed}	settlement velocity of RBCs in suspended cultures
$T_{agitation}$	agitation period
$p_{survival}$	probability for an IRBC to survive a complete infection cycle
ν	fraction of the RBC population replaced in subcultivation
$q(t_{INF})$	distribution of post-invasion times in IRBCs
f_{ts}, p_{ts}, n_{ts}	$q(t_{INF})$ of present, previous and next infection cycles at a time step ts

Chapter 3

$f_P(k, \lambda)$	Poisson distribution function of k events occurring with average appearance λ
RS, FCC, HD	spatial configurations within the hematocrit layer
pf	packing factor of RBCs
S	surface of the culturing vial covered by the HL
H_{total}	thickness of the culture system
V	culture volume
V_J	volume occupied by an individual entity of type J
ϵ	binding energy per unit surface of individual RBCs
L_C	capillary length of the HL
L_{EXC}	calculated boundary of the HL that excludes propagation of the infection
L	separation between the walls of the culturing vial
p_{fall}	velocity settlement
h, h_1, h_2	parameter is the population based models of the HL
D_{eff}	effective average diffusivity in the hematocrit layer
D_0	average diffusivity in the culturing medium
ϵ_w	fraction of culture medium within the HL
K_{eff}	kinetic constant for minimum glucose uptake
ψ	temporal scale of diffusion in the HL
t_{90}	time until reaching the stationary substrate distribution in the HL
d_0	depletion depth in the HL at the stationary state
d_{RBC}, d_{IRBC}	maximum depths for RBC and IRBC viability
ϕ	Thiele modulus
ϕ_0	Thiele modulus for zero-order uptake kinetics
ϕ_1	Thiele modulus for first-order uptake kinetics
δ	relative difference in substrate concentration
cf	diffusion calibration factor in the explicit methods

Chapter 4

L	spatial scale of the system
T	temporal scale of the system
λ	lag parameter of a bacterial culture
$\mathcal{D}$	difference between $q(m_i)$ of instant and exponential growth in bacterial cultures
m_i	class of biomass
$q(m_i)$	biomass probability distribution of a bacterial culture
q_i	general form of a normalized probability distribution in the population
i	one of the N classes used to set q_i
S	thermodynamic entropy
$\mathcal{S}$	Shannon entropy,/ diversity in the population
β	Lagrange multiplier
Z	normalization factor of the probability distribution function
$\widetilde{q_S}$	distribution function that maximizes diversity
μ	mean post-invasion time of the IRBC population
σ	deviation of post-invasion times of the IRBC population

Appendix B

Experiments carried out by EMG-GSK

The general culturing procedure follows the recomendations of the MR4 standardized protocols described in Section I:A in Methods for Malaria Research (MR4, 2008). Some especifications of the particular conditions of the experiments carried out by the EMG-GSK presented in this work are listed below.

Trial run of different serums

Outline

In order to select the most appropriate source of serum, the product of three different commercial providers were used for the one-week cultivation of the same strain of the parasite.

Experimental methods

The experiment comprises eight trial runs comparing four different commercial human sera (provided by the trademarks Cambrex, Biorreclamation, Sigma-Aldrich and PAA) to complement the RPMI culture medium. Cultures are inoculated with $\%I_0 = 0.5\%$ and $\%I_0 = 1\%$, and either with synchronous (showing other four combinations of the only ring forms of the IRBCs) or asynchronous inocula (consisting of rings and mature trophozoites), and maintained during seven days at most. Medium provided by Sigma-Aldrich is

tested with the four possible combinations of inocula (That is, synchronous/asynchronous with $\%I_0 = 0.5\%$ and $\%I_0 = 1\%$). Medium provided by Cambrex is tested with two degrees of synchronism and only for inocula with $\%I_0 = 1\%$. Mediums provided by Bioreclamation and PAA are tested only for synchronous inocula with $\%I_0 = 1\%$.

Samples of 4 ml of culture are maintained in flat-bottom plastic flasks of 25 cm^2 base surface under low oxygen atmospheres. Parasitaemia is measured daily through the extraction of samples with minimal perturbation of the culture. Subcultivations are carried out during the first 4 days, during the first two cycles of infection. When a first threshold parasitaemia is exceeded ($\%I > 3\%$) a subculture is inoculated with a sample of the harvest with a dilution ratio of $\frac{1}{2}$ in volume. When the parasitaemia exceeds a second threshold value ($\%I > 4\%$), subcultibvation is carried out with a $\frac{1}{4}$ dilution ratio. The experimental protocols are summarized in Table B.1.

Input measurements	Initial hematocrit ($\%H_0$), initial parasitaemia ($\%I_0$)
RBC source	Human AB-, supplied by Spanish Red Cross. No storage.
Initial Hematocrit ($\%H_0$)	5 % Hematocrit
Parasite source	*P. falciparum* 3D7A supplied by the MR4
Initial Parasitaemia ($\%I_0$)	0.5 % - 1 %
Culture medium [glucose] [lactate] pH	RPMI 1640, 25 mM HEPES, 10% human serum, 3 % Hypoxanthine. 2.67 mM/l 0 mM /l 7.2
Culturing method Medium renewal Culturing conditions Subcultivation Dilution ratio	Static cultivation 100% discrete renewal every 48h. Temperature: 37º C; Atmosphere: 5% O2, 5% CO2, 90% N2 Every 48, 72 or 96 hours, triggered by excessive parasitaemia ($\%I > 3\%$) 1/2 or 1/4 of culture volume, depending on the observed parasitaemia
Output measurements	Parasitaemia ($\%I$), through optical microscopy and Giemsa stained thin blood smears.
Data analysis	Worksheet and GrapPad Prism

Table B.1: *Specifications of the experimental procedure for the comparison of different culture sera.*

Conclusions

i) No significant differences are found between the serums from different commercial providers.

ii) No significant differences are found between cultures with synchronous and asynchronous inocula.

Agitation regimes in short-term culture systems

Outline

In order to study the geometric factors affecting the short-term proliferation of the parasite, *P. falciparum* is cultured in closed systems with several agitation patterns: no agitation of the culture, discrete agitation every 24 h, 48 h and 72 h, and continuous agitation.

Experimental methods

The experiment comprises ten trial runs comparing five different agitation regimes in a closed culture during four days, with no subcultivation of IRBCs, nor medium renewal. Parastiaemia is assessed every day with gentle sampling of the culture system (with minimal perturbation of the hematocrit layer). A static and three discretely agitated culture systems cultures with $V = 50\ ml$ are maintained in T-flasks with 25 cm^2 base surface. Discretely agitated cultures are shaken by hand with different agitation periods, in order to re-suspend the cells and modify the spatial structure of the hematocrit layer, but with minimal damage of RBCs.

A static and two continuously agitated cultures with $V = 40\ ml$ are kept in 100 ml glass bottles with 20 cm^2 base surface under similar conditions. The two replicas of the static cultures are carried out using different culturing sera, without significant difference in the resulting performance.

The source of RBCs, parasites and culture medium are detailed in Table B.1, together with the specifications of the culture medium and the culturing conditions. The specifications of the culture method (volume and agitation method) for the different regimes are outlined in Table B.2

Input measurements	Initial hematocrit ($\%H_0$), initial parasitaemia ($\%I_0$)
Output measurements	parasitaemia ($\%I$), through optical microscopy and Giemsa stained thin blood smears.
Initial Hematocrit ($\%H_0$)	5 % Hematocrit
Initial Parasitaemia ($\%I_0$)	0.5 %
Culturing method	**static cultivation in a closed system**
Volume:	50 ml
Base surface:	25 cm^2
Agitation regime:	- static (no agitation) - discrete shaking every 24 h, 48 h and 72 h
Culturing method	**suspended cultivation in a closed system**
Volume:	40 ml
Base surface:	20 cm^2
Agitation regime:	Stirred with magnetic stirrer v_{stirr} = 30 R.P.M. (revolutions per minute)

Table B.2: *Specifications of the experimental procedure for the comparison of different discrete agitation protocols.*

Conclusions

i) No significant differences are found between the static cultures and those agitated at discrete periods. Continuously agitated culture systems show increased performance.

ii) All the culture systems are exhausted after 4 days(at most) in a closed culture.

Renovation regimes for static cultivation

Outline

In order to study the viability of continuous cultivation of the parasite, *P. falciparum* is cultured under static conditions with two actuation patterns: simultaneous medium renewal and subcultivation of the hematocrit every 48 hrs and daily medium renewal with dilution of cultures every 72 hrs.

Experimental methods

The experiment comprises four (10-days) trial runs comparing two different culturing protocols with fixed medium renewal and subcultivation periods. Medium renewal re-

places 100% of the culture medium and the dilution fraction at subcultivation is variable and set to obtain a resulting parasitaemia of 0.5%.

Samples of 5 ml of culture are maintained in glass bottles of 20 cm^2 base surface under low oxygen atmospheres. Parasitaemia is measured daily through the extraction of samples with minimal perturbation of the culture. Daily parasitaemia measurements of the samples extracted from the two replica of each trial are compared with ANOVA. The experimental protocols are summarized in Table B.3.

Input measurements	Hematocrit, initial parasitaemia ($\%I_0$)
RBC source	Human AB-, supplied by Spanish Red Cross. No storage.
Initial Hematocrit ($\%H$)	5 % Hematocrit
Parasite source	*P. falciparum* 3D7A supplied by the MR4
Initial Parasitaemia ($\%I_0$)	0.5 %
Culture medium	Specified in TableB.1
Culturing method	**static cultivation**
Medium renewal	100% discrete renewal, daily or every 48h.
Culturing conditions	Temperature: 37º C; Atmosphere: 5% O2, 5% CO2, 90% N2
Subcultivation	Every 48 or 72 hours
Dilution ratio	Variable, depending on the observed parasitaemia
Output measurements	Parasitaemia ($\%I$), through optical microscopy and Giemsa stained thin blood smears.
Data analysis	GrapPad Prism and ANOVA

Table B.3: *Specifications of the experimental procedure for the comparison of different cultivation regimes.*

Conclusions

No significant differences are found between the cultures with a 48 h-period for both medium renewal and subcultivation, and those cultures with daily medium renewal and subcultivations every 72 h.

Renovation regimes for automatic cultivation

Outline

In order to determine the appropriate dilution of subcultures for the continuous automatic cultivation of the parasite, *P. falciparum* is cultured with two agitation patterns: under static undisturbed cultivation regime and under suspension culturing conditions, through the continuous agitation of the hematocrit with a magnetic stirrer.

Experimental methods

The experiment comprises four (10-days) trial runs comparing two different culturing protocols with fixed medium renewal and subcultivation periods. Medium renewal replaces 100% of the culture medium and the dilution fraction at subcultivation is variable and set to obtain a resulting parasitaemia of 0.5%.

Samples of 5 ml of culture are maintained in glass bottles of 20 cm^2 base surface under low oxygen atmospheres. Parasitaemia is measured daily through the extraction of samples with minimal perturbation of the culture. Daily parasitaemia measurements of the samples extracted from the two replica of each trial are compared with ANOVA. The experimental protocols are summarized in Table B.4.

Input measurements	Hematocrit, initial parasitaemia ($\%I_0$)
RBC source	Human AB-, supplied by Spanish Red Cross. No storage.
Initial Hematocrit ($\%H$)	5 % Hematocrit
Parasite source	*P. falciparum* 3D7A supplied by the MR4
Initial Parasitaemia ($\%I_0$)	0.5 %
Culture medium:	Specified in Table B.1
Culturing method	**static cultivation**
Medium renewal	100% discrete renewal, daily or every 48h.
Culturing conditions	Temperature: 37º C; Atmosphere: 5% O2, 5% CO2, 90% N2
Subcultivation	Every 48 hours
Dilution ratio	Variable, depending on the observed parasitaemia
Output measurements	Parasitaemia ($\%I$), through optical microscopy and Giemsa stained thin blood smears.
Data analysis	GrapPad Prism and ANOVA

Table B.4: *Specifications of the experimental procedure for the comparison of different cultivation regimes.*

Conclusions

i) Static culture systems can be continuously maintained during 5 cycles when the fraction of subcultured cells to the fresh healthy RBCs at subcultivation is equal to $\nu = \frac{1}{3}$.

ii) Stirred culture systems can be continuously maintained during 5 cycles when the fraction of subcultured cells to the fresh healthy RBCs at subcultivation is equal to $\nu = \frac{1}{11}$.

Long-term semi-automatized cultivation

Outline

In order to study the semi-automatized cultivation of the parasite, *P. falciparum* is cultured during one month using the methods and fractions presented in Section B.

Conclusions

Both static and suspended culture systems can be continuously maintained during one month. The population-based model defined in Section B correctly predicts the experimental behavior.

Measurement of the depth of the hematocrit layer and of the RBC sedimentation rate

Outline

In order to study the spatial structure of the hematocrit layer, the settlement of RBCs and the dimensions of the hematocrit layer are analyzed for different geometries (base surface) of the culturing vials.

Experimental methods

Different culture devices are filled with the same volume of culture $V = 20\ ml$ at 5% hematocrit, simulating ordinary culturing conditions. RBC sedimentation lasts until the culture is split into two phases: the hematocrit layer and the completely transparent culture medium. Height of RBC layer is measured once red blood cells finish settling down. RBC packing factor (pf) is calculated as $pf = \frac{HLD_{obs}}{HLD_{calc}}$. Where HLD_{obs} is the measured HLD and $HLD_{calc} = 0.05 \cdot H$, being the total culture height. Sedimentation rate (v_{sed}) is calculated with H and the span until reaching the settled hematocrit layer (t_{sed}), $v_{sed} = \frac{H}{t_{sed}}$.

Input measurements	Culture volume (V), hematocrit ($\%H$), base surface of the vial (S_{base})
Output measurements	Duration of the settling process (t_{settle}), height of the culture (H), height of the hematocrit layer (H_{obs})
Culture volume	$V = 20\ ml$
Hematocrit	$\%H = 5\%$
Glass tube	$S_{base} = 2.41\ cm^2$
Silicone tube 1	$S_{base} = 0.5\ cm^2$
Silicone tube 2	$S_{base} = 0.28\ cm^2$

Table B.5: *Specifications for the analysis of the hematocrit settlement for different geometries*

Conclusions

i) The observed sedimentation rate of RBCs is $v_{sed}(RBC) = (1.1 \pm 0.3) \cdot 10^{-5} cm/s$.

ii) The packing factor of RBCs within the settled the hematocrit layer is $pf = 0.87 \pm 0.07$

Multiple invasion of IRBCs

Outline

In order to study the spreading of the extracellular merozoite, the number of IRBCs with multiple invasions is counted for several samples of static *in vitro* cultures.

Experimental methods

Cultures of RBCs are carried out following the protocols described by the MR4 and specified in Section . The number of IRBCs with multiple invasions is measured for 51 samples of the culture through optical microscopy and Giemsa stained thin blood smears. The statistics of the counts are presented in Table B.6.

Statistic	**1-Ring**	**2-Rings**	**3-Rings**	**4-Rings**	**5-Rings**	**6-rings**	**+6-Rings**
Mean	66.37	22.43	9.31	1.65	0.23	0.06	0
St dev.	13.31	10.46	8.28	2.61	0.96	0.42	0

Table B.6: *Mean value and standard deviation of each category of multiparasited IRBCs after 51 samples of a static in vitro culture of P. falciparum.*

Effect of the base surface of the hematocrit layer on the culture performance

Outline

In order to study the influence of walls distance in static cultures, *P. falciparum* 3D7A cultures with different separation between walls are maintained in plastic dishes during two days and their performance is assessed ($\%I$ and GR).

Experimental methods

Closed static cultures with the same volume are implemented in petri dishes with different geometries. Several plastic pieces are attached to the plates in such way that they maintain the total base surface (S_{base}) and depth of the hematocrit layer (HLD) but measure different distances between side walls, and therefore split the hematocrit layer into areas of different culturing surfaces (S). Modified plates cover distances: $L = 9\ cm$ (plate with no pieces), $L = 4\ cm$ (2 separators), $L = 3\ cm$ (3 separators), $L = 1\ cm$ (5 separators) and $L = 0.75\ cm$ (7 separators). Each plate with separators leaves an enclosure of fixed volume, to compare the behavior of different trials (control area). Another plate, covered with the glue used to attach the pieces but without separators, is used in comparison with the plate with no separators as a toxicity control. The specifications for the culture method are described in Section B, the specifications of the culture system are shown in Table B.7. The specific geometries and the obtained results are presented in Table 3.3.

Input measurements	Culture hematocrit $\%H$, initial parasitemia $\%I_0$
Output measurements	Parasitemia (%), growth ratio (GR_{48})
Culture volume	$V = 16.5\ ml$
Total base surface	$S_{base} = 63.6\ cm^2$
Depth of the hematocrit layer	$HLD = 0.19\ mm$
Culture initialization	$\%H = 5\%$; $\%I = 0.5\%$

Table B.7: *Specifications for the P series: static closed cultivation of P. falciparum.*

Conclusions

i) Material used to attach separators to petri dish has not effect in multiplication ratios GR_{24} or GR_{48}.

ii) The performance observed in the control areas is the same for every trial culture.

iii) Separation distance between walls significantly affects the culture performance. Growth ratio decreases with decreasing distances.

Effect of the depth of the hematocrit layer on the culture performance (*W series*)

Outline

In order to study the influence of the thickness of the hematocrit layer in static cultures, *P. falciparum* 3D7A cultures with different culture volumes are maintained in plastic dishes during 10 days and their performance is assessed.

Experimental methods

Static cultures with different volumes (V) are implemented in 6-well petri dishes with the same total base surface (S_{base}). Therefore, the thickness of the hematocrit layer (HLD) varies from trial to trial. Static cultivation is maintained during 14 days, following the procedures proposed by the MR4. The specifications for the culture trials are outlined in Table B.8. The specific geometries and the obtained results are presented in Table 3.3.

Input measurements	Culture hematocrit $\%H$, initial parasitemia $\%I_0$
Output measurements	Parasitemia (%), growth ratio (GR_{48})
Culture initialization	$\%H = 5\%$; $\%I = 0.5\%$
Total base Surface	$S_{base} = 63.6\ cm^2$
Medium change	Every 48 hours
Subcultivation	Every 48 hrs, 96 hrs after inoculation
Dilution	Manual, to initial parasitaemia $\%I = 0.5\%$

Table B.8: *Specifications for the in the W series: static cultivation of P. falciparum following the candle-jar method.*

Conclusions

i) The performance of the culture significantly varies with the thickness of the hematocrit layer (HLD).

ii) Cultures with $HLD < 0.06\ mm$ show non viable or extremely hindered performances. For these geometries, the hematocrit sticks to the walls of the culturing device and does not form a thin layer.

Effect of the depth of the hematocrit layer on the culture performance (*B series*)

Outline

In order to study the influence of the thickness of the hematocrit layer in static cultures, *P. falciparum* 3D7A cultures with different culture volumes were maintained in glass bottles during 14 days and their performance was assessed.

Experimental methods

Static cultures with different volumes (V) are implemented in 100-ml glass bottles with the same total base surface (S_{base}). The hematocrit layer forms a sessile drop that does not cover S_{base}. The real culturing surface (S) and thickness of the hematocrit layer (HLD) vary from trial to trial with V. Static cultivation is maintained during 14 days following the procedures proposed by the MR4. The specifications for the culture trials are outlined in Table B.9. The specific geometries and the obtained results are presented in Table 3.3.

Input measurements	Culture hematocrit $\%H$, initial parasitemia $\%I_0$
Output measurements	Parasitemia (%), growth ratio (GR_{48})
Culture initialization	$\%H = 5\%$; $\%I = 0.5\%$
Total base Surface	$S_{base} = 19.64\ cm^2$
Medium change	Every 48 hours
Subcultivation	Every 48 hrs, 96 hrs after inoculation
Dilution	Manual, to initial parasitaemia $\%I = 0.5\%$

Table B.9: *Specifications for the B series: static cultivation of P. falciparum following the candle-jar method.*

Conclusions

i) The culture performance does not depend on the total culture volume (V).

ii) The performance of the culture significantly varies with the thickness of the hematocrit layer (HLD).

Appendix C

Statistical tests and tools

Tests of similarity

There are several standard statistical tests to compare different distribution functions and data sets. Some of them are briefly outlined below. For a detailed description and further technical considerations, a wide amount of specialized literature is available, see for instance: (Zar, 1999; Corder and Foreman, 2009). Further information can also be found on-line in open access handbooks, for instance the *NIST/SEMATECH e-Handbook of Statistical Methods*:
http://www.itl.nist.gov/div898/handbook/index.html,
and the e-Handbook of Biological Statistics: http://udel.edu/~mcdonald/statintro.html.

- ANOVA: ANalysis Of VAriance is a general statistical procedure that can be used to test the hypothesis that the means among two or more groups of data are equal, under the assumption that the sampled populations are normally distributed. The guiding principle behind ANOVA is the decomposition of the sums of squares of the data samples and mainly focuses on the sum of differences, or SSE.
  A *factor* affecting the datasests is any modification on the source of data, which can be independently controlled. ANOVA can be used to assess the effect of either one experimental factor or of multiple factors simultaneously. If a single factor is evaluated, this technique allows measuring the similarity of the means, but if multiple factors are evaluated, ANOVA gives information regarding the similarity of the means and the correlation of the factors.

- SSE and MSE: The Sum of Squared Errors (or differences between two samples)

and the Mean Square Error are two basic *objective functions,* or measures of the similarity of two distributions, and are also employed to evaluate the likelihood of a particular observation or simulation outcome.
When performing any statistical test, the hypothesis regarding similarity or likelihood with a prefixed degree of confidence is usually assessed by relating the measured SSE or MSE with a parameter defined by the test ($p - value$). Then, this parameter is compared with its expected value when the hypothesis of likeliness is true with a certain degree of confidence. Finally, the tested hypothesis is rejected if the calculated $p - value$ exceeds the expected one.

- Student's t-test (t-test): It is usually employed to compare one nominal variable with the mean value of one measured variable. Slight variations of the t-test are used to compare a single observation to a sample, or to compare a sample mean to a theoretical mean. The t-tests assume that the observations within each group are normally distributed and the variances are equal in the two groups.
  T-test is not particularly sensitive to deviations from these assumptions, but if the distribution of values within the samples strongly differ from normal distributions, then other tests shall be used.

- Kolmogorov-Smirnov test (KS-): The Kolmogorov-Smirnov test is used to decide if a sample comes from a population with a fully specified distribution function. A fully specified distribution is completely characterized by an analytical function. The comparison is not carried out by using the SSE or MSE but the cumulative distribution function. The hypothesis regarding the distributional form is rejected if the test statistic is greater ($p - value$) than a critical value obtained from a table.

- Kruskal-Wallis test (KW-test): The Kruskal-Wallis Test is used to compare populations with unknown distributions. It is a non parametric test, this meaning that it compares raw data rather than their mean or variance. It tests the hypothesis that k samples from different populations actually originate from similar populations.

Methods for parameter estimation

Parameter estimation is essential to set and to evaluate the predictive capability of a computational modeling. Basically, it consists on determining the set of parameters in the model that provide a response most similar to the one observed in real systems. A parameter estimation problem is usually formulated as the optimization of an objective function between the observed data and simulation outcome under fixed conditions, while

varying just a few experimental factors (e.g.: the best fit between the simulated and observed growth ratios for different temperatures).

Because of different objective functions and optimization techniques, parameter estimation can be solved in many ways. The most usual objective functions are MSE and SSE. Three optimization techniques are compared below. They correspond to three methods tested by the group in previous research (Prats, 2008).

- Grid search: it consists in exploring the phase space of the parameters of the model step by step. It evaluates all the possible combinations of parameters. For each combination, a set of simulations is run. Finally, we select the set of parameters that provides the best-fitting outcome to the experimental observations. This method is not effective because many sense-less and bad performing combinations of parameters are evaluated. In fact, grid search becomes unfeasible for IbM, where there are many parameters to be evaluated and where each simulation lasts for several minutes.

- NMTA: the Nelder Mead Threshold Accepting is a direct search method to optimize the exploration of the parameter space. The basic unit of this method is the *simplex*: a geometric figure in an n-dimensional space that is a convex hull of n+1 vertexes, with each vertex representing a certain combination of the n parameters to be estimated. The value of the objective function is assessed in each simplex vertex, and a new simplex that is closer to the objective function is constructed in each step.
  The rules for constructing each new simplex are based on a set of geometric operations: *reflection, expansion, outside contraction, inside contraction* and *shrinkage*. The value of the objective function determines whether new vertexes are accepted or rejected. If a new vertex is accepted, a new simplex is constructed by rejecting the worst existing vertex, always according to the objective function values. The selected combination of parameters can be set after a certain number of simplex explorations.
  In order to avoid tending towards local minima of the objective function rather than towards an absolute minimum, new vertices are accepted even if the value of the objective function is slightly greater than the preceding values, unless the acceptance threshold is exceeded.

- NEWUOA: The NEWUOA algorithm fits a quadratic approximation of the objective function, which is valid inside a trust region. NEWUOA finds the minimum of the quadratic within the trust region and compares it to the real cost: if the decrease

in the real value is less than the decrease predicted by the model, the radius of the trust region is scaled down; otherwise, the radius is not changed. The optimization is achieved by successive quadratic approximations. The algorithm finishes when the radius of the trust region achieves a lower prefixed threshold.

Computational methods for solving systems of equations with finite differences

Partial differential equations are used to formulate problems involving several variables, for instance, to model transport phenomena. Unfortunately, closed analytical expressions for their solutions can be found only in very special circumstances, which are mostly of limited theoretical and practical interest. Therefore, such kind of problems are usually tackled by means of computational methods (Crank, 1990; Hoffmann and Chiang, 2004). In this thesis we compare two very basic approaches to the diffusion of substrate, an explicit method, Forward-Time-Centered-Space (FCTS) and the Crank-Nicholson approach.

Appendix D

Own publications

Published works that comprise this thesis

JCR journals

- **Ferrer, J.**, Prats, C., López, D. and Vives-Rego, J., 2009. Mathematical modelling methodologies in predictive food microbiology: a SWOT analysis. *Int. J. Food Micro.* 134 (1-2), 2-8.

- **Ferrer, J.**, Rosal, M. D., Vidal, J. M., Prats, C., Valls, J., Herreros, E., López, D. and Gargallo, D., 2008. Effect of the hematocrit layer geometry on *Plasmodium falciparum* static thin-layer *in vitro* cultures. Malaria Journal 7: 203.

- **Ferrer, J**., Prats, C. and López, D., 2008. Individual-based Modelling: an essential tool for microbiology. *J. Biol. Phys.* 34 (1-2), 19-37.

- **Ferrer, J.**, Vidal, J., Prats, C., Valls, J., Herreros, E., López, D., Giró, A. and Gargallo, D., 2007. Individual-based model and simulation of *Plasmodium falciparum* infected erythrocyte *in vitro* cultures. *J. Theor. Biol.*, 248 (3), 448-459.

Oral communications in international conferences

- Prats, C., **Ferrer, J.**, López, D. and Vives-Rego, J, 2009. Mathematical Modelling Methodologies in Predictive Food Microbiology: a SWOT Analysis. The 21st International ICFMH Symposium "Evolving Microbial Food Quality and Safety" (FoodMicro 2008)

- **Ferrer, J.**, Prats, C. and López, D., 2008. INDISIM, an Individual-based Model to simulate microbial cultures. SwarmFest 2008.

Publications in international peer reviewed journals of conferences

Abstracts of poster presentations

- **Ferrer, J.**, Herreros, E., Valls, J., Prats, C., López, D. and Gargallo-Viola, D., 2005. Individual based modelling and simulation of *Plasmodium falciparum* Infected erythrocytes *in vitro* cultures (3D model). In: The American Journal of Tropical Medicine and Hygiene, volume 73 number 6 (Suppl.), p. 113-114 [American Society of Tropical Medicine and Hygiene (ASTMH) 54th Annual Meeting, Washington (USA), December 2005]

- **Ferrer, J.**, Prats, C., Valls, J. and López, D., 2005. Individual Based Model and simulation of *Plasmodium falciparum* Infected Erythrocytes *in vitro* cultures (2D model). In: Abstract book. Medicine and health in the tropics, p. 235 [XVIth International Congress for Tropical Medecine and Malaria, Marseille (França), September 2005]

Additional works during this period

JCR journals

- Prats, C., **Ferrer, J.**, Gras, A. and Ginovart, M., 2010. Individual-based modeling and simulation of microbial processes: yeast fermentation and multi-species composting. *MCMDS, in press.*

- Prats, C., Giró, A., **Ferrer, J.**, López, D. and Vives-Rego, J., 2010. On the evolution of cell size distribution during bacterial growth cycle: experimental observations and Individual-based Model simulations. *AJMR,* 4 (5) 400-407.

- Prats, C., Giró, A., **Ferrer, J.**, López, D. and Vives-Rego, J., 2008. Analysis and IbM simulation of the stages in bacterial lag phase: basis for an updated definition. *J. Theor. Biol.*, 252, 56-68.

- Prats, C., López, D., Giró, A., **Ferrer, J.** and Valls, J., 2006. Individual-based modelling of bacterial cultures to study the microscopic causes of the lag phase. *J. Theor. Biol.*, 241, 939-953.

Publications in international peer reviewed journals of conferences

Full papers

- Prats, C., **Ferrer, J.**, Flix, B., Giró, A., López, D. and Vives-Rego, J., 2007. Evolution of biomass distribution during bacterial lag phase through flow cytometry, particle analysis and Individual-based Modelling. In: G. Nychas et al. (Ed.), 5th International Conference on Predictive Modelling in Foods. Proceedings, p. 301-304, Agricultural University of Athens, Greece (ISBN: 978-960-89313-7-4) [5th International Conference on Predictive Modelling in Foods (ICPMF 2007), Athens (Greece), September 2007]

- Prats, C., **Ferrer, J.**, Giró, A., López, D. and Valls, J., 2006. Spatial properties in Individual-based Modelling of microbial systems. Study of the composting process. In: Méndez-Vilas, A. (Ed.), Modern Multidisciplinary Applied Microbiology. Exploiting Microbes and their Interactions, p. 461-465, Wiley-VCH Verlag, Weinheim, Germany (ISBN: 3-527-31611-6) [International Conference on Environmental, Industrial and Applied Microbiology (BioMicroWorld 2005), Badajoz (Spain), April 2005]

Abstracts

- Prats, C., Giró, A., **Ferrer, J.**, López, D., Valls, J., and Vives-Rego, J., 2006. A Mathematical Analysis of the Stages in Bacterial Lag Phase. In: Food Micro 2006. Food safety and food biotechnology: diversity and global impact. Book of abstracts, p.550 [The International ICFMH Symposium. Food safety and food biotechnology: diversity and global impact (Food Micro 2006), Bologna (Italy), August 2006]

- Prats, C., Giró, A., **Ferrer, J.**, López, D. and Valls, J., 2006. A physical interpretation of the microbial growth rate temperature dependence. In: XX Sitges Conference. Physical Biology: from Molecular Interactions to Cellular Behavior. Book of abstracts, p. 161-163 [XX Sitges Conference. Physical Biology: from Molecular Interactions to Cellular Behavior, Sitges (Spain), June 2006]

- Prats, C., **Ferrer, J.**, Giró, A., López, D. and Valls, J., 2005. Spatial properties in Individual based modelling of microbiological systems. Study of the composting process In: BioMicroWorld-2005 Book of Abstracts, p. 642 [International Conference on Environmental, Industrial and Applied Microbiology (BioMicroWorld 2005), Badajoz (Spain), April 2005]

Links

The following link leads to the web of the group MOSIMBIO: http://mosimbio.upc.edu/
In this web, the interested reader may find some updated and complementary material:

- Outline of the model and simulator INDISIM-RBC
- Sample simulations, executable and code source of INDISIM-RBC.*v2D.2*
- Sample simulations, executable and code source of INDISIM-RBC.*v3D*
- Additional models and simulators on malaria spreading.

Bibliography

Åkerlund T., Nordström K. & Bernander R., 1995. Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. *J. Bacteriol.* 177 (3), 6791-6797.

Abbey, H., 1952. An examination of the Reed Frost theory of epidemics. *Human Biol.* 24, 201-233.

Ahmed, E. and Agiza, H.N., 1998. On modeling epidemics. Including latency, incubation and variable susceptibility. *Physica A.* 253 (3-4), 347-352.

Aiello O.E., Haas V.J., daSilva M.A.A. and Caliri, A., 2000. Solution of deterministic-stochastic epidemic models by dynamical Monte Carlo method. *Physica A.* 282 (3-4), 546-558.

Alder, B. J. and Wainwright, T. E., 1959. Studies in Molecular Dynamics. *I. General Method. J.Chem Phys.* 31 (2), 459-466.

Allan, R.J., Rowe, A. and Kwiatkowski, D., 1993. *Plasmodium falciparum* varies its ability to induce tumor necrosis factor. *Infect. Inmmunol.* 61, 4772-4776.

Alpkvist, E., Picioreanu, C., van Loosdrech, M.C.M. and Heyden A., 2006. Three-dimensional biofilm model with individual cells and continuum EPS matrix. *Biotechnology and Bioengineering* 94, 961-979.

Anderson, R.M. and May, R.M., 1979. "Population Biology of Infectious Diseases: Part I." *Nature* 280, 361-367.

Bailey, V.A. , 1933. On the interaction between several species of hosts and parasites *Proc. Roy. Soc. A* 143 (4), 75-88.

Ballou, W.R. and Cahill, C.P., 2007. Two decades of commitment to malaria vaccine development: GlaxoSmithKline biologicals. *Am. J. Trop. Med. Hyg.* 77 (Suppl 6), 289-295.

Barnwell, J. w. and Galinski, M. R., 1998. Invasion of vertebrate cells: Erythrocytes. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 93-117.

Baranyi, J. and Roberts, T.A., 1994. A dynamic approach to predicting bacterial growth in food. *Int. J. Food Micro.* 23, 277-294.

Baranyi, J. and Tamplin, M.L., 2004. ComBase: A combined database on microbial responses to food environments. *J. Food Protect.* 67, 1967-1971.

Barker G.C. and Grimson M. J., 1993. A cellular automaton model of microbial growth. *Binary* 5, 132-137 .

Baton, L.A. and Ranford-Cartwright, L.C., 2005. Spreading the seeds of million-murdering death: metamorphoses of malaria in the mosquito. *Trends in Parasitology* 21, 573-580.

Bermúdez, J., López, D., Valls, J. and Wagensberg, J. 1989. On the analysis of microbiological processes by Montecarlo simulation techniques. *Cabios* 5 (4), 305-312.

Bernaerts, K., Versyck, K. J. and Van Impe, J. F., 2000. On the design of optimal dynamic experiments for parameter estimation of a Ratkowsky-type growth kinetics at suboptimal temperatures. *Int. J. Food Micro.* 54, 27-38.

Berryman, J. W., 1983. Random close packing of hard spheres and disks. *Phys. Rev. A* 27 (2) 1053-1061.

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. *PNAS* 99 (Suppl. 3), 7280-7287.

Bozdech, Z., Llinás, M., Pulliam, B. L., Wong, E. D., Zhu, J. and DeRisi, J. L., 2005. The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. *PLoS Biol.* 3(12) e426.

Brehm-Stecher, B. F. and Johnson E. A., 2004. Single-cell microbiology: Tools, technologies, and applications. *Microbiol. Molec. Biol. Rev.* 68, (3) 538-559.

Buchanan, R.L., 1991. Using spreadsheet software for predictive microbiology applications. *J. Food Safety* 11, 123-134.developing countries

Burden, J.L. and Faires, J.D., 2004. Numerical Analysis. Ed: Thomson Brooks.

Butcher G. A., 1981. A comparison of static thin layer and suspension cultures for the maintenance in vitro of Plasmodium falciparum. *Ann. Trop. Med. Parasitol.* 75 (1), 7-17.

Butcher G. A., 1982. The behavior of different strain of *Plasmodium falciparum* in suspension and static cultures. *Trans. Roy. Soc. Trop. Med. Hyg.* 76 (3), 407-409.

Capps, T.C. and Jensen, J. B., 1983. Storage requirements for erythrocytes used to culture *Plasmodium falciparum. J. Prasitol.* 69 (1),158-162.

Chin, W. and Collins, W.E., 1980. Comparative studies of three strains of *P. falciparum* isolated by the culture method of Trager and Jensen. *Am. Trop. Med. Hyg.* 29 (6), 1143-1146.

Chitnis, C.E. and Blackman M. J., 2000. Host Cell Invasion by Malaria Parasites. *Parasitol. today* 16 (10), 411-415.

Chotivanich, K. T., Dondorp, A. M., White, N. J., Peters, K., Vreeken, J., Kager, P. A. and Udomsangpetch, R., 2000. The resistance to physiological shear stresses of the erythrocytic rosettes formed by cells infected with *Plasmodium falciparum. Ann. Trop. Med. Parasitol.* 94 (3), 219-226.

Chyiaka C., Garira, W. and Dube, S., 2008. modeling inmune response and drug therapy in hman malaria infection. *Comp. Math. Meth. Med.* 9 (2), 143-163.

Clough, B., Atiola, F. and Pasvol, G., 1998.The role of rosetting in the multiplication of *Plasmodium falciparum*: rosette formation neither enhances nor targets parasite invasion into uninfected red cells. *Brit. J. Haematol.* 100, 9-104.

Corder, G. W. and Foreman, D. I., 2009. *Non parametric statistics for non-staticians. A step by step approach.* Wiley, New Jersey.

Crank, J., 1990. *The mathematics of diffusion.* Oxford Science Pub. Clarendon Press, Oxford.

Dens, E.J. and Van Impe, J.F., 2001. On the need for another type of predictive model in structured foods. *Int. J. Food Microbiol.* 64, 247-260.

Dietz, K., 1988. Mathematical models for transmission and control of malaria In: *Malaria*, Churchill Livingstone, Edinburgh (1988), 1091-1133.

Diggs, C., Pavanand, K., Permpani, B., , Numsuwan, V., Haupt, R. and Chuanak, N., 1971. Penetration of human fetal erythrocytes by *Plasmodium falciparum in vitro*. *J. Parasitol* 57 (1), 187-188.

Doona, C. J., Feherry, F. E. and Ross, E.W., 2004. A quasi-chemical model for the growth and death of microorganisms in foods by non-thermal and high-pressure processing. *Int. J. Food Microbiol.* 100, 21-32.

Escalante, A.A. and Ayala, F., 1995. Evolutionary origin of Plasmodium and other Apicomplexa based on rRNA genes. *Prod. Natl. Acad. Sci. USA* 92, 5793-5797.

Fairlamb, A. H., Warhurst,D.C. and Peters, W., 1985. An improved technique for the cultivation of *Plasmodium falciparum in vitro* without daily medium change. *Ann. Trop. med. Parasitol.* 79 (4), 379-384.

Ferrer, J., Vidal, J., Prats, C., Valls, J., Herreros, E., López, D., Giró, A. and Gargallo, D., 2007. Individual-based model and simulation of *Plasmodium falciparum* infected erythrocyte *in vitro* cultures. *J. Theor. Biol.,* 248 (3), 448-459.

Ferrer, J., Prats, C. and López, D., 2008. Individual-based modeling: an essential tool for microbiology. *J. Biol. Phys.* 34 (1-2), 19-37.

Ferrer, J., Rosal, M. D., Vidal, J. M., Prats, C., Valls, J., Herreros, E., López, D. and Gargallo, D., 2008. Effect of the hematocrit layer geometry on *Plasmodium falciparum* static thin-layer *in vitro* cultures. *Malaria Journal* 7: 203.

Ferrer, J., Prats, C., López, D. and Vives-Rego, J., 2009. Mathematical modeling methodologies in predictive food microbiology: a SWOT analysis. *Int. J. Food Micro.* 134 (1-2), 2-8.

Foty, R.A. and Steinberg, M.S., 2005. The differential adhesion hypothesis: a direct evaluation. *Develop. Biol.* 278, 255-263.

Gallup, J.L. and Sachs, J.D., 2001. The economic burden of malaria. *Am. J. Trop. Med. Hyg.* 64 (1-2), 85-96.

Geeraerd, A.H., Valdramidis, V.P., Devlieghere, F., Bernaert, H., Debevere, J. and Van Impe, J.F., 2004. Development of a novel approach for secondary modeling in predictive microbiology: incorporation of microbiological knowledge in black box polynomial modeling. *Int. J. Food Micro.* 91, 229-244

Ginovart, M., 1996. *Desenvolupament d'una metodologia fonamentada en la simulació discreta per a l'estudi de sistemes biològics d'interès tecnològic i científic.* PhD Thesis. Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Spain.

Ginovart, M., López, D. and Valls, J., 2002a. INDISIM, an individual-based discrete simulation model to study bacterial cultures. *J. Theor. Biol.* 214, 305-319.

Ginovart, M., López, D., Valls, J. and Silbert M., 2002b. Simulation modeling of bacterial growth in yoghurt. *Int. J. Food Microbiol.* 73, 415-425.

Ginovart, M., López, D., Valls, J. and Silbert M., 2002c. Individual-based simulations of bacterial growth on agar plates. *Physica A* 305, 604-618.

Ginovart, M., López, D. and Gras, A., 2005. Individual-based modeling of microbial activity to study mineralization of C and N and nitrification process in soil. *Nonlinear Anal. Real World Appl.* 6, 773-795.

Ginovart, M., López, D., Giró, A. and Silbert, M., 2006. Flocculation in brewing yeasts: a computer simulation study. *BioSystems* 83, 51-55.

Ginovart, M., Xifré, J., López, D. and Silbert, M., 2007. INDISIM-YEAST, an individual-based model to study yeast population in batch cultures. In: Méndez-Vilas, A. (Ed.), *Communicating Current Research and Educational Topics and Trends in Applied Microbiology*, Microbiology Book Series 1, vol. 1, p. 401-409, Formatex, Badajoz, Spain (ISBN: 978-84-611-9422-3).

Ginsburg, H. and Hoshen M.D., 2002. Is the development of falciparum malaria in the human host limitted by the availability of uninfected erythrocytes? *Malaria J.* 1, 18.

Giró, A., Gonzalez, J.M., Padró J.A. and Torra V., 1980. The structure of liquid lead at 670K through molecular dynamics. *J. Chem. Phys.* 73 (6), 2970-2972.

Giró, A., Padró, J.A., Valls, J. and Wagensberg, J., 1985. Monte Carlo simulation of an ecosystem: a matching between two levels of observation. *Bull. Math. Biol.* 47, 111-122.

Giró, A., Valls, J., Padró, J.A. and Wagensberg, J., 1986. Monte Carlo simulation program for ecosystems. *CABIOS* 2, 291-296.

Glushakova, S., Yin, D., Li, T. and Zimmerberg, J., 2005. Membrane transformation during malaria parasite release from human red blood cells. *Curr. Biol.* 15, 1645-1650.

Godin, M., Bryan, A. K., Burg, T.P., Babcock, K. and Manalis, S.R., 2007. Measuring the mass, density, and size of particles and cells using a suspended microchannel resonator. *J. Appl. Phys.* 91, 123121.

Garnham, P., 1988. Malaria parasites of man: life-cycles and morphology (excluding ultrastructure). In: *Malaria Principles and Practice of Malariology* (Vol. 1). Ed.: Wernsdorfer W. H. M. I.,Churchill, Livingstone, Edinburgh, 61-96.

Gras, A., 2004. *INDISIM-SOM. Un model discret per a l'estudi de la dinàmica de la matèria orgànica i la nitrificació en sòls.* PhD Thesis, Escola Tècnica Superior d'Enginyeria Agrària, Universitat de Lleida, Spain.

Gras, A., Prats, C. and Ginovart, M., 2006. An Individual-based model to study the main groups of microbes active in composting process. In: *Modern Multidisciplinary Applied Microbiology. Exploiting Microbes and their Interactions*, 14-18, Ed.: Méndez-Vilas, A. , Wiley-VCH Verlag, Weinheim, Germany (ISBN: 3-527-31611-6).

Gravenor, M. Mb. and Kwiatkowski D., 1998. An analysis of the temperature effects of fever on the intra-host population dynamics of *Plasmodium falciparum. Parasitol.* 117, 97-105.

Greenwood, B. and Mutabingwa, T., 2002. Malaria in 2002. *Nature* 415 (6872), 670-673.

Grimm, V., Frank, K., Jeltsch, F., Brandl, R., Uchmánski, J. and Wissel, C., 1996. Pattern-oriented modeling in population ecology. *Sci. Total. Environ. 183, 151-166*

Grimm, V., 1999. Ten years of Individual-based modeling in ecology: what have we learned and what could we learn in the future? *Ecol. Model.* 115, 129-148.

Grimm, V. and Railsback, S.F., 2005. *Individual-based modeling and Ecology.* Princeton University Press, Princeton and Oxford, New Jersey, USA.

Grimm V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H. H., Weiner, J., Wiegand,T. and DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310 (5750), 987-991.

Grimm V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jorgensen, C., Mooij, W.M., Muller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Ruger, N., Strand, E., Souissi, S., Stillman, R.A., Vabo, R., Visser, U., DeAngelis, D.L., 2006. A standard protocol for describing individual-based and agent-based models. *J. Ecol. Model.* 198 (1), 115-126.

Gujer, W., Henze, M., Mino, T., van Loosdrecht, M.C.M., 1999. Activated sludge model no. 3. *Water Sci. Technol.* 39, 183-193.

Guyton, A. C., 1996. *Manual de Fisiología Humana. 6ª edición.* Nueva editorial Interamericana, México DF.

Hawking, F., 1975. Circadian and other rythms of parasites. *Adv. Parasitol.* 13, 123-182.

Haydon, D.T., Matthews, L., Timms, R. and Colegrave, N., 2003. Top-down or bottom-up regulation of intra-host blood-stage malaria: do malaria parasites most resemble the dynamics of prey or predator? *Proc. R. Soc . Lond.* B 270 (1512), 289-298.

Heddini, A. 2002. Malaria pathogenesis: a jigsaw with increasing number of pieces. *Int. J. Parasitol.* 32 (13), 1587-1598.

Hellweger, F. L. and Bucci, V., 2009. A bunch of tiny individuals? Individual-based modeling for microbes. *Ecol. Model.* 220, 8-22.

Hochmuth, R. M. and Marcus, M. D., 2002. Membrane tethers formed from blood cells with available area and determination of their adhesion energy. *Biophys. J.* 82 (6), 2964-2969.

Hoffmann, K. A. and Chiang S. T., 2004. Parabolic partial diferential equations. *In: Computational fluid dynamics. Volume 1.* Ed. EEESbooks, Wichita, USA, 60-96.

Hoppensteadt, F.C. and Peskin, C.S., 2001. *Modeling and simulation in medicine and life science.* Ed: Springer, Maryland, 295-306.

Jamshidi, N. and Palsson, B. Ø., 2006. Systems biology of the red blood cell. *Blood cell. Mol. Dis.,* 36, 239-247.

Jaynes, E. T., 1957. Information theory and statistical Mechanics. *Phys. Rev.,* 106 (4), 120-130.

Jensen J. B. and Trager W., 1978. Some recent advances in the cultivation of Plasmodium falciparum. *Isr. J. Med. Sci.* 14, 563-570.

Jensen M. D., Conley M. and Helstowski L. D., 1983. Culture of Plasmodium falciparum: the role of pH, glucose and lactate. *J. Parasitol.* 69, 1060-1067.

Jensen, J. B., 1988. *In vitro* cultivation of malaria parasites: erythrocytic stages. *In: Malaria. Principles and practice in malariology . Vol. 1.* Churchill Livingstone, Edinburgh, 307-320.

Jimenez-Diaz, M.B., Mulet, T., Gomez, V., Viera, S., Alvarez, A., Garuti, H., Vazquez, Y., Fernandez, A., Ibanez, J., Jimenez, M., Gargallo-Viola, D. and Angulo-Barturen, I., 2009. Quantitative Measurement of *Plasmodium*-Infected Erythrocytes in Murine Models of Malaria by Flow Cytometry Using Bidimensional Assessment of SYTO-16 Fluorescence. *Cytometry-A* 75 A, 225-235.

Jou, D., 1985. *Introducció a la termodinàmica de processos biològics.* Ed.: Institut d'Estudis Catalans, Barcelona.

Karsenti E., 2009. Self-organisation in cell biology: a brief history. *Nature Rev. Mol. Cel. Biol.* 9 (4), 255-262

Kauffman, K., Pajerowski, J., Jamshidi, N. , Palsson, B. and Edwards, J., 2002. Description and Analysis of Metabolic Connectivity and Dynamics in the Human Red Blood Cell. *Biophys. J.* 83 (2) 646-662.

Kiatkowski, D. 1989. Febrile temperatures can synchronize the growth of Plasmodium falciparum in vitro. *J. Exp. Med.* 169, 357-361

Kirk, K. 2001. Membrane transport in the malaria-infected erythrocyte. *Phisiological Rev.* 81 (2) 495-537.

Kittle, C, 1958. Fundamental principles of statistical mechanics. In: *Elementary Statistical Physics.* Ed.: Courier Dover Publications, 16-35.

Koch, A.L., 1966. Distribution of cell size in growing cultures of bacteria and applicability of Collins-Richmond principle. *J. Gen. Microbiol.* 45(3), 409.

Koch, A.L. and Higgins, M.L., 1984. Control of wall band splitting in *Strptococus-faecalis. J. Gen. Microbiol.* 130, 735-745.

Kreft J. U., Booth G. and Wimpeney J. W. T., 1998. BacSim, a simulator for individual-based modeling of bacterial colony growth. *Microbiology* 144, 3275-3287.

Kuchel, P. K., 2004. Current status and challenges in connecting models of erythrocyte metabolism to experimental reality. *Prog. Biophys. Mol. Biol.* 85, 325-342.

Kuske, R., Gordillo, L.F. and Greenwood, P., 2006. Sustained oscillations via coherence resonance in SIR. *J. Theor. Biol.* 245 (3), 459-469.

Kutalik, Z., Razaz, M., Baranyi, J., 2005. Conection between stochastic and deterministic modeling of microbial growth. *J. Theor. Biol.* 232, 285-299.

Lambros, C. and Vanderberg, J . P., 1979. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol.* 65 (3), 418-420.

Leclerc, M.C, Hugot J.P., Durand, P. and Renaud, F., 2004. Evolutionary relationships between 15 *Plasmodium* species from new and old world primates (including humans): an 18S rDNA cladistic analysis. *Parasitol.* 129, 677-684.

Leete, T. H. and Rubin, H., (1996). Malaria and the cell cycle. *Parasitol. Today* 12, 442-444.

Leitch, J. M. and Carruthers, A., 2007. ATP-dependent sugar transport complexity in human erythrocytes. *Am. J. Physiol. Cell. Physiol.* 292 (2), C974-C986.

Lew , V. l., Tiffert, T. and Ginsburg, H., 2003. Excess hemoglobin digestion and osmotic stability of *Plasmodium falciparum* infected red blood cells. *Blood* 101 (10), 4189-4194.

Li, X.Z., Gupur, G. and Zhu G.T., 2001. Threshold and stability results for an age-structured SEIR epidemic model. *Comp. Math. Appl.* 42, 883-907.

Li, H., Xie, G., Edmondson, A., 2007. Evolution and limitations of primary mathematical models in predictive microbiology. *Brit. Food J.* 109 (8), 608-626.

López, D., 1992. *Simulació de cultius bacterians. Estudi cinètic i termodinàmic.* PhD thesis. Departament de Física Fonamental, Divisió de Ciències Experimentals i Matemàtiques, Universitat de Barcelona, Spain.

López, S., Prieto, M., Dijkstra, J., Dhanoa, M.S. and France, J., 2004. Statistical evaluation of mathematical models for microbial growth. *Int. J. Food Micro.* 96, 289-300.

Lotka, A. J. 1925. *Elements of Physical Biology.* Ed.: Williams and Wilkins, London.

Lurié, D., Valls, J. and Wagensberg, J., 1983. Thermodynamic approach to biomass distribution in ecological systems. *Bull. Math. Biol.* 45 (5), 869-872.

Margalef R, 1963. Certain unifying principles in ecology. *Am. nat.* 97 (897), 357-368.

Margalef, R., 1991. *Teoría de los sistemas ecológicos.* Ed.: Univ. Barcelona, Spain. pp 290.

Matthäus, F., Jagodîc, M and Dobnikar, J., 2009. *E. coli* Superdiffusion and Chemotaxis—Search Strategy, Precision, and Motility. *Biophys. J.* 97 (4) 946–957.

Mauritz, J.M.A., Esposito, A., Ginsburg, H., Kaminski, C.F., Tiffert, T. and Lew, V.L., 2009. The Homeostasis of *Plasmodium falciparum*-Infected Red Blood Cells. *PLOS Comp. Biol.* 5 (4) , e1000339.

McKellar, R.C., Lu, X., Delaquis, P. J., 2002 A probability model describing the interface between survival and death of Escherichia coli O157:H7 in a mayonnaise model system. *Food Microbiol.* 19, 235-247.

McKenzie, F.E. and Bossert, W. H., 2005. An integrated model of Plasmodium falciparum dynamics. *J. Theor. Biol.* 232 (3), 411-426.

McMeekin, T. A.,Baranyi, J., Bowman, J., Dalgaard, P., Kirk, M., Ross, T., Schmid, S. and Zwietering, M.H., 2006. Information systems in food safety management. *Int. J. Food Micro.* 112, 181-194.

Metropolis, N. and Ulam, S., 1949. The Monte Carlo Method. *J. A. Stat. Ass.* 44, 335-341.

Mideo, N., Day, T. and Read, A.F., 2008. Modeling malaria pathogenesis. *Cell. microbiol.* 10 (10), 1947-1955.

Miller, L.H., Baruch, D.I., Marsh, K. and Doumbo O.K., 2002. The pathogenic basis of malaria. *Nature* 415 (6872), 673-679.

MMV- Medicines for Malaria Venture, 2009, Website. <www.mmv.org>

Morton, K. W. and Mayers, D. F., 1996. Numerical Solution of Partial Differential Equations. Cambridge University Press, Cambridge.

MOSIMBIO, 2009, Website. <http://mosimbio.upc.edu>

MR4- Malaria Research and Reference Reagent Resource Center, 2008, Website. <http://www.mr4.org/>

Naughton, J. A. and Bell, A., 2007. Studies on cell-cycle synchronization in the asexual erythrocytic stages of *Plasmodium falciparum*. *Parasitology*. 134 (3), 331-337.

Nauta, M.J., 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. *Int. J. Food Micro.* 57, 9-18.

Neu, B., and Meiselman, H. G., 2001. Sedimentation and Electrophoretic Mobility Behavior of Human Red Blood Cells in Various Dextran Solutions. *Langumir* 17, 7973-7975.

Nicholson, A. J. , 1933. The balance of animal populations. *J. Animal Ecol.* 2, 132-178.

Odum, E., P., 1969. Strategy of ecosystem development. *Science* 164 (3877) 262.

Palsson, B. Ø., Narang, A. and Joshi A., 1989. Computer model of hman red blood cell metabolism. *Prog. Clin. Microbiol. Res.* 319, 133-154.

Pasvol, G. Weatherall, D.J. and Wilson R.J.M., 1980. The increased susceptibility of red blood cells to invasion by the parasite *Plasmodium falciparum*. *Brit. J. Haematol.* 45, 285-295.

Pavanand K., Permpani B., Chuanak N. and Sookto P., 1974. Preservation of Plasmodium-falciparum-infected erythrocytes for in vitro cultures. *J. Parasitol.* 60 (3), 537-539.

Picioreanu, C., Xavier, J.V., van Loosdrecht, M.C.M., 2004. Advances in mathematical modeling of biofilm structure. *Biofilms* 1, 337-349.

Pinder, J.C., Fowler, R.E. Bannister, L.H. Dluzewski, A.R. and Mitchell, G.H., 2000. Motile Systems in Malaria Merozoites: How is the Red Blood Cell Invaded? *Parasitol. Today* 16 (6), 240-245.

Prats, C., López, D., Giró, A., Ferrer, J. and Valls, J., 2006. Individual-based modeling of bacterial cultures to study the microscopic causes of the lag phase. *J. Theor. Biol.* 241, 939-953.

Prats, C., Ferrer, J., Giró, A., López, D. and Valls, J., 2006. Spatial properties in Individual-based modeling of microbial systems. Study of the composting process. In: *Modern Multidisciplinary Applied Microbiology. Exploiting Microbes and their Interactions*, Wiley-VCH Verlag, Weinheim, 461-465.

Prats, C., Ferrer, J., Flix, B., Giró, A., López, D. and Vives-Rego, J., 2007. Evolution of biomass distribution during bacterial lag phase through flow cytometry, particle analysis and Individual-based modeling. In: G. Nychas et al. (Ed.), *5th International Conference on Predictive modeling in Foods. Proceedings*, Agricultural University of Athens, Greece (ISBN: 978-960-89313-7-4), 301-304.

Prats, C., 2008. *Individual-based modelling of bacterial cultures in the study of the lag phase*. PhD thesis. Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Spain.

Prats, C., Ferrer, J., Giró, A., López, D. and Vives-Rego, J., 2008. Analysis and Individual-based modeling simulation of the stages in bacterial lag phase: basis for an updated definition. *J. Theor. Biol.* 252, 56-68.

Prats, C., Ferrer, J., López, D., Vives-Rego, J. and Giró, A., 2010. On the evolution of cell size distribution during bacterial growth cycle: experimental observations and Individual-based Model simulations. *AJMR.* 4 (5) 400-407.

Prats, C., Ferrer, J., Gras, A. and Ginovart, M., 2010. Individual-based modeling and simulation of microbial processes: yeast fermentation and multi-species composting. *MCMDS, in press.*

Prigogine, I. 1955. *Introduction to Thermodynamics of Irreversible Processes,* Ed.: John Wiley, New York.

Radfar, A. Méndez, D., Moneriz, C. Linares, M., Mariín-García, P, Puyet, A. Diez, A. and Bautista,J. M., 2009. Synchronous culture of *Plasmodium falciparum* at high parasitaemia levels. *Nature Protocols* 4 (12) 1899-1914.

Rapoport, T.A., Henrich R. and Rapoport, S.M., 1976. The regulatory principles of glycolysis in erythrocytes in vivo and in vitro. A minimal comprehensive model describing steady states, quasi-steady states and time-dependent processes. *Biochem. J.* 154 (2), 449-469.

Rayner, J. C., 2006. Erythrocyte exit: Out, damned merozoite! Out I say! *TRENDS Parasitol.* 22 (5), 189-191.

Malaria Foundation International. <http://www.rollbackmalaria.org/index.html>

Rhodes, C. J. and Anderson R. M., 1996. Persistence and Dynamics in Lattice Models of Epidemic Spread. *J. Theor. Biol.* 180 (2) 125-133.

Rich, M. N. and Ayala, F., 2000. Population structure and recent evolution of *Plasmodium falciparum. PNAS.* 97 (13), 6994-7001.

Rojas, M. O. and Wasserman, M., 2007. Effect of Low Temperature On the In Vitro Growth of *Plasmodium falciparum. J. Euk. Microbiol.* 40 (2), 149-152.

Rosenthal, P. J. and Meshnik, R. S., 1998. Hemoglobin processing and metabolism of amino acids, heme and iron. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 145-159.

Roth, E., 1990. *Plasmodium falciparum* carbohydrate metabolism: a connection between host cell and parasite. *Blood cells* 16, 453-460.

Rouzine, I. M. and McKenzie, F.E., 2003. Link between immune response and parasite synchronization in malaria. *Proc. Natl. Acad. Sci. U.S.A.* 100 (6), 3473-3478.

Rowe, J.A., Obiero, J., Marsh, K. and Raza, A., 2002. Positive correlation between rosetting and parasitaemia in *Plasmodium falciparum* clinical isolates. *Am. J. Trop. Med. Hyg.,* 66(5), 458-460.

Sachs, J. and Malaney, P., 2002. The economic and social burden of malaria. *Nature* 415 (6872), 680-685.

Sarkey, K.J., 2008. Deterministic epidemilogical models at the individual level. *J. Math. Biol.e* 57, 311-331.

Sherman, I. W., 1998 Carbohydrate metabolism of Asexual stages, 1998. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 135-145.

Sherman, I. W., 1998 Purine and pyrimidine metabolism of Asexual stages, 1998. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 177-185.

Sherman, I.W., Eda, S. and Winogard, E., 2004. Erythrocyte aging and malaria. *Cell. Mol. Biol.* 50 (2) 159-169.

Shung, K. K., Krisko, B.A. and Ballard, J.O., 1982. Acoustic measurement of erythrocyte compressibility. *J. Acoust. Soc. Am.* 72 (5), 1364-1367.

Silvert, W., 2001. Modeling as a discipline. *Int J. Gen. Sys.* 30 (3), 261-282.

Simpson, J.A., Silamut. K., Chotivanich, K., Pukrittayakamee, S. and White, N.J., 1999. Red cell selectivity in malaria: a study of multiple-infected erythrocytes. *Trans. Ry. Soc Trop. Med. Hyg.* 93 (2), 165-168.

Solé, R.V. and Valls, J., 1992. Nonlinear phenomena and chaos in a montecarlo simulated microbial ecosystem. *Bull. Math. Biol.* 54, 939-955.

Solé, R.V., López, D., Ginovart, M. and Valls, J., 1992. Self-organized critically in Monte Carlo simulated ecosystems. *Phys. Lett.* A 172, 56-61.

Sole-Mauri, F., Illa, J., Magrí, A., Prenafeta-Boldú F.X. and Flotats, X., 2007. An integrated biochemical and physical model for composting process. *Bioresource Technology*, 98, 3278-3293.

Standaert, A., 2007. *From Macroscopic to microscopic modeling methodologies in predictive microbiology.* PhD Thesis. Department Chemische Ingenieurstechniekene. Katholieke Universiteit Leuven.

Stewart, P. S., 1996. Theoretical aspects of antibiotic diffusion into microbial biofilms. *Antimicrob. Agents. Chemother.* 40 (11), 2517-2522.

Stewart, D., 1998. A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. *Biotechnol. bioeng.* 59 (3), 261-272.

Stewart, P. S., 2003. Diffusion in biofilms. *J. Bacteriol.* 185 (5), 1485-1491.

TeGiffel, M. C. and Zwietering, M. H., 1999. Validation of predictive models describing the growth of Listeria monocytogenes. *Int. J. Food Microbiol.* 46, 135-149.

Torii, M. and Aikawa, M., 1998. Ultrastructure of Asexual stages, 1998. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 123-134.

Toussaint, O. and Schneider, E. D., 1998. The thermodynamics and evolution of complexity in biological systems. *Comp. Biochem. Phys. A.* 120, 3-9.

Trager W. and Jensen J.B., 1976. Human malaria parasites in continuous culture. *Science* 193, 673-675.

Trager, W., 1994. Cultivation of Malaria parasites. *Meth. cell. biol.* 45, 7-26.

Uchmanski, J. and Grimm, V., 1996. Individual-based modeling in ecology: what makes the difference? *TREE* 11 (19), 437-441.

United Nations, 2000. Millenium project. Website. <http://www.unmillenniumproject.org/>

Valls, J., 1986. *Simulació de sistemes de n-cossos interactius. Aplicació als ecosistemes.* PhD thesis. Departament de Termologia, Facultat de Física, Universitat de Barcelona, Spain.

Vial H. J., Ancelin M. L., 1998. Malaria lipids. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 177-185.

Vives-Rego, J., Resina, O., Comas, J., Loren, G. and Julià, O., 2003. Statistical analysis and biological interpretation of the flow cytometric heterogeneity observed in bacterial axenic cultures. *J. Microbiol. Meth.* 53, 43-50.

Wagensberg, J., López, D. and Valls, J., 1988a. Statistical aspects of biological organization. *J. Phys. Chem. Solids* 49, 695-700.

Wagensberg, J., Valls, J. and Bermúdez, J., 1988b. Biological adaptation and the mathematical theory of adaptation. *Bull. Math. Biol.* 50, 445-464.

Wahlgren, M., Fernandez, V., Scholander, C. and Carlson, J., 1994 Rosetting. *Parasitol. today* 10 (2), 73-79.

White, N.J., Chapman, D. and Watt, G., 1992. The effects of multiplication and synchronicity on the vascular distribution of parasites in *falciparum* malaria. *Trans. Roy. Soc. Trop. Med. Hyg.* 86, 590-597.

White, N.J., 1998. Malaria pathophisiology. In: *Malaria: parasite biology, pathogenesis and protection.* ASM Press, Washington DC, 371-385.

World Health Organisation, 2000. Severe falciparum malaria. *Trans. Roy. Soc Trop. Med. Hyg.* 94 (Suppl. 1), S1/1-S1/77.

World Health Organisation, 2008. Websites. <http://www.who.int/topics/malaria/en/>, <http://apps.who.int/malaria/wmr2008/malaria2008.pdf>.

Wickham, M.E. et al. (2003) Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J. Biol. Chem.* 278, 7658-3766.

Winograd, E., Clavijo, C. A., Bustamante, L. Y. and Jaramillo, M., 1999. Release of merozoites from *Plasmodium falciparum*-infected erythrocytes could be mediated by a non-explosive event. *Parasitol. Res.* 85, 621-624.

Xavier, J. B., Picioreanu, C. and van Loosdrecht M. C. M., 2005. A framework for multidimensional modelling of activity and structure of multispecies biofilms. *Env. Microbiool.*7 (8), 1085-1103.

Yu, L., 2007. Overview of some theoretical approaches for derivation of the Monod equation. *Appl. Microbiol. Biotechnol.* 73, 1241-1250.

Zar, J.H., 1999. *Biostatistical analysis. 4th edition.* Prentice Hall, Upper Saddle River, NJ.

Zolg, J. W., MacLeod, A. J., Dickson, I. H. and Scaife, J.G., 1982. *Plasmodium falciparum*: modifications of the *in vitro* culture conditions improving parasitic yields. *J. Parasitol.* 68 (6), 1072-1080.

Acknowledgements

Personal thanks

This Thesis has been elaborated in collaboration with many other researchers. First, I would like to thank my directors and tutors: Dr. Daniel López, and Dr. Joaquim Valls for their trust, confidence, enthusiasm and insight. Next, I give thanks to the team from the group EMG-GSK: Dr. Jaume Vidal, Dr. Domingo Gargallo Viola, Dr. Esperanza Herreros, Dr. Iñigo Angulo Barturen, and the technical staff for their expertise and skillfulness in malaria research, which provided with the motivation and basis for the work here presented.

Special thanks to my research mate Dr. Clara Prats for her sagacity, her assistance and guide, but for her kindness and friendship above all. Thanks to all the members of the group MOSIMBIO: Dr. Marta Ginovart, Dr. Anna Gras, Dr. Antoni Giró, Dr. Josep Vives-Rego, Dr. Moises Silbert, Roger Dalmau, Marta Dies, Jordi Valls, Xavier Portell and Dr. Rosa Carbó.

The community of physicists in the ESAB (Escola Superior d'Agricultura de Barcelona) and the CBL (Campus del Baix Llobregat) has also contributed with help and support. Particularly, I would like to give thanks to Eva, who is closer than a sister to me, and to Dr. Eloi Pineda, Dr. Trinitat Pradell, Dr. Pere Bruna, Dr. Francesca Ribas, Dr. David Pino, Dr. Pablo Loren, Dr. Jorge Serrano and Dr. Daniel Crespo (and also to my my recent work fellows, Wei and Radostin), for making me enjoy going to work and meeting them.

Thanks to all the people from the Department of Ecological Modeling in the UFZ (Center for Environmental Research) for their instruction during my stay in Leipzig. In particular: Dr. Volker Grimm, Dr. Roger Jovani, Thomas Banitz, Dr. Karin Johst, and to all the colleagues and friends that made the experience more pleasant and the winter less cold: Leandro, Álvaro, Jasmine, Irene, Jörg, Elias, Jorgina, Rafa, among others.

Thanks to the staff and colleagues from the PhD course in the DTU (Technical Uni-

versity of Denmark): Dr. Joao Xavier, Dr. Jan Ulrich Kreft, Ines Schrottenbaum and Dr. Christophe Ahhstof, among others.

My most sincere gratitude for Marcel Pié, who collaborated in the graphical design of this book and to the rest of friends from the band Cheb Balowski, which brought music to my live and provided the maintenance during my first steps as a researcher.

Last but not least, I would like to give deep personal thanks to all the people that have been my true lifeline. To my parents Isidre and Nona, to my brother Claudi, to my nearest cousins and closest family, to Sergi, to Esther, to my flat mates Blanca, Quique and Xavi, to the friends of the UB, the ones of the high school J. Boscà and of the elementary school Orlandai. Thanks to my teacher Jordi Sanchez, who showed me the beauty of science and to Hector Martínez, who helped me finishing what I had started.

Financial support

I acknowledge the financial support of GlaxoSmithKline, for the initiatory scholarship, and of the Ministerio de Educación y Ciencia and the Ministerio de Ciencia y Tecnología for their grants to the research projects in which I have participated: Plan Nacional I+D+i CGL 2004-01144 and CGL 2007-65142. I want also to give thanks to the Universitat Politècnica de Catalunya, for hiring me as an assistant teacher during these last years.