Flux balance analysis of *Plasmodium falciparum* growth and energy metabolism

by

Robert William Burger

Thesis presented in partial fulfilment of the requirements for the degree of Master of Biochemistry in the Faculty of Science at Stellenbosch University

Department of Biochemistry
University of Stellenbosch
Private Bag X1, 7602 Matieland, South Africa

Supervisors:

Prof. J. L. Snoep Dr. J. Eicher

December 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2017

Abstract

Flux balance analysis of *Plasmodium falciparum* growth and energy metabolism

R.W. Burger

Department of Biochemistry
University of Stellenbosch
Private Bag X1, 7602 Matieland, South Africa

Thesis: MSc Bsc (Biochem.)

December 2017

Genome-scale network reconstructions serve as an ideal tool for the modeling of large, complex metabolic networks. In this study we use flux-balance analysis to investigate key amino acid dependencies of *Plasmodium falciparum* growth. Our model was able to produce a realistic flux distribution geared for the optimal production of biomass. We were able to simulate parasite growth in a variety of conditions and assess the favourability of these conditions for optimal growth. To demonstrate the applicability of the model we additionally simulate the response of the model to the introduction of an antimalarial agent, atovaquone.

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations:

- Supervisors Prof. Jacky Snoep and Dr. Johann Eicher for their guidance and insight throughout the project,
- Laboratory manager, Arrie Arends, for ensuring the smooth running of the laboratory and always making himself available for assistance,
- My fellow researchers at the JJJ laboratory, for always being open to lend their council and for maintaining a positive, friendly, working environment,
- The financial support of the NRF, of which none of my research would have been possible without it.

Contents

Chapter 1

Introduction

1.1 Background and motivation

As a disease that puts almost half the world's population at risk, malaria has been a worldwide public health challenge [78]. In 2015 there were an estimated 212 million cases of malaria, with 429 000 of those cases estimated to have been fatal [78]. Malaria is a mosquito-borne infectious disease caused by the protozoan parasite, *Plasmodium*. Carried by the female *Anopheles* mosquito, *Plasmodium* sporozoites are transferred to humans during a blood meal where they infect the host's liver and use the host nutrients to rapidly proliferate. At this point patients remain asymptomatic. During the next phase, parasites are released from liver cysts, entering the bloodstream in the form of merozoites, where they will invade red blood cells and begin to reproduce asexually. It is only once the infected erythrocytes reach the point of rupture that symptoms will begin to present themselves. The subsequent release of burst red blood cell contents, sequestration of infected erythrocytes to vascular endothelial cells and a multitude of metabolic changes that occur, can lead to anaemia, fever and respiratory distress [68, 15, 35]. In the case of *Plasmodium falciparum*, there is a risk of infection spreading to the brain, vastly diminishing the likelihood of survival and increasing the chance of permanent damage. *Plasmodium falciparum* infections account for 99% of malaria-associated deaths worldwide [78]. It is for this reason that *Plasmodium falciparum* will be the focus of this study. Due to emergent resistant strains, it is of great importance that novel or improved anti-malarial drugs and administration procedures be developed. As mentioned

by Liu *et al.* [60], understanding of the parasite's molecular mechanisms and interactions are essential for the identification of new drug targets.

Mathematical modeling allows us to investigate the characteristics of *Plasmodium falciparum* metabolism prior to conventional experimentation, enabling the identification of significant metabolic interactions or occurrences, thus promoting a more focused approach in experimental design and investigation. Furthering the biochemical knowledge of this organism helps in the discovery of potential drug targets and improving existing treatment methods. The genome of *Plasmodium falciparum* has been sequenced, providing a library of parasite proteins and regulatory elements [29, 10]. Adding to this, transcriptomic [9], proteomic [24, 54], and metabolomic [77] studies have been conducted to further investigate key cellular interactions and metabolic pathways. All these data are added to specific databases that are freely available. These databases can be used to computationally generate metabolic models that are able to simulate an organism's metabolism. Based on mathematical methods such as linear programming, flux balance analysis can be used to predict the flux distributions of a metabolic network for a specified set of conditions [22].

1.2 Aims

To date, our group has largely focused on modeling the glycolytic pathways of *P. falciparum*. Using genome-scale network reconstructions we can create a network of reactions to acquire insight into the amino acid metabolism of the parasite, and study how these pathways interact with those of the glycolytic pathway. A metabolic network of *Plasmodium falciparum* was created to simulate parasite growth inside a human red blood cell with a focus on amino acid metabolism resulting from haemoglobin digestion, and the subsequent amino acid transport dynamics as accurately as possible. From these simulations, the following questions can be addressed:

- Is the model able to reproduce experimental observations in terms of amino acid incorporation into biomass and amino acid export, using haemoglobin catabolism as a primary source, and extracellular amino acid scavenging to make up for any deficits?
- Are we able to reproduce the findings of Liu *et al.* [59] that *Plasmodium falciparum* is able to grow in culture medium containing isoleucine as the only supplemented amino acid?

- Are we able to identify a steady-state flux distribution for the model system that is realistic to what we can observe in nature?
- From the flux distribution outputs, can we identify any key reactions or metabolites that could be investigated further as potential drug targets?

In *P. falciparum* metabolism, a clear dependence on specific amino acids for the optimal growth of the parasite has been established [59, 16]. The aim of this study was to use established biological knowledge of *Plasmodium falciparum* to construct a computational model able to simulate parasite growth and energy metabolism, with the purpose of investigating the dependence of each individual amino acid for the optimal rate of biomass formation. A network reconstruction constructed by Plata *et al.* [89] was used as a foundation to develop such a model. Their model contained 1001 compartmentalised reactions and 1024 metabolites, which was an adequately comprehensive network containing all the key transport reactions, a biomass objective function, and the incorporation of haemoglobin import into the network. The 1001 reactions included comprised 75% of the total enzymatic reactions. The enzymatic reactions removed were those with no literature support, and those not applicable to biomass production. The Plata model was originally used to simulate the effect of gene knock-outs on biomass production, and using constraints, was able to predict direction of concentration changes for external metabolites with 70% accuracy. Using the Plata model as a foundation meant that some of the necessary model infrastructure would already be in place for the construction of our own model. Sub-cellular localization of metabolites, substrate and cofactor specificities, pathway gaps, and reaction directionalities were for the most part accounted for. Using flux-balance analysis we were able to obtain a flux distribution profile for amino acid and growth metabolism for a variety of simulated conditions. Samples obtained from *in vitro* cultures were subjected to amino acid analysis using UPLC separation with UV detection (after 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate derivitization). The metabolic fluxes were determined from the experimental data, which compared well with model-simulated fluxes. Furthermore, we were able to make key observations into the metabolic dependencies of *Plasmodium falciparum* growth, and test the response of the model to the effects of a popular antimalarial compound, atovaquone.

Chapter 2

Literature overview

2.1 Biology of *Plasmodium falciparum*

Malaria is a disease caused by several species of the *Plasmodium* genus. Classified as obligate intracellular protozoa, these parasites are dependent on a vertebrate host for the asexual stage, and the female *Anopheles* mosquito, for the sexual stage of their life-cycle. During a blood meal, mosquito saliva containing sporozoites are transferred into the bloodstream of the host and travel to the liver where they invade hepatocytes and proliferate to form liver-stage schizonts. Each of these schizonts can release thousands of merozoites [30, 31, 100, 36, 101]. It is believed that these schizonts actually bud off from hepatocytes as membrane-bound clusters, termed merosomes [112, 109]. These merosomes exit the liver and travel towards the lungs where they will release merozoites into the pulmonary micro-circulation to invade host erythrocytes. Merozoites bind to and enter erythrocytes via an apicomplexan invasion organelle [1]. During infection of the erythrocyte, the merozoite forms a vesicle from the erythrocyte membrane [14]. The merozoite proceeds to mature into an early trophozoite, known as the ring stage, named after the ring-like appearance observed microscopically. The ring stage lasts about 16 hours, accounting for one third of the erythrocytic life cycle, but has been regarded as metabolically nondescript [32]. During the subsequent trophozoite stage, notably the most metabolically active phase of development, the parasite will digest up to 80% of the host cell's haemoglobin, meanwhile growing in size [5, 70, 79]. At around the 38 hour mark of development the trophozoite will begin to undergo schizogeny, whereby the parasite will undergo multiple rounds of mitosis to produce 15 to 30

merozoites within the parasitophorous vacuole membrane, collectively referred to as a schizont [6]. After 48 hours of development, *Plasmodium falciparum* infected red blood cells will rupture, releasing merozoites into the bloodstream to re-invade new erythrocytes. It is only in this blood-stage of infection where clinical symptoms will present themselves.

2.2 Pathogenesis of malaria

Patients with malarial infections progressing to the blood-stage will initially display flu-like symptoms such as fever, nausea, muscle pain and convulsions, as well as symptoms of anemia.There is an array of factors responsible for this. Adhesin proteins encoded by the parasite facilitate the binding of infected cells to the blood vessel walls of the host, preventing their removal by macrophages in the spleen. Sequestration of infected cells in the lungs (and brain in the case of *P. falciparum* infections) leads to partial obstruction of blood flow, endothelial barrier breakdown, and inflammation [72]. Infected erythrocytes have a glucose uptake rate of up to 100 times higher than uninfected erythrocytes [95]. Infection of host red blood cells also causes an immune response by the host, whereby macrophages destroy infected cells, but also uninfected cells [43]. Reduced deformability of uninfected red blood cells and intrinsic and extrinsic changes to these cells increase recognition by macrophages [69]. Erythrocyte production by host bone marrow is unable to keep up [53] with this concomitant loss and an anemic state will result. Anemia and various other factors present in malaria infections will cause an insufficient oxygen supply to tissue that may cause a metabolic shift in the host to anaerobic glucose metabolism and increased lactic acid production [50]. Together, these shifts in metabolism will present symptoms of hypoglycemia, lactic acidosis, and brain infection in the case of *P. falciparum*.

Proteolysis of haemoglobin releases a reactive heme group. Humans have a heme oxygenase responsible for the cleaving of heme rings, which produces biliverdin, ferrous iron and carbon monoxide. However, *Plasmodium* lack the heme oxygenase enzyme that vertebrates use to degrade free heme. Free heme has been shown to threaten parasite development by having the potential to inhibit enzyme activity [34, 41], peroxidise membranes [111], and cause oxidative free radical accumulation in the digestive vacuole [3]. In malarial infections, host erythrocytes are lysed at a much higher rate than normal, leading to an increased amount of free heme in the host blood-stream. These free heme groups

are detoxified by the parasite via polymerization of heme groups into inert crystalline structures known as hemozoin, which accumulate in the digestive vacuole. Carbon monoxide has also been suggested to bind to ferrous iron in haemoglobin to prevent heme release [83].

2.3 *Plasmodium falciparum* metabolism

2.3.1 Protein metabolism

Over a 48 hour life cycle of *P. falciparum* within a red blood cell, it is accepted that up to 80% of haemoglobin is digested by the parasite; however, only about 16% of the released amino acids are actually utilised [96, 51]. The colloid-osmotic hypothesis proposes that parasites digest host cell cytosol to prevent the red blood cell reaching critical haemolytic volume (CHV) which would result in premature lysis of the infected erythrocyte [32, 122]. During this process, haemoglobin is rapidly degraded and the amino acids produced are largely exported out the cell, in what would otherwise be considered a metabolically demanding process. The increase in osmotic pressure as a result of freed amino acids due to haemoglobin digestion alone is not enough to determine the time of lysis of the erythrocyte. Another hypothesis is that haemoglobin degradation necessity may also be linked to protease inhibitors [27, 57, 94, 92, 93]. Inhibition of plasmepsin I and cysteine protease activity causes the haemoglobin degradation process to halt, resulting in osmotic swelling, impairment of vacuolar functions and eventually parasite death [26].

The majority of haemoglobin degradation occurs during the 6-12 hour trophozoite stage [119]. New permeation pathways (NPP's) have been shown to be induced in the membrane of the erythrocyte [33, 48]. Mauritz *et al.* [67] have hypothesised that upon entry of this metabolically intensive trophozoite stage, NPP development is induced by the parasite. Host cell haemoglobin is ingested by the acidic vacuole of the parasite and proteolysed to its amino acid constituents. A small portion of the amino acids are retained for parasite growth and the rest exported with the aid of these newly formed NPP's. *Plasmodia* have evolved to no longer be capable of amino acid biosynthesis, instead relying on haemoglobin catabolism and extracellular amino acid supply [29]. With regards to *Plasmodium* growth, haemoglobin is a deficient source of methionine, cysteine, glutamine, glutamate, and contains no isoleucine at all. This indicates a strong likelihood that these amino acids are scavenged from extracellular sources. Divo *et al.* [16] discovered that *P. falciparum* were able to

grow in media containing only five exogenous amino acids (cysteine, glutamate, glutamine, methionine and isoleucine). Liu *et al.* [59] were then further able to show that *P. falciparum* is able to grow in media supplemented with only a single amino acid, isoleucine. Isoleucine is absent from haemoglobin, and research has shown that *Plasmodia* growth is not possible without the exogenous supplementation of isoleucine [59]. This finding suggests that aside from isoleucine, *P. falciparum* should be able to grow solely on amino acids released via haemoglobin catabolism.

2.3.2 Carbon metabolism

Glycolysis

Glucose metabolism by *Plasmodium spp.* is achieved via the Embden-Meyerhof-Parnas (EMP) pathway of glycolysis. Glucose is rapidly metabolised by the parasite, up to 100 times that of uninfected erythrocytes [95]. However, around 60-70% of glucose consumed by *Plasmodium* is incompletely oxidised to lactate and excreted, contrasting with the 90% glucose-to-lactate conversion observed in uninfected erythrocytes, reflecting the increased flux of glucose-sourced carbon into biomass generation [42]. Only a small fraction of glucose consumed is completely oxidized to carbon dioxide, in agreement with the belief that the parasite tricarboxylic acid plays a minor role in energy metabolism. A hexose transporter encoded by *Plasmodium* is targeted to the host cell membrane to allow for the increased demand for glucose [102].

Tricarboxylic acid cycle

It is well established that the genome of *P. falciparum* encodes all the tricarboxylic acid (TCA) cycle enzymes and are expressed during the asexual life stage [29, 9]. Isotopic-labelling studies have also shown this TCA cycle to be canonical [46]. The TCA cycle is mostly fueled by acetyl-coA (and oxaloacetate, to a smaller degree) obtained from glycolysis [108]. However, Cobbold *et al.* [12] have shown that as the parasite matures, a transition towards glutamine-derived TCA-intermediates over that obtained via glycolytic acetyl-coenzyme A (acetyl-CoA) production, giving more evidence for an adaptive metabolism employed by *Plasmodium.* Glutamine is deaminated to glutamate; where after an additional amino group is cleaved from glutamate to produce α-ketoglutarate for use in the TCA cycle [76]. During this process, pyruvate receives the amino group producing alanine, which ties up with the large spike in alanine export from the cell

observed during late trophozoite stages of growth. This also coincides well with the large consumption of glutamine released from an accelerating haemoglobin catabolism.

Pentose phosphate pathway

The pentose phosphate pathway (PPP) is a critically conserved pathway in almost all cells capable of of utilising carbohydrates as a carbon source. *P. falciparum* is no different, and PPP-enzymes were found to be encoded in the parasite's genome [29], as well as expressed [9]. The PPP comprises two interconnected branches. The oxidative arm generates riboses required for nucleic acid synthesis, and NADPH, a cofactor for biosynthetic reactions and redox control. The non-oxidative arm comprises a set of reversible reactions interconverting 3, 4, 5, 6 and 7-carbon sugar phosphates, enabling the recycle of ribose-5-phosphate back into glycolytic intermediates when there is an NADPH shortage. In human red blood cells the PPP is one of the major metabolic pathways, consuming 3-11% of glucose metabolised under normal conditions, and is the only source of NADPH required to reduce glutathione in response to oxidative stress [121]. In trophozoite-stage infected red blood cells, the activity of the oxidative branch increased to 78-fold that of uninfected erythrocytes [4]. In free parasites 82% of this activity can be accounted for, indicating *P. falciparum* is responsible for the majority of the flux increase, and also for the up-regulation of of the erythrocyte-PPP by 24-fold. By up-regulating NADPH production, the parasite is able to hamper programmed cell death initiation by the host immune system, by maintaining a high ratio of reduced to oxidised glutathione.

2.3.3 Parasite metabolism as a target for antimalarials

The arms race between disease and treatment solutions is an ongoing war that researchers and clinicians have fought throughout the ages. Humans have always been one step behind in this war; a constant struggle to identify new solutions to the ever-evolving resistances that pathogens develop in response to the cures and treatments we develop. In the case of microbial infections it is often a question of: "What is the most efficient solution to kill the unwelcome parasites, while doing as little harm as possible to the host?" An important factor with drug target identification strategies based on metabolomic approaches is the differences in metabolisms of host and parasite. Blocking the availability of metabolite X in host and pathogen could be a viable approach if the host has

no vital requirement of metabolite X for survival, whereas even a 10% disruption in metabolite Y production or depletion might kill both host and parasite. Therefore core strategies often initially require the identification of differences in metabolic network structure as well as robustness to change thereof.

Unlike the majority of eukaryotic cells, *Plasmodia* species do not rely on mitochondrial electron transport systems for ATP production. However, inhibition of the Cyt bc_1 complex of electron transport is usually lethal to the parasite, due to interruption of pyrimidine biosynthesis and destabilisation of mitochondrial membrane potential [65, 66, 115]. An observation was made by Painter *et al.* [81] that *P. falciparum* likely maintain mitochondrial electron chain activity for the sole metabolic purpose of regenerating ubiquinone, the electron acceptor of dihydroorotate dehydrogenase (DHODH). DHODH is responsible for the oxidation of dihydrooorotate to orotate, and is an essential reaction in the synthesis of pyrimidines required for DNA synthesis. Atovaquone, an analog of ubiquinone, has been used in conjunction with proguinil in the drug malarone for the treatment of malaria since 2000. Atovaquone occupies the quinol oxidase site of mitochondrial Cyt b, inhibiting the electron flux through the Cyt bc1 complex thereby preventing pyrimidine biosynthesis and collapse of mitochondrial membranes. At concentrations as low as 5 nM, atovaquone is still able to inhibit 90% of DHODH activity [40].

2.4 Mathematical models and metabolism

Mathematical modeling has been used as a powerful predictive tool across a diverse spectrum of fields including the natural sciences, engineering, economics and social sciences. In biology, metabolic networks can be described mathematically, using sets of equations to represent network reactions and a set of variables that define the metabolic environment. The model, represented as a set of functions describing interactions between variables, is optimised for a desired set of conditions and a set of prediction outputs are acquired. The fact that a metabolic environment can be defined means that modeling has implications in physiological studies, pathway gap-filling and flux manipulation in biological systems [80]. For example, by constraining certain reactions in a bacterial species, we can simulate a growth medium or even gene knock-outs and then use a mathematical approach such as flux balance analysis (FBA), to predict the yield of key metabolites and cofactors.

As biological systems consist of a number of low-level components acting to-

gether to display high-level system behaviour, when we construct a mathematical model of a biological system, there are two approaches. A bottom-up approach involves analysis of the detailed experimental data of the components and their interactions and then including only what is deemed necessary. This type of model is not fitted on the system behavior, but rather predicts the behaviour of the system as a function of the interacting components. Alternatively, a top-down approach uses measured system behaviour to construct a model consisting of apriori selected mechanisms. Method choice is based on the purpose of the model and the experimental data available, however a combination of the two approaches can be used during the construction of a mathematical model.

Depending on the scope of the project, a researcher must also determine the elements required of their model for their investigations. Static models are time-invariant and calculate the flux distribution of the system in equilibrium (steady-state), whereas a dynamic models are able to account for time-dependent changes in the system. This often becomes a question of model complexity. Based on the size of the model and the available data, complexity is an important consideration during model design, and boils down to a trade-off between simplicity and accuracy. Metabolic models tend to either contain only a few metabolic reactions described to high kinetic detail (kinetic models) or comprise a large set of reactions with little or no kinetic detail (stochiometric, or constraint-based models). Large models such as genome-scale network reconstructions may seem more biologically relevant as they can account for global inter-dependencies and the effects of multiple linked metabolisms and compartments, but this is often accompanied by the lack of kinetic data leading to a lower model resolution. The yeast consensus model contains more than 1000 reactions, but merely define the network stoichiometry, which can only be studied using Elementary Mode Analysis or Flux Balance Analysis [106]. Kinetic models on the other hand are limited to a smaller number of reactions due to computational issues, such as numerical instability and the computational costs of higher resolution simulations.

There are many kinetic models of yeast as well, however their coverage does not extend far beyond the glycolytic pathway and typically comprise less than 20 reactions, but allow for dynamic simulations and metabolic control analysis. Ideally it would be desirable to construct genome-scale kinetic models, and attempts have been made to do this using methodologies such as linlog rate laws or approximation methodologies, whereby generic rate laws are incorporated [103, 104]. The current methodologies however, do not produce models of a satisfactory calibre and research on this topic is ongoing. Therefore, depending on the focus of the research, when initiating model design at this point in time,

a researcher must decide on a top-down or bottom-up approach and whether to define enzyme kinetics for the model.

2.5 Genome-scale network reconstructions of biochemical networks as an approach to modeling metabolism

2.5.1 Establishing a genome-scale network reconstruction

Traditionally, metabolic network reconstructions were based on primary literature and biochemical characterization [64, 84, 85]. Today, thanks to high-throughput sequencing, we are able to sequence and annotate on a genome level, allowing the creation of genome-scale network reconstructions. This can even be done for organisms with minimal published biochemical data. Network reconstructions consist of a network of linked reactions that are mass-balanced and the system is at steady-state. To construct a network reconstruction, certain knowledge of the enzymes catalysing each reaction is required. For each enzyme, the actionable substrates and products and the stoichiometric coefficients of each metabolite participating in said reaction needs to be known. In practice, these reconstructions occur using automated software to map out data from bibliomic databases. Organism-specific genome annotations can be found in genomic databases such as EcoCyc [116] and SGD [45]. These genome annotations are each directly linked to a metabolic enzyme and indicate how these gene products interact and catalyze metabolic reactions in the network.

Metabolic databases such as KEGG [44] and Transport DB [91] are used to assign reaction specificity to each of the enzymes present in the system. Many of these databases link an enzyme commission (EC) number or transport commission (TC) number to reactions observed in a different organism to that of the model. It is, however, important to note that substrate specificities and enzyme activities can differ between enzymes with the same EC or TC number. Additionally, key information required for the network reconstruction, such as sub-cellular localization and reaction directionality, may not be present. This highlights the importance of curating the initial reconstruction.

2.5.2 Curating the network reconstruction

Automatically reconstructed networks will often be incomplete and require manual curation. There will often be reaction gaps or reaction mismatches, and this can be quite a tedious procedure. Using available literature, organism-specific databases and even consultations with experts that have specialist knowledge of the organism, reaction errors and gaps can be solved. These curated network reconstructions build on the biochemical, genomically and genetically (BiGG) structured knowledge base available to all researchers working on the target organism. The reactions placed in a BiGG knowledge base form a genome-scale network reconstruction (GENRE). These GENREs are formed iteratively as new experimental data is released and the BiGG knowledge base grows [22].

2.5.3 Transforming a genome-scale reconstruction into a computational model

Before a network reconstruction can be used to make simulations, it needs to be converted into a mathematical format that can be computationally read. Converting a GENRE to a genome-scale model (GEM), opens up the model to computational tools that can be used to evaluate the network in its ability to accurately predict outcomes, and where necessary, make changes to the network to more accurately align these prediction outcomes with published data. Necessary for these analyses, is the composition of cellular biomass, the composition of minimal growth media, and a set of experimental data which includes growth rates and substrate uptake rates of the specific organism under a specified set of conditions.

Once in a GEM format, a biomass objective function can be applied to simulate the ability of the network to support growth. The biomass objective function is an artificial reaction added to the model as a representation of the macromolecular components required to synthesize a model-defined amount of an organism. Most commonly, it is the millimole of each component required to form a gram of organism dry weight. Chavali *et al.* [11] describes an approach whereby the dry weight distribution of protein, carbohydrate, DNA, RNA, lipid, and polyamine percentage is determined experimentally or via available literature of the organism or similar organisms. The relative abundance of each molecular component comprising each macro component can then be estimated based on its percentage prevalence. For example, in *Leishmania major*, Chavali *et al.* [11] estimated that the protein fraction constituted about 45 % of the organism dry

weight. To determine the amount of arginine in this fraction, the amino acid sequences associated with each open-reading frame in the *L. major* genome can be acquired from GeneDB. Using the percentage prevalence of each amino acid and each respective molecular weight, we can determine a dry weight distribution for the various amino acids. At a percentage prevalence of 7.19 and a molecular weight of 175.11 g/mole, we can calculate the total mass contribution of arginine as 9.95%. This can then be converted to a mmol/gDW value, indicating the millimole demand of each metabolite to form one gram of dry weight biomass.

There are additional factors to consider during model curation, being that of strain-specific parameters and non-metabolic activities. For example, the stoichiometry of translocation reactions and cell maintenance reactions are major factors that need to be acknowledged [21, 117]. Translocation reactions often need to be considered post-reconstruction as a result of the difficulty to experimentally obtain data on the mass and energy balances of ion-pump components [22]. To analyse the accuracy of the modeled network reconstruction, model simulations can be conducted for a specified set of conditions, of which experimental data has already been published.

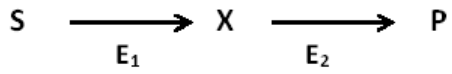
2.5.4 Large-scale data processing of genome-scale models

Completed GENREs and GEMs can be integrated to cover a range of cellular activities with a single model. Network reconstructions can be performed on a variety of an organism's cellular activities. In addition to metabolic processes, network reconstructions can be applied to transcriptional and translation processes, transcriptional regulation, and other signaling systems. Once multiple types of network reconstructions have been established for a given organism, these network reconstructions can be integrated to evaluate genome-scale 'omics data sets with a higher biological relevancy and accuracy. Integration of transcriptional regulation networks (TRNs) and metabolic networks has been a popular choice for researchers [13, 37, 55]. TRNs add an extra step of regulation as active enzyme concentrations are regulated by transcriptional activity. In turn, gene expression is regulated by metabolic precursors of transcription, and a kind of positive feedback is established. The same approach has been taken with integration of sRNA (small RNA) into genomic models. One of the functions of sRNAs is to regulate gene expression through mRNA binding, thus acting together with transcriptional regulation to determine eventual enzyme expression.

2.6 Flux balance analysis as a mathematical approach to biological problems

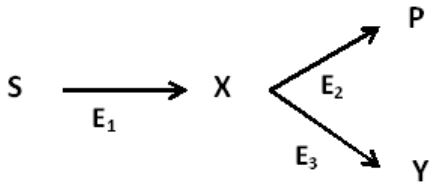
In biochemistry, flux balance analysis (FBA) is a widely used approach for modeling genome-scale biochemical networks [18, 21, 23, 75]. By means of linear programming, FBA mathematically analyses the flow of metabolites through a metabolic network, allowing for the prediction of metabolic outcomes for a given set of parameters. For example, given a set of constraints, FBA can be used to predict the growth rate of an organism for a specific set of biological conditions.

In a simplistic 2-step reaction, a substrate "S", will be converted by enzyme "E_1" into an intermediate metabolite "X". Enzyme "E_2" will then convert metabolite "X" into product "P".



In this reaction the rate of production of product "P" is dependent on how fast enzyme "E_1" can convert substrate "S" into intermediate "X", and how fast enzyme "E_2" can convert intermediate "X" into product "P". At steady-state, these reaction rates are known as steady-state fluxes. Considering the activities of "E_1" and "E_2", whichever enzyme has the slowest activity will determine the rate at which "P" can be formed. Even if "E_1" is able to convert 5 "S" to 5 "X" per minute, if "E_2" is only able to convert 2 "X" to 2 "P" per minute, only 2 "P" can be produced every minute. The same is true for vice versa. If "E_2" is able to convert 5 "X" to 5 "S" per minute, but "E_1" is only supplying 2 "X" per minute, only 2 "P" will be produced per minute. This rate-limiting effect on "P" production by enzyme "E_1" or "E_2", is known as a system constraint.

If we expand our simple reaction network further, introducing a third enzyme, which competes with enzyme "E_2" for intermediate "X", there will be two possible end-products that can be formed from metabolite "X".

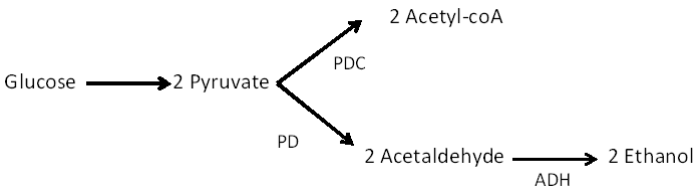


In a dynamic system, the concentrations of substrate "S", and products "P" and "Y" can affect the reaction rates. In a system consisting only of a saturating concentration of substrate "S", reaction throughput by the "E_1", "E_2" and "E_3" reactions will be a lot higher than a system where there is very little available substrate "S" and say, a high concentration of metabolite "X". Assuming reaction "E_1" is reversible, the presence of a large concentration of "X" and a low concentration of substrate "S", may push the reaction flux in the reverse direction, producing "S" from "X". A reaction network is said to be at steady-state when the concentrations of internal metabolites reach a stable concentration, and each reaction displays a constant, net flux. When this system is at steady-state, depending on reaction favourabilities, there will be a flux distribution between enzymes "E_2" and "E_3", represented as steady-state fluxes "J_2" and "J_3". As intermediate "X" can only be consumed as fast as it can be produced, the sum of "J_2" and "J_3" will thus equal "J_1".

$$J_1 = J_2 + J_3$$

There can be multiple different flux distributions for "J_2" and "J_3", all fulfilling $J_1 = J_2 + J_3$. We may wish to optimise the production of metabolite "Y". In a modeled system, we could set an objective function for the production of metabolite "Y", and using FBA, find the most optimal solution for maximizing the rate of "Y" production.

A simplistic branched metabolic pathway can be used to illustrate glycolysis in yeast. During glycolysis, glucose is processed by a pathway of consecutive reactions to form pyruvate. Depending on the growth conditions of a yeast cell, pyruvate ("X") can either undergo aerobic metabolism, converting pyruvate into acetyl-coA, or undergo anaerobic fermentation, whereby pyruvate is processed into ethanol. For the sake of simplicity, the glycolytic pathway has been condensed into a single reaction step.



Each enzyme in the metabolic network will have a maximal and minimal activity. This is known as the upper and lower bounds, and serves as a flux constraint for each reaction. Depending on the parameters we set for a modeled system of yeast metabolism, FBA will determine a flux-distribution for the fate of pyruvate molecules produced by glycolysis. One may wish to optimise their network so that the yeast cell favours ethanol production over aerobic respiration. An objective function for the model can be set to optimise the production of ethanol. Using FBA, an optimal solution space can be calculated for the model, which can reveal key factors responsible for optimising ethanol production. Once a stable, working model has been achieved, a variety of new conditions for the modeled yeast cell can be tested. For example, one can simulate a hypothetical GMO-yeast strain that produces a lower concentration of pyruvate dehydrogenase complex enzymes by constraining the flux through this pathway in the model, and test whether this decrease in enzyme concentration will lead to a higher portion of pyruvate being converted to ethanol than acetyl-coA.

2.6.1 FBA and biology

Using experimental data obtained from various databases, such as KEGG [44], BRENDA [98] and MetaCyc [49], we are able to generate a network reconstruction. These days automated software is readily available that can perform this task efficiently. When converting a metabolic network reconstruction into a computational model, the first step is to mathematically represent the metabolic reactions of the model. This is done by setting up a numerical matrix of the stoichiometric coefficients of each reaction (figure 1b). From this numerical matrix, a system of linear equations can be derived, representing the stoichiometry of each reaction in the network. This set of mass balance equations act as constraints ensuring that the total amount of any metabolite produced must be equal to the total amount consumed at steady state. This is also the first constraint that is applied to the model.

The second constraint on the system comes in the form of upper and lower bounds applied to each reaction in the model (figure 1c). These bounds define

the maximum and minimum allowable flux for each reaction. Reaction flux is a representation of the rate at which each metabolite is consumed or produced by each reaction. These two types of constraints together define the possible flux distributions for a system, also known as the solution space.

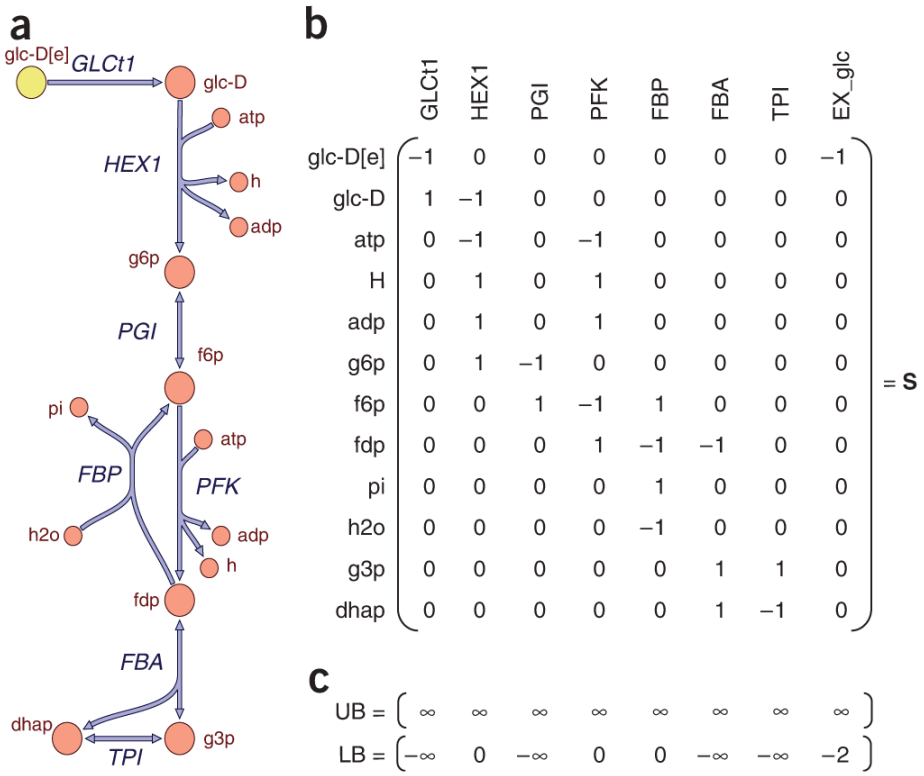


Fig.1 A network reconstruction consisting of metabolic pathways (a) is represented in the form of a stoichiometric matrix (b), where each column displays a reaction in the network, and each row displays the list of metabolites involved in the network. To analyse growth from model simulations, an additional column representing an artificial biomass reaction can be added. Stoichiometric coefficients are assigned to metabolites for each reaction. An upper and lower bound is applied to each reaction (c), defining the maximal and minimal allowable flux through each reaction in the pathway [7].

Most biological network reconstructions are large models where there are more

reactions than metabolites, meaning there is more than one possible solution to the set of linear equations. It is thus necessary to use some kind of optimisation technique to find the optimal solutions. To investigate a desired phenotype in a model, an objective function is defined, which will either be maximised or minimised mathematically. A desired phenotype is investigated by setting an objective function for the model and applying the necessary constraints to the model. For example, for predicting growth rate of an organism, the objective set would be a reaction representing biomass production. In the case of biomass production, an artificial 'biomass reaction' is created that mathematically represents the rate at which metabolites are converted into biomass constituents. In terms of the stoichiometric matrix, an extra column is thus introduced containing stoichiometric values for metabolic constituents that are consumed to simulate biomass production. The composition of this biomass reaction is based on experimental data on biomass components, such as nucleic acids, proteins and lipids. The biomass flux is scaled to the exponential growth rate (μ) of the organism. Predicting the maximum possible growth rate can then be achieved by using FBA to calculate the possible conditions that will result in the maximum flux through the biomass reaction. It is possible that multiple reactions contribute to a phenotype of interest. Mathematically, an objective function quantitatively determines how much each reaction contributes to a phenotype. Linear programming algorithms are used to solve for optimal solutions for a modeled set of conditions. Due to the size of most models, computational linear programming algorithms are necessary to solve for these solutions. Many of these computational algorithms are freely available on the internet or on request by author. Algorithm software, such as the COBRA Toolbox, is able to efficiently solve for optimal solutions to large systems of equations [7]. FBA can also be used to determine a phenotype for a variety of conditions. Constraints can be altered by adjusting the bounds of reactions in the system. Growth rates of an organism can thus be simulated for a range of substrate conversion rates.

2.6.2 Linear programming: The foundation of FBA

Linear programming is a method of optimising a mathematical model, where all variables are represented by linear relationships. An objective function is assigned and optimized, while confined to the linear equality and linear inequality constraints defined in the model.

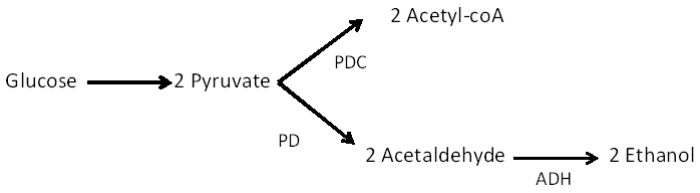
Linear problems can be expressed mathematically in a canonical form:

- maximise $Z = c^{T}v$

- subject to Sv $\leq$ b
- and v $\geq$ 0

where "v" represents the vector of all the variables and "c" defines the weighting of each flux in the objective function. "b" Is the vector of all the constraints, "S" is the matrix of coefficients and "$(.)^{T}$" is the matrix transpose. "$c^{T}v$" Is the expression being optimized ie. the objective function. The inequalities "Sv $\leq$ b" and "v $\geq$ 0" are the constraints depicting the convex polytope over which the objective function is to be optimized. This convex polytope serves as a graphical illustration of the solution space of the model ie. all the possible model outcomes, based on the array of possible flux distributions (figure 2).

Returning to our simplistic branched metabolic pathway, we can illustrate the principles of linear programming.



Our linear objective function is to maximize the production of ethanol. By increasing the proportion of pyruvate produced via pyruvate kinase activity (PYK) consumed by pyruvate decarboxylase (PD) rather than by the pyruvate dehydrogenase complex (PDC), more acetaldehyde can be produced. More acetaldehyde means more ethanol can be produced by alcohol dehydrogenase (ADH).

Glycolysis produces two quantities of pyruvate for every quantity of glucose consumed. The stoichiometry of a reaction network directly applies to the reaction fluxes observed, and as a result, the rate of pyruvate formation is effectively double that of the glucose consumption for a given glycolytic flux. All downstream fluxes will also be affected by this stoichiometric factor and so it is acknowledged when writing out our objective function:

maximize PD: $2\ v_{PD} = v_{PYK} - 2\ v_{PDC}$

with the following constraints:

$2\ v_{PD} + 2\ v_{PDC} \leq v_{PYK}$

The rate of pyruvate consumption between the PDC and PD reactions can only be equal to, or less than half the rate of glycolysis.

$-\ 0.5\ v_{PYK} + v_{PD} \leq -\ v_{PDC}$

The total flux through PDC must be equal to or less than 50% of the glycolytic flux minus the flux through the PD branch

$-\ 0.5\ v_{PYK} + v_{PDC} \leq -\ v_{PD}$

Similarly to the PDC reaction constraint above, the total flux through PD must be equal to or less than 50% of the glycolytic flux minus the flux through the PDC branch

$-\ v_{PD} \leq -\ v_{ADH}$

The flux of ADH can only be equal or less than the flux through PD

and non-negative variables:

$v_{PYK} \geq 0$

$v_{2PD} \geq 0$

$v_{2ADH} \geq 0$

These non-negative constraints define that the reactions PYK, PD, and ADH cannot have negative flux values. Basically, in order to maximise the production of acetaldehyde, it is required that these three reactions have a non-negative net flux.

More complex linear programs are usually expressed in matrix form:

$\max\{c^T v \mid Sv \leq b \; \lambda \; v \geq 0\}$

As described earlier, our objective function "$c^T v$" or "Z" is maximised, within the feasible solution space created by the constraints "$Sv \leq b$" and "$v \geq 0$".

which can be expanded and shown in standard form as:

$$\text{maximise} \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

with constraints:

S x v $\leq$ b , v $\geq$ 0 :

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \leq \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In an FBA-based model, the stochiometric matrix ("S") will consist of rows of reactions ("n") and columns of metabolites ("m"), representing the quantity of each metabolite consumed and produced for every network reaction (figure 1b). The vector of all the reaction fluxes ("v") is multiplied through the rows of reactions in the stochiometric ("S") matrix ("S x v"). These outputs are constrained by the system of inequalities "$Sv \leq b$" and "$v \geq 0$", which together define a set of minimal and maximal bounds for each reaction flux (figure 1c). Based on the bound-restrictions established for each reaction, we can graphically display the reaction vectors and map out a feasible region (figure 2).

By factoring in our constraints, a solution space is defined in the form of a convex polytope. This solution space is more commonly referred to as the feasible region. To solve for optimal solutions to an established objective function ("Z"), a linear programming algorithm can solve for the points in this polyhedron where the objective function value is optimal (either maximised or minimised).

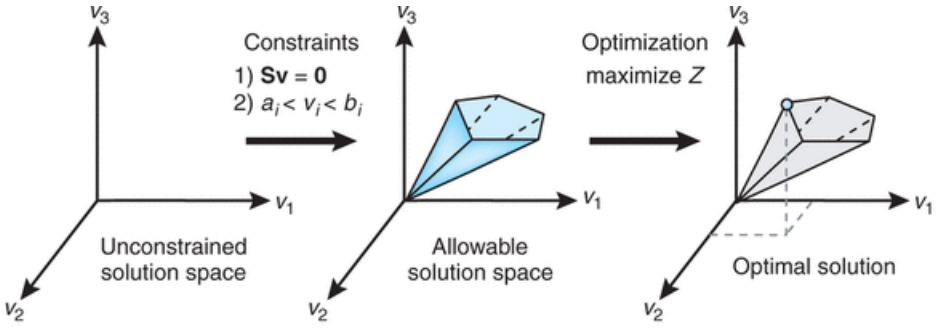


Fig2. Before applying constraints, the solution space (feasible region) is limitless, but after applying constraints in the form of linear inequalities, an allowable solution space is established. Linear programming algorithms then work to find optimal solutions for the system along the edge of the feasible region [80].

With just four reactions and five metabolites in the example above, there are only a handful of variables that can affect the outcome in the optimization of ethanol production. One could determine an optimal solution using just a pen and paper. However, in reality, biological systems are far more complex and the number of reactions can extend into the hundreds and thousands, each with minimal and maximal allowable fluxes (bounds). The potential outcomes of such a system can get exponentially more complex with the introduction of each new variable. There is thus a need for a higher solving power than a man and his pen and paper. Today we make use of solving-algorithms based on the principles of linear optimization to simulate feasible solutions to computationally-modeled systems.

2.7 Genome-scale network reconstructions, then and now

Before the existence of GENREs, constraint-based modeling was used as a method to determine theoretical pathway yields and metabolite overflows for simple metabolic pathways. The ability to sequence whole genomes made it possible to utilise constraint-based modeling at a genome scale, allowing for the assessment of phenotypic functions based on the complete metabolic gene content of an organism. The first GENRE was built for *Haemophilus influenzae* in 1999, demonstrating that the genotype-to-phenotype relationship of metabolic pathways could be discerned mechanistically at a genome-scale [20]. GENREs of model organisms followed soon after, consistently being updated and improved as modeling techniques became more sophisticated and additional data became available. Updates to the GENRE for *Escherichia coli* brought about the development of a set of standards for modeling the relationship between genes, proteins and reactions involved in a particular biochemical transformation [90]. Later, updates to the GENRE for *Saccharomyces cerevisiae* suggested a standard way to describe cellular compartmentalisation [25]. The development of these guidelines facilitated the study of global organization of cellular behavior, such as pathway structure [107], adaptive evolution end points [39], metabolic fluxes [2] and bacterial evolution [82, 86].

The age of "omic" data opened the door to much richer, diverse GENREs. The ability to compare transcriptional changes for a given metabolic network allowed for context-specific modeling, and further aided accuracy in these models. As a result, reconstructions of dynamic metabolism such as human metabolisms [18, 113] and photosynthetic processes [74] were now possible.

Today, highly curated and validated GENREs enable meaningful predictions of biological processes. Lobel *et al.* [62] investigated the role of metabolism in the pathogenesis of *Listeria monocytogenes*. Through constraint-based modeling they were able to contextualize expression data and algorithmically correct for incorrectly up-regulated transcripts where the rest of the associated pathway is inactive. This higher predictive accuracy enabled the research group to focus their experimentation on highly active pathways and to elucidate the role of amino acid metabolism in *L. monocytogenes* pathogenesis via knockout-strain simulations. Interaction networks describe the phenomological interactions between different biomolecules, such as genes, proteins and transcription factors. Szappanos *et al.* [110] were able to discover the mechanistic principles that underlie the global properties of *S. cerevisiae* genetic interaction

networks. Through the use of single-deletion mutants, they were able to disrupt the production of multiple metabolite precursors necessary for cellular growth and show the contribution of associated gene interactions in the functionality of certain genes. As a result, model refinement led to the identification of one of two NAD^+ biosynthetic pathways from amino acids in their model as a source of inaccurate predictions. Growth experiments using mutant strains were able to confirm that this second biosynthetic pathway was not present in *Saccharomyces cerevisiae*. Nakahigashi *et al.* [71] discovered a novel pathway and previously uncharacterized enzymatic functions in *Escherichia coli* directly as a result of genetic perturbation simulations to the central carbon system of *E. coli*. This prompted experimentation with a variety of knockout-strains of *E. coli*. Metabolomic analysis identified a metabolite previously uncharacterized in *E. coli*, "sedoheptulose-1,7-biphosphate", suggesting a novel reaction. It was found that phosphofructokinase carries out this reaction and that glycolytic aldolase can split the seven-carbon sugar into three- and four-carbon sugars. In short, the use of constraint-based modeling led to the discovery of two new catalytic functions of classic glycolytic enzymes.

GENREs have also aided in the field of metabolic phenotype design and biochemical engineering. Yim *et al.* [120] designed an *E. coli* strain able to produce 1,4-butanediol (BDO) at high yields. This is a significant feat, considering BDO is not a naturally occurring compound in any organism. The authors used a pathway prediction algorithm to determine the necessary transformations required to convert an endogenous *E. coli* metabolite to BDO. Using FBA they were able to identify the most favourable pathway based on thermodynamic feasibility, theoretical yield, and topological distance of the pathway from the central carbon system. Another hurdle was getting *E. coli* to produce BDO at a significant rate. Using a knock-out algorithm, a strategy was developed to block out the production of natural fermentation products, forcing the organism to channel carbon flux through BDO production in order to maintain redox balance. The use of constraint-based modeling greatly accelerated the industrial strain design process.

Conventional drug target discovery is based on gene-centric constraints. However, metabolite- and reaction-knockouts are also possible. Using a metabolite-centric approach, enabled Kim *et al.* [47] to search for structural analogues of metabolites essential to the Gram-negative pathogen, *Vibrio vulnificus*, that would inhibit the enzymes that relied on them as substrates. Using a constraint-based model they were able to approach antibiotic discovery from multiple perspectives, and as a result, a compound more effective than current anti-bacterials was identified experimentally.

GENREs and constraint-based modeling has also led to the coupling of various cellular processes. Lerman *et al.* [56] integrated a model of *Thermotoga maritima* metabolism with another model of transcription and translation. The hopes are that an integrated model may address some of the challenges that limited metabolic models face. This integrated model is able to account for the variability of cellular composition at different growth rates, whereas metabolic models only use a biomass function for growth rate optimization. The authors were able to validate predictions of differential experimental transcriptome and proteome levels across varying conditions. With GENRE models such as these, as the content of these models increase the ability of such models to explain and predict biological functions grows in scope.

2.8 Network reconstruction curation conducted by Plata *et al.*

A network reconstruction was created by Plata *et al.*[89] to simulate biomass production rates and changes in metabolite exchange reactions as a result of single and double gene deletions. In their model, enzyme-coding genes of *P. falciparum* were mapped to their corresponding metabolic reactions using well-studied metabolic models, such as the iAF1260 model for *E. coli*, the iND750 model for *S. cerevisiae* and the genome-scale human metabolic network by Duarte *et al [19].* Additional mapping was conducted using the KEGG database. Stoichiometry and compartmentalisation of reactions were conducted. Transport and exchange reactions were added based on PlasmoDB and MPMP databases, as well as additional literature reported values. A biomass-producing objective function was created and added to the model based on a modified version of the yeast objective function found in the iND750 model [19]. Lipid composition was modified as reported for *Plasmodium* by Hsiao *et al* [38]. Amino acid and nucleotide components were adjusted based on proteome and genome expression data obtained by Llinas *et al.* [61]. The percentage prevalence of each ribonucleotide and amino acid across all open reading frames (ORFs) was calculated as the relative frequency of each monomer. Where available, the counts for each ORF were multiplied by their expression levels. The percentage prevalence of dNTPs was derived from the genome A+T content of 80.6% and then converted to mmol/gDW as described by Chavali *et al* [11]. Once the network reconstruction was able to synthesize or import all biomass components, the reactions with no literature support and not essential for biomass production were removed from the network.

Chapter 3

Materials and methods

3.1 Setting up an initial model

A literature search of published genome-scale network reconstructions of *P. falciparum* metabolism was conducted. FAME is web-based modeling tool that combines the tasks of creating, editing, running, and analyzing/visualizing stoichiometric models into a single program [8]. Appropriate models were analysed using the FAME modeling tool, and compared using a set of criteria:

- Does the model contain all the key reactions and metabolites required to accurately simulate intra-erythrocytic *P. falciparum* growth?
- Are all the necessary transport reactions included in the model?
- Are the network reactions compartmentalised?
- Does the model include a biomass objective function to represent parasite growth?
- Does the model include a haemoglobin import function?

3.2 Model curation

Using the framework of a selected network reconstruction, we constructed our model with a focus on the amino acid metabolism in *P. falciparum*. We set a

maximum biomass formation rate based on a determined specific growth rate as well as a maximum rate that haemoglobin can be digested for amino acid supply to the parasite. We then set objective functions on biomass formation rate, haemoglobin catabolism and glucose consumption. Based on model simulation outputs, we assessed the flux outputs of the model and applied constraints to the model to address any model stability issues. This process was repeated multiple times to test the response of the model to the applied constraints.

Model simulations were conducted in the Python programming language, and using the COnstraints-Based Reconstruction and Analysis for Python (COBRA) toolbox [97]. Additional packages were installed including LMFIT, MATPLOTLIB, NETWORKX, NUMPY, LIBSBML, SCIPY AND SYMPY.

3.3 Parasite culturing

P. falciparum (Strain D10) cultures were grown in airtight, standard 250 mL culturing flasks containing 50 mL of RPMI 1640 formulation 6504 medium (10.4 g/L) supplemented with 0.5% *m/v* Albumax II ®, 22.2 mM glucose (33.3 mM glucose final concentration), 25 mM HEPES, 3 mM hypoxanthine, 25 mM sodium bicarbonate, and 50 µg/mL gentamycin sulphate, and pH 7.2 at 37 °C , as described by Trager & Jensen [114]. Cultures were maintained at 4% haematocrit (A+ erythrocytes obtained from anonymous donors through the Western Province Blood Bank, South Africa) and a gas mixture consisting of 3% oxygen, 4% carbon dioxide, and 93% nitrogen, was used to aerate cultures. All reagents were obtained from Sigma-Aldrich (St. Louis, Missouri, USA), unless otherwise stated. RBC's were washed twice in standard medium before incubation.

Culture synchronization was conducted using sorbitol synchronization as described by Lambros and Vanderberg [52] leaving only ring-infected and uninfected RBC's intact. The more mature life stage infected RBCs, containing trophozoites and schizonts would be lysed. Centrifuge-isolated (750 x *g*, 3 min) infected erythrocytes were suspended in a pre-warmed, sterile sorbitol solution in distilled water (5%) at 37 °C for 5 minutes. This creates a 16 hour window of synchronisation approximately (ring stage parasites are observed between the 6 to 22 hour time points in the 48 hour life cycle). Tighter synchronisation windows were obtained using multiple synchronisations at start and endpoints of ring stage, allowing for approximately 90% uniformity in a synchronisation window of about 4 hours.

3.4 Growing *P. falciparum* parasites on a minimal medium

The minimal media used by Liu *et al.* [59] was recreated to test haemoglobin-dependent growth of erythrocytic parasites. RPMI 1640 formulation R8999-04A (US Biological) w/o amino acids and sodium phosphate was used as the base media. R8999-O4A RPMI media (8.59 g/L) was supplemented as in the standard culture media, but with the addition of 0.8 g/L sodium phosphate dibasic that was lacking in the custom formulation. Single amino acids were added in a variety of combinations to test the dependence of parasite growth on the presence of key amino acids (isoleucine, methionine, glutamine, glutamate, cysteine; Sigma-Aldrich). Amino acids concentrations were based on the concentrations found in standard RPMI 1640 media formulation. The pH was adjusted to 7.2 at 37 °C for all custom media created.

3.5 48 Hour life cycle sampling of cultured parasites

After sorbitol-synchronisation of parasites, there is a possibility that parasite growth may be affected (metabolic stress). For this reason, parasites were allowed an additional life cycle to be sure parasite growth was stable post-synchronisation. Parasitemia change and synchronicity was monitored microscopically via blood-slides. Synchronised parasites used for experimentation were harvested during early schizont stage and split into two triplicate lines; one line of triplicates representing low parasitemia (9%) growth, and the other, high parasitemia (21%). Culture lines were split in the schizont stage (40-44h life-stage), accounting for an established 4x multiplication factor during parasite re-invasion. The motivation behind schizont stage splitting is to avoid parasite loss present in early merozoite based splitting, whereby newly released merozoites that have not re-invaded yet are lost during parasite harvesting. A 0h timepoint was sampled immediately after culture splitting. Parasites were allowed to re-invade new red blood cells and timepoint-based sampling was continued at: 24h; 36h; 48h post-split. Samples were immediately flash-frozen in liquid nitrogen and stored at -20 °C. Important to note is that culture media was refreshed after the 24 hour sample point, as 48 hour experiments with no media refresh showed stressed growth not long after 24 hours of incubation for both low and high parasitemia counts. Stressed growth was observed by the

change in colour of culture medium from a bright red to a dark, brownish-red. When microscopically inspected, some intra-erythrocytic parasites appeared as dark, solid balls, often observed in cultures under metabolic stress. This is due to hemozoin accumulation and the pH drop resulting from lactic acid accumulation.

3.6 Extracellular amino acid analysis using UPLC-UV

Timepoint samples used for analysis were defrosted and centrifuged at 1200 x g for 5 minutes to remove large particulate matter and macromolecules. Supernatants were transferred to Nanosep 10K Omega centrifugal devices (PALL Life Sciences ®) for finer filtration to isolate extracellular amino acids by centrifugation at 14000 x g for 12 minutes. Supernatants were then submitted to Central Analytical Facility at Stellenbosch University, South Africa for amino acid analysis. Submitted samples were subjected to the Waters AccQ Tag Ultra Derivatization Kit and placed in a heating block for 10 minutes at 55 °C. Amino acid separation and detection was performed using a Waters Acquity Ultra Performance Liquid Chromatograph (UPLC) fitted with a photodiode array (PDA) detector. A Waters UltraTag C18 (1.7 µm, 2.1 x 50 mm) column was used for the separation, and instrument control and data acquisition was performed by MassLynx software, which integrates the peaks at the defined retention times and plots calibration curves for each amino acid based on the peak response (peak area/internal standard peak area) against concentration. UPLC parameters were as follows: 700 µL/min flow rate, water as solvent A, and acetonitrile as solvent B.

3.7 Model validation

3.7.1 Evaluating model prediction accuracy by comparison to flux values obtained via experimental data

We established objective functions for our model, applied a set of reaction constraints for general model stability and addressed model errors that occurred,

such as infinite-looping of metabolites. We then validated our model simulations by comparing our model outputs with experimental fluxes obtained from a UPLC data set.

3.7.2 Simulation of atovaquone, a popular antimalarial

At concentrations as low as 5 nM, atovaquone has been shown to be able to inhibit 90% of dihydroorotate dehydrogenase (DHODH) activity in previous experimentation[40]. In our model a reaction named "DHORD2_mt" represents the activity of mitochondrial DHODH, responsible for the oxidation of dihydroorotate to orotate. As atovaquone ultimately affects the oxidation of dihydroorotate to orotate by hampering mitochondrial DHODH activity, we simulated the effect of the drug by constraining the flux through this reaction in the model to investigate how biomass production is affected.

Chapter 4

Results

In this chapter we will present the results of our model simulations as well as the results of our experimental work. In literature claims have been made that *P. falciparum* could be grown quite efficiently on a variety of nutrient-restricted media [16, 99, 58]. This is a significant aspect in the investigation of the role of amino acid metabolism in biomass production. We will first present the results of our culturing of *P. falciparum* on a variety of nutrient-restricted media, followed by the UPLC-UV analysis of standard cultures for the main purpose of model validation. We will then move on to the construction of our model and the accompanying simulations, and finally displaying the results of our investigations using the constructed model.

4.1 Minimal media growth

The ability to grow *P. falciparum* on a nutrient-restricted medium, while still maintaining a growth rate comparable to that of standard growth conditions was attempted. If successful, this would allow for a more tractable UPLC analysis of the amino acid metabolic profile of *P. falciparum*, as well as the degree to which the parasite can rely on haemoglobin as a primary source of amino acids.

A range of growth media conditions were used to test the dependence of *P. falciparum* on serum-sourced amino acids for growth. Being able to grow parasites on a minimal medium that still produces similar growth characteristics as

that of standard media would allow for a much clearer picture of haemoglobin metabolism and its importance as an amino acid source for biomass production. It would also allow for a higher resolution of amino acid metabolic profiles acquired from UPLC-UV analysis. Life cycle progression and re-invasion efficiency was monitored for a spectrum of growth media configurations. Culturing of *P. falciparum* in standard RPMI 1640 medium results in approximately four-fold increases in parasitemia after a 48 hour life cycle (figure 3).

For limited media cultures, life cycle progression is retarded (figure 4) and takes roughly 72 hours for merozoites to advance through schizogeny and subsequent re-invasion of daughter merozoites (data not shown). Re-invasion efficiency is also suspected to be lower in limited media cultures, as once all parasites in each culture had progressed through schizogeny, parasitemia counts after re-invasions were 25-40% lower for 5x and I+M culture lines when compared to standard cultures. This seems to be a case of a fraction of intra-erythrocytic parasites not progressing to schizogeny and therefore not producing a new generation of merozoites. In limited media culture lines, some early stage parasites would not progress to mature trophozoites, instead stalling in what appeared as enlarged merozoites or compacted early trophozoites. This phenomenon has been noted in cultures that have been subjected to some form of stress. In this case it is most likely due to a shortage of specific amino acids required for maturation of parasites. Ultimately these factors add up to a higher doubling time and lower specific growth rate for parasites. Under standard culturing conditions we observe a four-fold increase in parasitemia every two days (˜24 hour doubling time). What we observe with the limited media cultures however, is a doubling time of 50.2 hours for 'I+M' and 67.3 hours for '5x' (table 1). Cultures containing only isoleucine as an exogneous amino acid 'I' displayed a negative specific growth rate, indicating additional exogenous amino acid types were required for *P. falciparum* to increase in biomass.

Table 1: Calculated specific growth rates and doubling times for *P. falciparum* grown in a variety of media conditions

Culture line	Specific growth rate (μ)	Doubling time (h^{-1})
Standard 1	0.0302	23.0
Standard 2	0.0297	23.3
Standard 3	0.0284	24.4
Ile,Met,Gln,Glu,Cys	0.0103	67.3
Ile	-0.005	-138.6
Ile,Met	0.0138	50.2

4.2 48 hour sampling of parasite growth

When parasites were inoculated in fresh medium and grown without refreshing the medium, at the 36 hour timepoint, culture serum appeared to be darkening and life cycle progression appeared sluggish when observed microscopically. Blood slides would show unhealthy parasites to appear as dark, compact spheres within erythrocytes (figure 5). Further evidence of stressed parasite growth is shown in appendix A: figure 1, contrasting the healthy progression of parasites to schizonts in comparison with the halted growth of stressed parasites. In follow-up experiments we incorporated a 24 hour media-refresh. This means that UPLC-determined concentrations for 36 and 48 hour samples will not display the changes in metabolite concentrations from the 24 hour timepoint, but rather show the changes in concentrations from effectively "a 0 hour" metabolite concentration profile. As a result of this, metabolite concentrations recorded for 36 and 48 hour samples were normalised to the concentrations obtained at the 24 hour mark. To accomplish this, we used the following formula to determine the concentrations for each metabolite expected if refreshing of culture media did not reset the metabolite concentrations at the 24 hour timepoint:

$$Conc_{exp} = Conc_{36h} + (Conc_{24h} - Conc_{0h})$$

In other words, the expected 36 hour concentration for each metabolite we would observe in a normal culture ($\text{Conc}_{\text{exp36h}}$), is equal to the sum of the recorded concentration at the 36 hour timepoint (Conc_{36h}), and the difference

between concentrations recorded at 24 hours ($\text{Conc}_{24\text{h}}$) and that of fresh medium ($\text{Conc}_{0\text{h}}$).

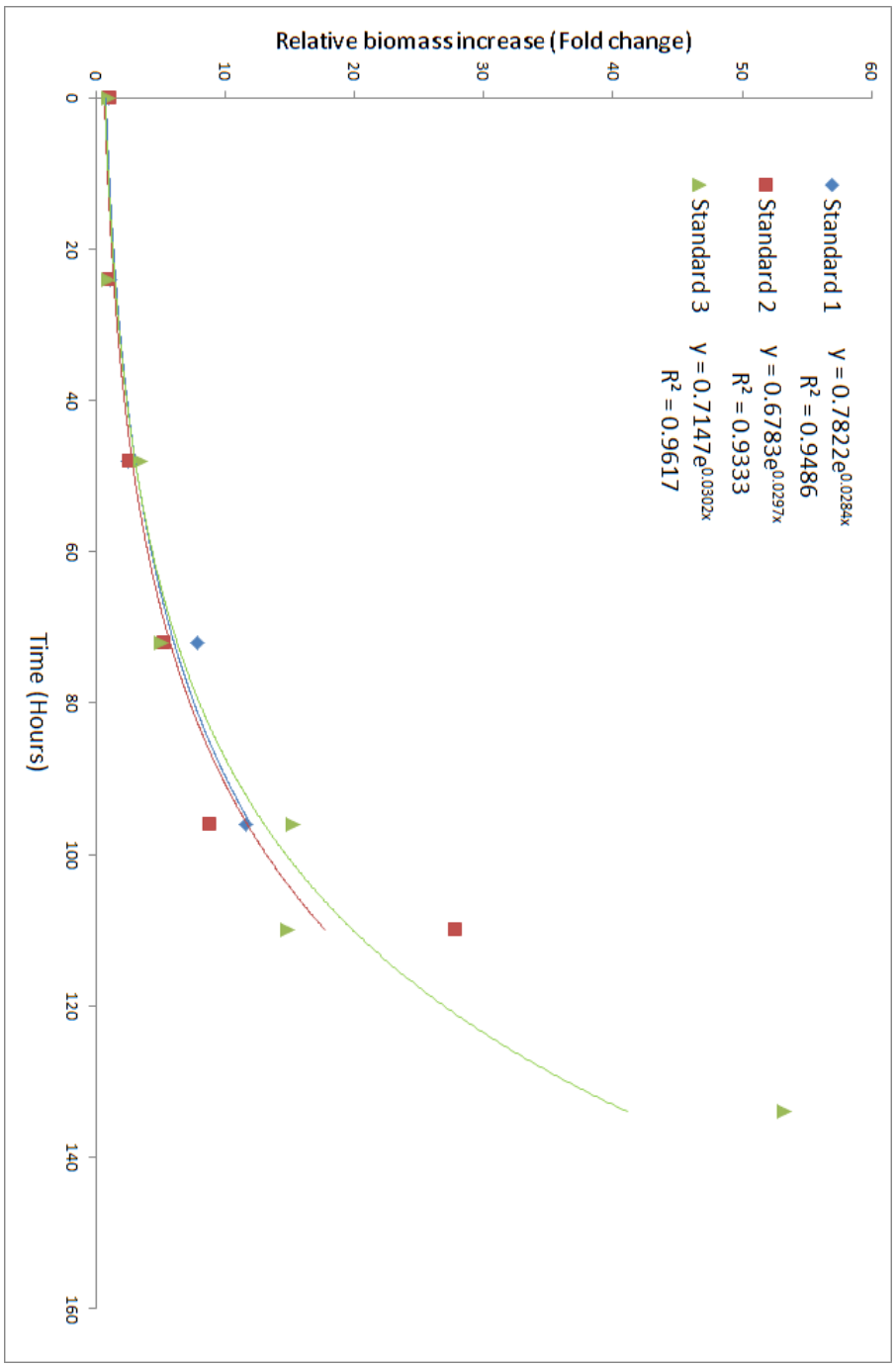


Fig.3 Three independent *P. falciparum* cultures (Standard 1, Standard 2, Standard 3) were grown in standard RPMI 1640 and parasitemia counts were taken at 24 hour intervals. Based on the relative increases in parasitemia over time a projection of biomass growth was generated.

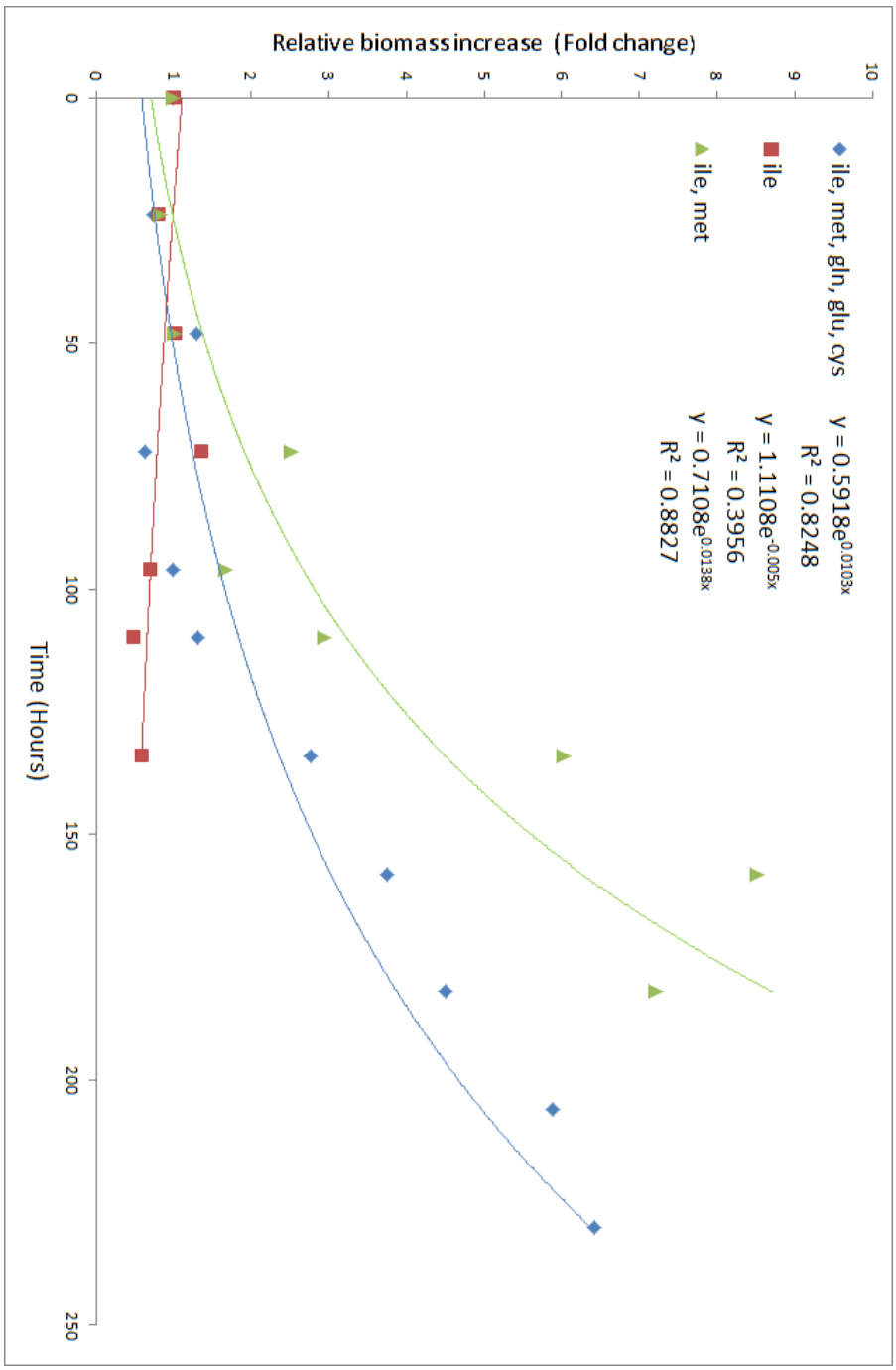


Fig.4 Growth rates observed for *P. falciparum* grown in a variety of limited media, with 24 hour sampling intervals. Parasites were cultured in three limited RPMI 1640 media conditions: The suspected core 5 amino acids (isoleucine; glutamine; glutamate; cysteine; methionine) 'Ile,Met,Gln,Glu,Cys', an isoleucine only line, 'Ile', and a line containing isoleucine and methionine 'Ile,Met'.

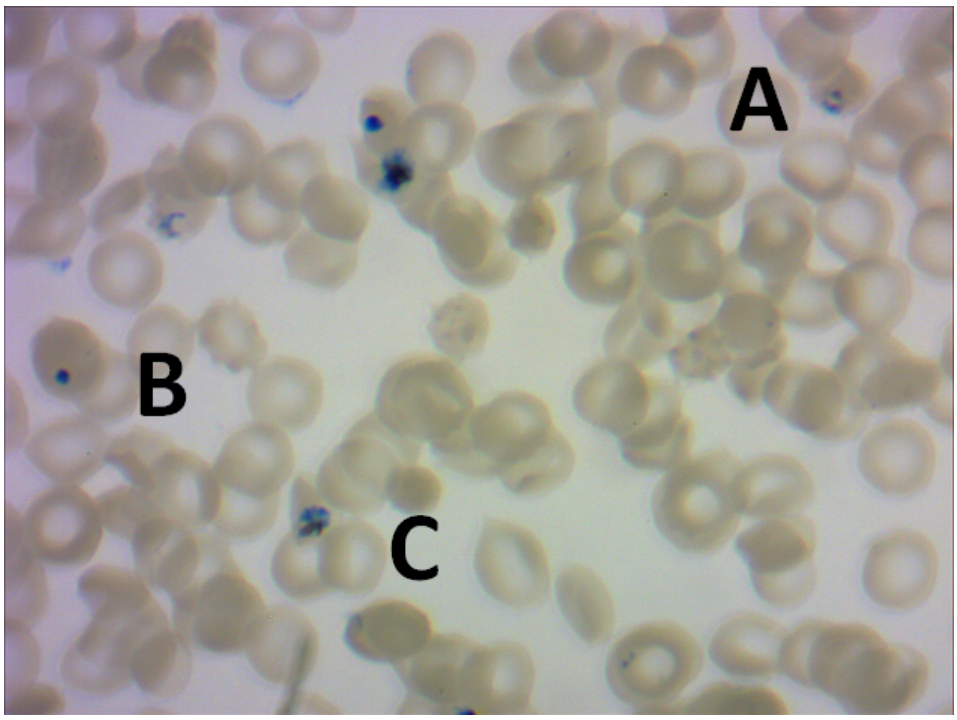


Fig.5 Microscope image captured of a blood slide of a *P. falciparum* culture undergoing metabolic stress. "A" = healthy ring-stage parasite; "B" = parasite exhibiting metabolic stress, observed as a solid, compact mass; "C" = healthy trophozoite-stage parasite.

4.3 UPLC-UV analysis

To test the accuracy of the supplier-stated concentrations for RPMI media, concentrations of amino acids calculated for RPMI 1640 growth media (based on information supplied by Sigma-Aldrich) were compared to UPLC analysed concentrations of prepared RPMI medium (figure 6). Furthermore, controls of known concentration of the five core amino acids (glu; gln; met; ile; cys) were analysed (figure 7). These controls would be run along-side experimental samples.

Six *P. falciparum* cultures (3 x 9% parasitemia; 3 x 21% parasitemia) were sampled at 12 hour intervals for 48 hours and immediately flash-frozen in liquid nitrogen. Additionally, parasitemia and parasite health were monitored microscopically for each sample. The sampling of cultures was initiated when parasites were entering the late-schizont stage of development to avoid the potential loss of merozoites yet to invade erythrocytes that could occur if sampling was initiated during the merozoite stage. Samples from high parasitemia and low parasitemia

culture lines were submitted to the Central Analytical Facility at Stellenbosch University, for UPLC-UV analysis (figure 8). Chromatograms from the amino acid analysis provide a visual display of the amino acid concentrations represented by the peak heights for amino acids eluting at their respective retention times. Figure 9 contains the chromatograms of two controls included in the analysis, five amino acids of known concentration (above) and standard RPMI 1640 (below). Figure 10 provides a visual comparison of the changes in peak heights over 48 hours for *P. falciparum* lifecycle. Observed is general accumulation of amino acids over the 48 hour timeframe, such as ornithine, glutamate and alanine. Amino acids such as arginine and isoleucine are visibly decreased in concentration by the 48 hour timepoint.

Concentrations of the majority of amino acids were increased by the end-stage of parasite maturation, barring isoleucine, methionine, arginine, glutamine, cysteine and aspartic acid, which decreased. *P. falciparum* metabolic activity is known to be highest during the trophozoite stage of development [77]. Our experimental data reflects this where we see large consumption of glutamine between the 24-36 hour mark. In line with this we observe a rise in exported glutamate and alanine. This accounts for the 24 hour timepoint shift by the parasite from glucose being favoured for the production of TCA-intermediates to glutaminolysis-derived α-ketoglutarate from glutamine [76]. Along with the increase in glutamine metabolism from 24-36 hours, we also see a high rate of arginine depletion from the medium, and a resultant spike in ornithine export. This is mostly due to the increase in the parasite arginase activity [77]. Parasites are at a late trophozoite stage at the 36 hour mark, so between 36-48 hours, we expect to see an increase in the rate of energy metabolism and macromolecular biosynthesis as parasites enter the process of schizogeny. We observed a linear consumption of isoleucine throughout the first 36 hours of growth, and then consumption increased slightly during the late trophozoite stage of growth. In line with this, we saw a large increase in accumulation of alanine, glutamate and ornithine for this final 12 hour period of parasite maturation, but also to a lesser extent, leucine, histidine, glycine, threonine, proline, valine, and phenylalanine (figure 8).

Arginine metabolism seemed to be the highest of all amino acids, most notably between the 24-36h time-points, which was reflected closely by ornithine accumulation (figure 11). An interesting observation that after correcting for the 24 hour media refresh, as observed in figure 8, we predict that high parasitemia cultures of *P. falciparum* (>10% parasitemia) will deplete exogenous arginine supply if left to grow in the same medium for longer than 24 hours. Combining the arginine and ornithine concentrations as a total concentration indicate a

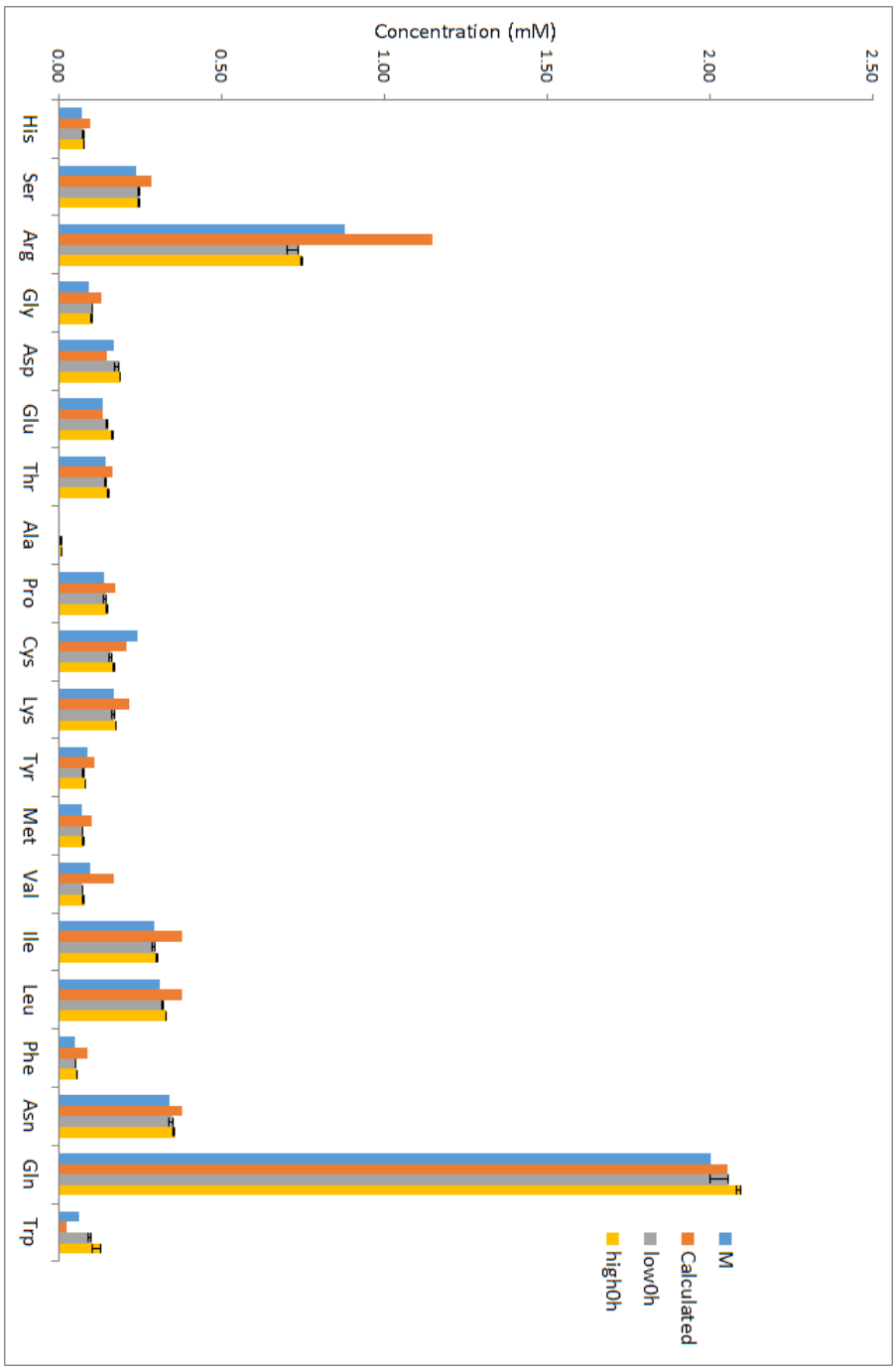


Fig.6 Concentrations of amino acids, as stated by Sigma-Aldrich (St. Louis, Missouri, USA), were compared to UPLC-UV analysed samples of standard culture media. "M" = un-inoculated RPMI 1640 medium; "Calculated" = millimolar calculated concentrations based on information supplied by Sigma-Aldrich (St. Louis, Missouri, USA); "low0h" = 9% parasitemia culture right after inoculation; "high0h" = 21% parasitemia culture right after inoculation.

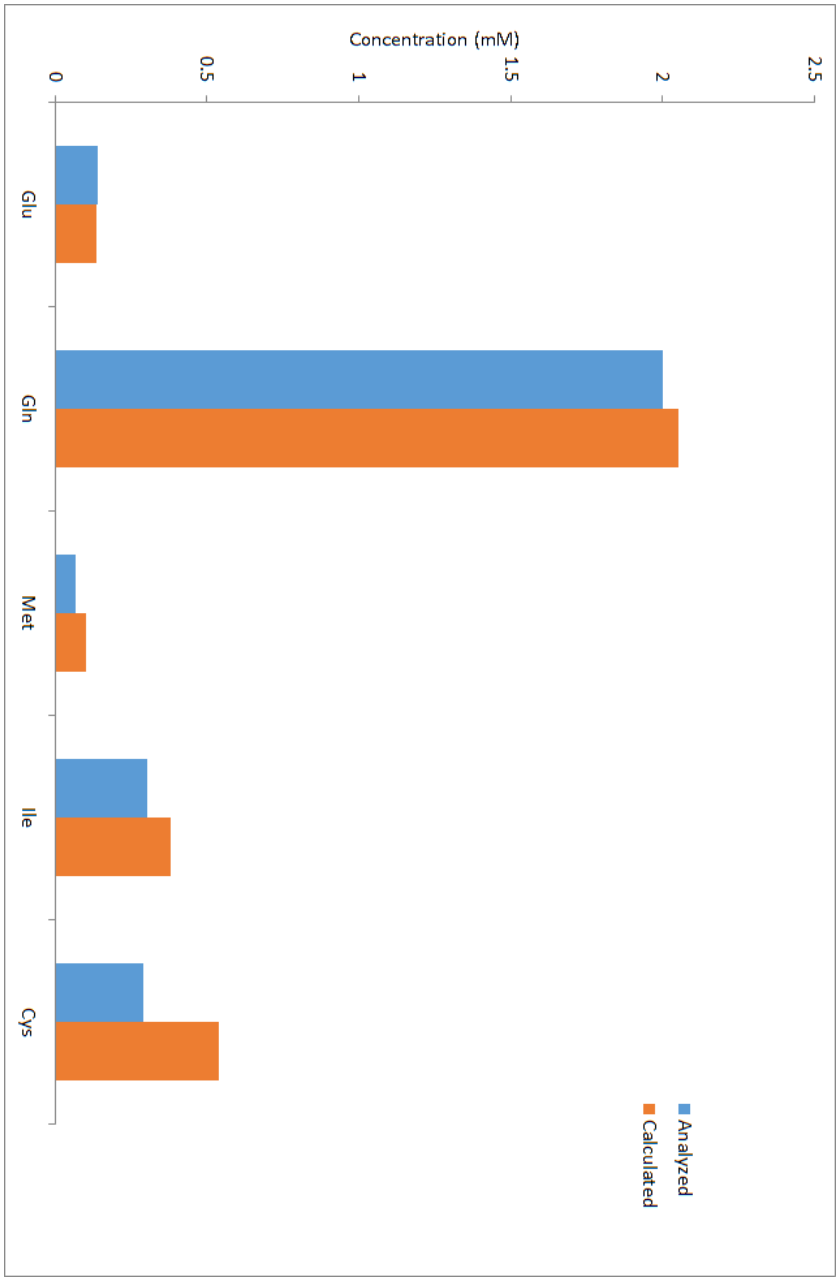


Fig.7 Control samples of known concentration ("Calculated") were analysed by UPLC-UV ("Analyzed") and compared to evaluate the accuracy of UPLC-UV analysis.

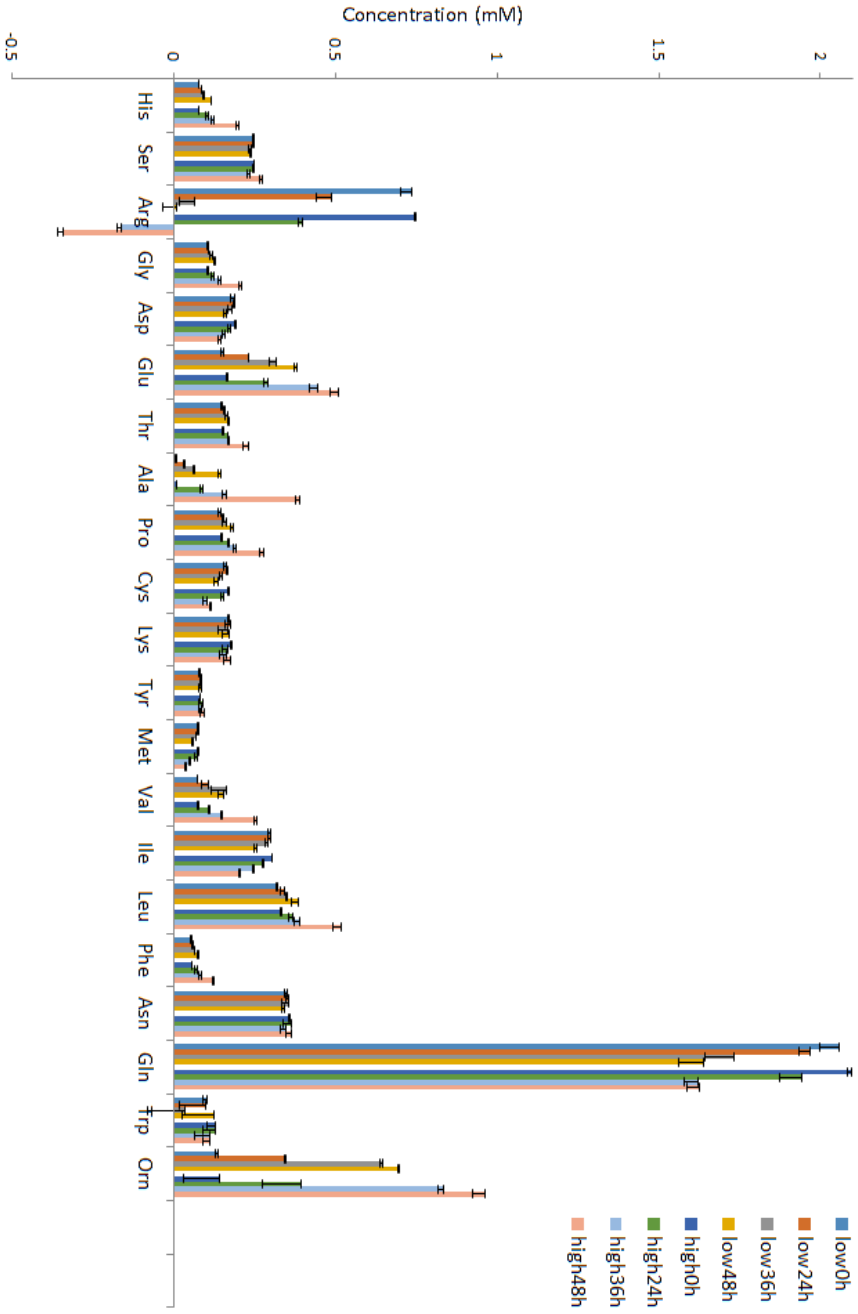


Fig.8 Amino acid concentration profile over a single 48 *hour P. falciparum* lifecycle, accounting for the reset of serum amino acid concentrations due to a media refresh at the 24 hour halfway point. Amino acids were analysed for 3 lines of 9% parasitemia cultures ("Low") and 3 lines of 21% parasitemia cultures ("High") and the average concentrations for each are shown.

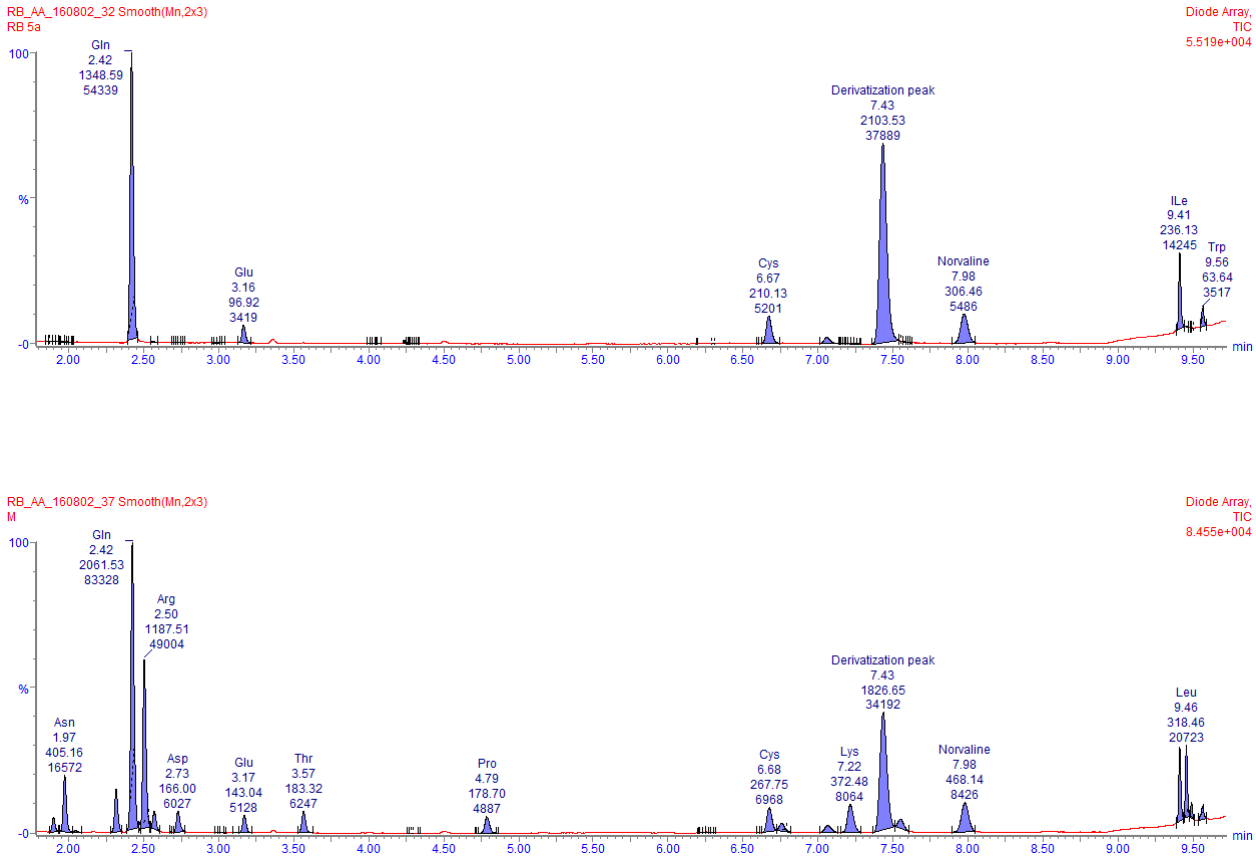


Fig.9 Integrated chromatograms generated from UPLC-UV analysis of five amino acids (gln; glu; met; cys; ile) in phosphate buffer (above), and fresh RPMI 1640 culture medium (below).

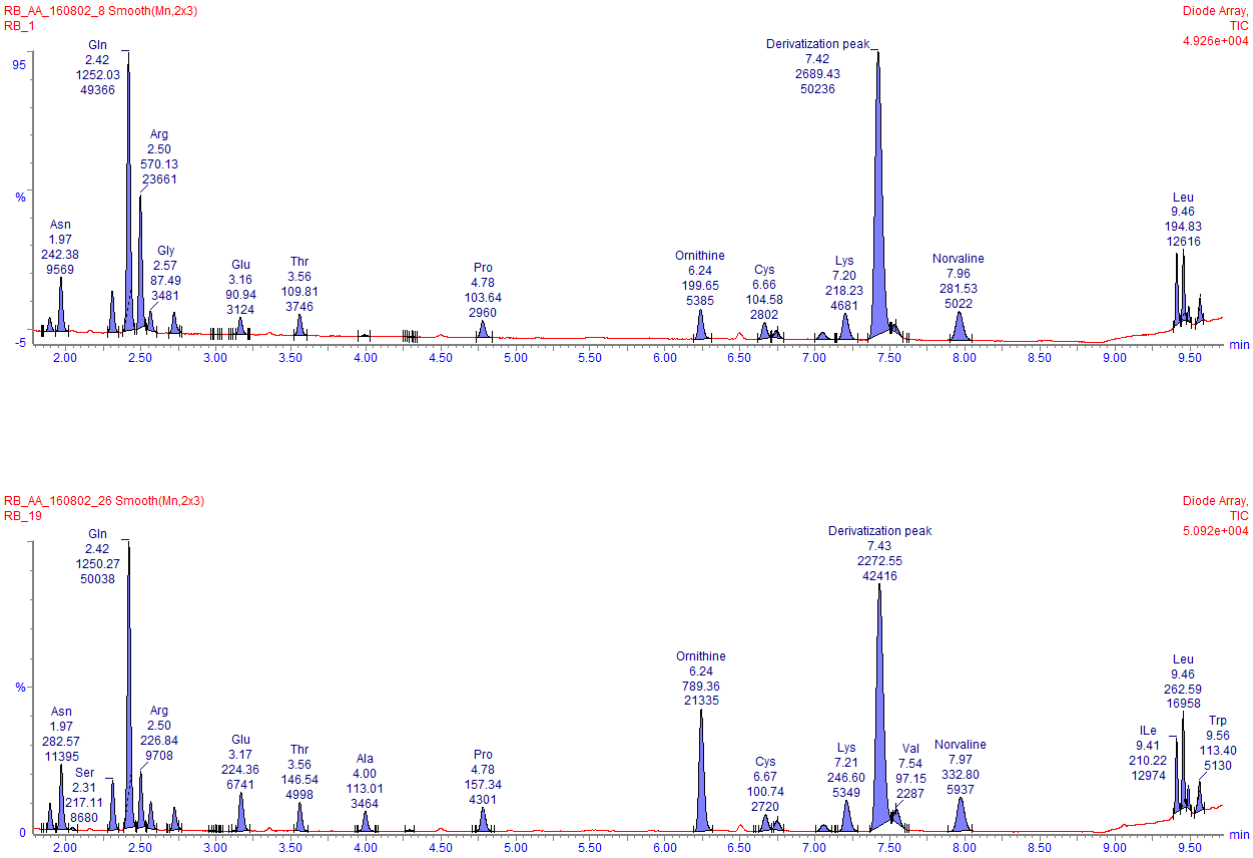


Fig.10 Integrated chromatograms generated from UPLC-UV analysis of amino acid profiles for timepoint zero (above) and 48 hour endpoint (below) of a 9% parasitemia *P. falciparum* culture.

strong favourability towards arginase primarily processing arginine, where the gradual loss in total concentration ("Low48h"; "High48h") is most likely attributed to minor conversion of arginine to citrulline by an arginine deiminase (figure 11).

4.4 Model Construction

The network reconstruction conducted by Plata *et al. [89]* was used as a starting point for the construction of a model to investigate *P. falciparum* growth dependencies in relation to amino acid metabolism. The first step was making sure all network reconstruction data was successfully imported to allow for model simulations. In order to recreate the core growth metabolism of the parasite, we would have to recreate the mass conversion of glucose to lactate observed in the intra-erythrocytic parasite, in conjunction with the rapid catabolism of haemoglobin featured in the parasite's complex amino acid metabolism. The terms "exchange reaction" and "transport reaction" will appear quite often in this section. Before explaining our process of model construction we will explain the difference between these two terms. An exchange reaction represents the net flux of a metabolite between the model organism and the extracellular environment, whereas a transport reaction is a reaction responsible for the transfer of a specific metabolite between different intracellular compartments. Exchange reactions are thus important indicators of the modeled system's decisions to import or export a variety of metabolites and the magnitude at which it occurs.

This would require us to set our solving algorithm to prioritize the digestion of haemoglobin for amino acid metabolic requirements, while still optimizing for biomass production. The primary objective functions of the model were set for optimising biomass production and haemoglobin digestion. A minor prioritization was also set for glucose uptake to ensure its favouring as a carbon source. This was done by setting an additional objective function on the glucose exchange reaction to optimise glucose import. The biomass objective function was, however, still the primary optimisation and this was maintained by setting optimisation weighting values of the biomass reaction, haemoglobin reaction, and glucose exchange reaction to 1.0, -1.0, and 0.01 respectively.

The core function of the model was to prioritise biomass production via the harvesting of amino acids provided by haemoglobin catabolism and glucose as the primary carbon source. Any insufficiencies in amino acid supply through

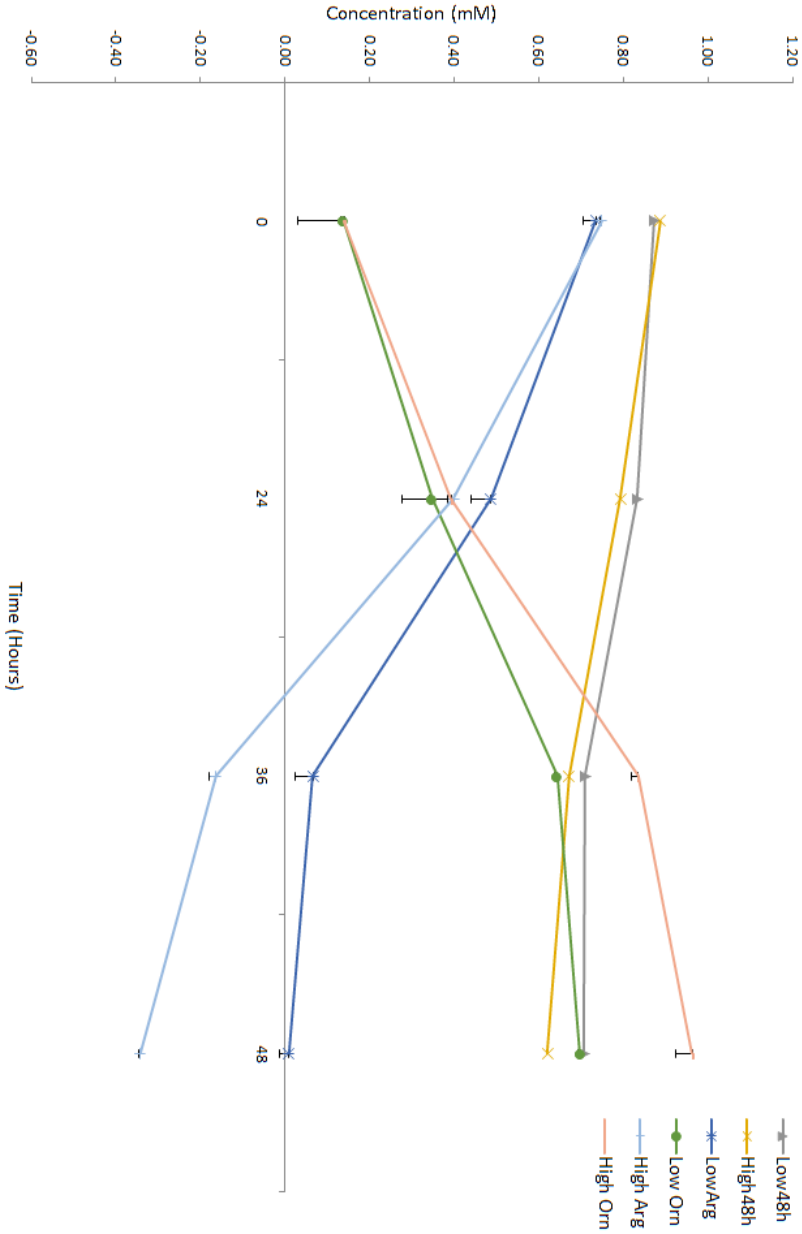


Fig.11 Concentration shifts for arginine and ornithine over a 48 hour life cycle of *P. falciparum*, accounting for a 24 hour media refresh. "Low" = 9% parasitemia cultures. "High" = 21% parasitemia cultures. "Low48h" and "High48h" are summed concentrations of arginine and ornithine concentrations.

haemoglobin would then be made available via amino acid exchange reactions mimicking exogenous amino acid supply. Excess amino acids produced via haemoglobin degradation would be exported via these same exchange reactions. In essence, we would be able to predict a metabolic profile of the amino acid metabolism of *P. falciparum* in a minimal medium environment, using the the amino acid exchange reaction fluxes simulated by the model. Some biologically relevant constraints were added to the model, especially for haemoglobin digestion rates and *P. falciparum* growth rates.

In the boundless state of the model, there are no biological constraints on the model (no maximal or minimal reaction rates have been defined), and one can see how the model attempts to consume as much glucose as it possibly can, as this is what the model has been programmed to do (figure 12). Constraints therefore needed to be applied so that the output fluxes of exchange reactions representing the core metabolism of the parasite were more in line with what can be observed in nature. In the following sections each of these constraints are detailed, and how they were determined.

4.4.1 Some initial calculations

Model flux outputs are represented in the units mmol / gDW / hour and model simulations would be conducted under the conditions of a 50 mL, 9% parasitemia culture at 4% haematocrit. Applying any biological constraints would therefore require conversion to this specific set of conditions.

As seen in figure 12, without defining an acceptable rate of biomass formation ("B"), the model will produce biomass at a rate that is only constrained by the rate at which it can acquire the necessary components. Without any constraints on the exchange reactions providing these building blocks, import fluxes and export fluxes appear at rates at the maximal and minimal bounds established for the model. It is thus important to parameterize the model, and establish some fundamental metabolic rates. We started by determining the specific growth rate ("μ") for *Plasmodium falciparum.*

Specific growth rate:

$$\frac{dB}{dt} = \mu B$$

$\int \frac{1}{B} dB = \mu \int dt$

$ln(B) = \mu t + c$

$B(t) = e^{\mu t + c} = e^c e^{\mu t} = B_0 e^{\mu t}$

$B(t) = B_0 e^{\mu t}$

A young intra-erythrocytic merozoite has an approximate volume of 5 femtolitres and over a 48 hour life-cycle progresses to a schizont of 70 femtolitres (14 times increase) [17]. Substituting in these parameters, we can establish a maximal specific growth rate:

$14 = 1e^{48\mu}$

$ln(\frac{14}{1}) = 48\mu$

Specific growth rate: $\mu = \frac{ln(14)}{48} = 0.0550h^{-1}$

This rate of $0.0550h^{-1}$ was used to define the maximal amount of biomass that could be formed per hour in the model.

As the model uses the unit of mmol/gDW/h, we needed to determine the dry mass of a single parasite, which could then be scaled up to the parameters established for the model. We calculated the parasite dry weight for a 9% parasitemia culture in a 50 mL volume. Penkler *et al.* [87] established that the intracellular volume of a trophozoite had a volume to protein dry mass ratio of $= 4.67\mu L.mg^{-1}$, which when inverted is $= \frac{1}{4.67} = 0.214mg.\mu L^{-1}$. Assuming a 45% protein mass of total mass, we estimated a biomass per volume ratio of $\frac{1}{0.45}0.214 = 0.476mg.\mu L^{-1}$ [11].

Using 5 fL as the volume of a single merozoite [17], we can calculate the dry mass of a single merozoite using our established biomass per volume ratio:

$5\ fL \times 0.476 \times 10^{-9} mg.fL^{-1} = 2.38 \times 10^{-9} mg = 2.38 \times 10^{-12} gDW$

The total dry parasite mass in a standard 9% parasitemia can be calculated using the dry mass of a single merozoite. Our standard cultures consisted of

growing parasites at 4% haematocrit in 50 mL culture flasks. Therefore 2mL of this volume is expected to be compact red blood cells. At an average volume of 90 fL per erythrocyte [73], there will be about 2.22×10^{10} total red blood cells.

Culture parasitemia is determined as the percentage of infected erythrocytes of total erythrocytes. Given a 9% parasitemia culture, there will be about 2.00×10^9 infected red blood cells. Above, the average mass of a single merozoite was calculated to be 2380 femtograms. Multiplying this mass by the 2.00×10^9 infected red blood cells, a total dry mass of 0.00476 g parasites can be estimated for a 9% parasitemia culture at 4% haematocrit.

The next goal was to estimate the rate of haemoglobin catabolism. In a 48 hour parasite life-cycle, approximately 75% of haemoglobin is digested by the parasite [63]. Using an established mass of haemoglobin per red blood cell of $\mathbf{27.2e^{-12}}$g [88] and 64458.0 grams per mole as the molecular weight of haemoglobin, a haemoglobin digestion rate in millimoles per hour can be calculated if the rate was linear over time.

Assuming haemoglobin consumption is proportional to the amount of parasite biomass ("B"),

$$\frac{dHb}{dt} = -cB(t)$$

$$\frac{dHb}{dt} = -cB_0e^{\mu t}$$

We can express this rate of haemoglobin digestion as: $Hb(t) = Hb(0) - cB_0e^{\mu t}$, where we can calculate our haemoglobin remaining at a given time "Hb(t)" is equal to our starting haemoglobin in a given erythrocyte "Hb(0)" minus the haemoglobin digested by the relative amount of parasite biomass calculated for the given time ("c").

Using the mass of Hb per erythrocyte as = $\mathbf{27.2e^{-12}g}$ [88], and MW_{Hb} = 64458.0 g/mol, we can calculate the moles of Hb per cell as:

$\frac{Hb_{cell}}{MW_{Hb}} = \frac{27.2\times10^{-12}}{64458} = 4.2 \times 10^{-16}\text{moles} = 4.2 \times 10^{-13}\text{mmoles}$

Therefore, assuming complete Hb consumption for a 48 hour life cycle:

$Hb(0) = 4.2 \times 10^{-13}\text{mmoles}$

$Hb(48) = 0.0\text{mmoles}$

Using the values at time-points 0 and 48 hours, we can calculate the mmoles of haemoglobin digested per single intra-erythrocytic parasite over a 48 hour life cycle:

$Hb(48) = Hb(0) - cB_0e^{\mu(48)}$ $(mmol = mmol - (c)(gDW)(e^{(h^{-1})(h)}))$
$0.0 = 4.2 \times 10^{-13} - ((c)(1.28 \times 10^{-12})(e^{0.0550(48)}))$
$1.79 \times 10^{-11}(c) = 4.2 \times 10^{-13}$ (mmol = (c)(gDW)(e))
$c = 2.346 \times 10^{-2} mmol/gDW$ (per 48 hour life cycle)

Using a linear rate of haemoglobin consumption over the 48 hour life cycle, we calculated the mmoles of Hb digested per gram of parasite dry mass per hour (mmol/gDW/h) by dividing by 48, to give us a haemoglobin digestion rate of 4.89×10^{-4} mmoles/gDW/h.

4.4.2 Applying constraints to the model

While defining haemoglobin catabolism parameters, it was found that the haemoglobin catabolising reaction in the model produced only half the amino acid constituents of a full human haemoglobin tetramer as a result of incorrect haemoglobin composition in the model file, as published in [89]. Additionally the stoichiometry of haemoglobin erroneously included two extra methionine components. This was found to be the start codons which are cleaved during post-translational modification of haemoglobin and were thus deducted from the number of methionines produced per haemoglobin consumed.

The core constraints were applied to the model as calculated above and then examined model simulations for any remaining fluxes that were extremely high. It was observed that the model was still finding alternative pathways to obtain metabolites for growth and energy metabolism. From the extremely high import

fluxes in figure 13, we can determine the reaction pathways that would need to be addressed.For the sake of model stability, even a boundless reaction with a theoretical maximum flux of positive or negative infinity requires an upper and lower bound to prevent model simulation crashes in the result of flux calculations to infinity. Upper and lower bounds of 1 000 000 mmol/gDW/h and -1 000 000 mmol/gDW/h were set for all reactions. Therefore, when reaction fluxes reaching one million mmol/gDW/h are observed, it is understood that the reaction flux is approaching infinity. This is an un-natural flux calculation, and is addressed by pathway flux analysis to identify problematic reactions causing these un-intended flux predictions. It is important to note that the application of the following constraints mentioned hereafter are implemented for the sake of model stability and do no affect the amino acid metabolism in the model in any way. A large portion of the metabolites shown in figure 13 are involved in the glycolytic pathways of the parasite, so our first strategy was to lock these metabolites within glycolysis and the TCA cycle by preventing their export out of the cell or any import of pathway intermediates. This was done by constraining the exchange reactions for any metabolites involved in glycolysis or the TCA cycle that displayed a tendency for significant import or export. The import reactions showing extremely high fluxes (EX_akg_e, EX_ade_e, EX_dgsn_e, and EX_ins_e), were addressed initially, as many of the export reactions with extremely high fluxes were directly linked to the activity of these import reactions (figure 13). Simulations were conducted again and any new problematic exchange fluxes that appeared were addressed. This process was conducted until natural carbon metabolism was observed, where no intermediates were being imported or exported unnecessarily (figure 14).

We then addressed any other metabolites that were entering or leaving the cell at a very high rate. We observed a very high rate of oxygen and water uptake. This was attributed to the superoxide dismutase and glutathione peroxidase reactions. The model was set to only take up oxygen and only export carbon dioxide, and the peroxide exchange reaction was turned off.

The biomass reaction requires a purine source to produce biomass units, which can be acquired via inosine, adenine or hypoxanthine. However adenosine and inosine were shown to be potential metabolites for carbon metabolism, so they were constrained to zero and hypoxanthine was used exclusively.

As it is not a component of haemoglobin, isoleucine import is required for biomass production and the exchange reaction was set to import only.

Infinite loops can occur during the simulation of network models, whereby a product of one reaction is used as the substrate of another reaction, however

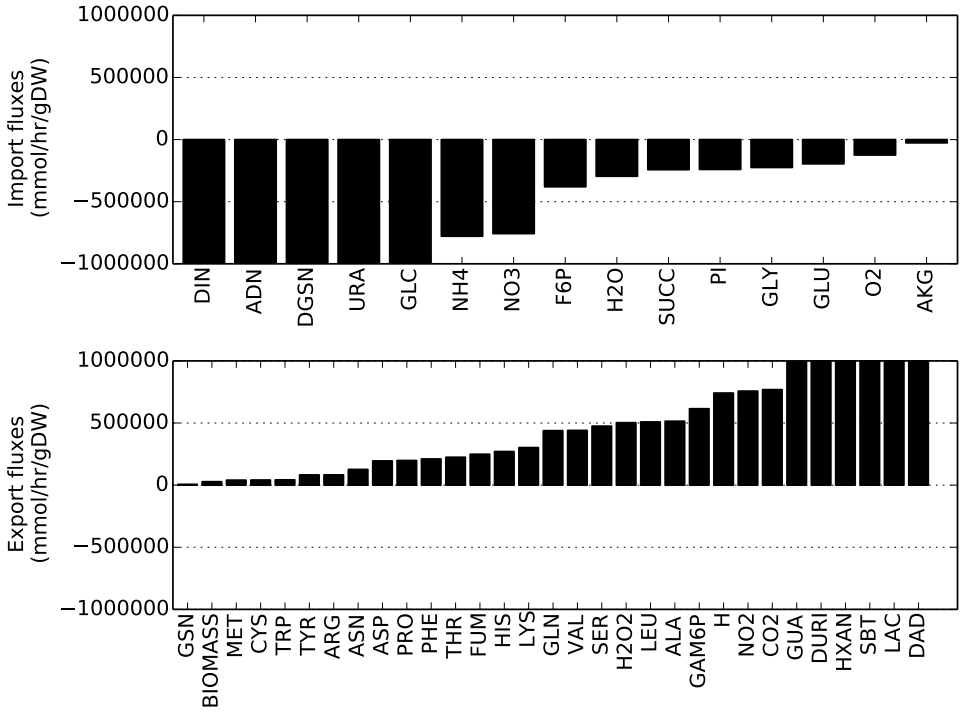


Fig.12 Flux distribution output of the boundless model after establishing the model objective functions for glucose catabolism, haemoglobin catabolism and biomass production. DIN = deoxyinosine; ADN = adenosine; DGSN = deoxyguanosine; URA = uracil; GLC = glucose; NH4 = ammonium; NO3 = nitrate; F6P = fructose-6-phosphate; H2O = water; SUCC = succinate; PI = inorganic phosphate; O2 = oxygen; AKG = alpha-ketoglutarate; GSN = guanosine; FUM = fumarate; H2O2 = peroxide; GAM6P = glucosamine-6-phosphate; H = protons; NO2 = nitrite; CO2 = carbon dioxide; GUA = guanine; DURI = deoxyuridine; HXAN = hypoxanthine; SBT = sorbitol; LAC = lactate; DAD = deoxyadenosine.

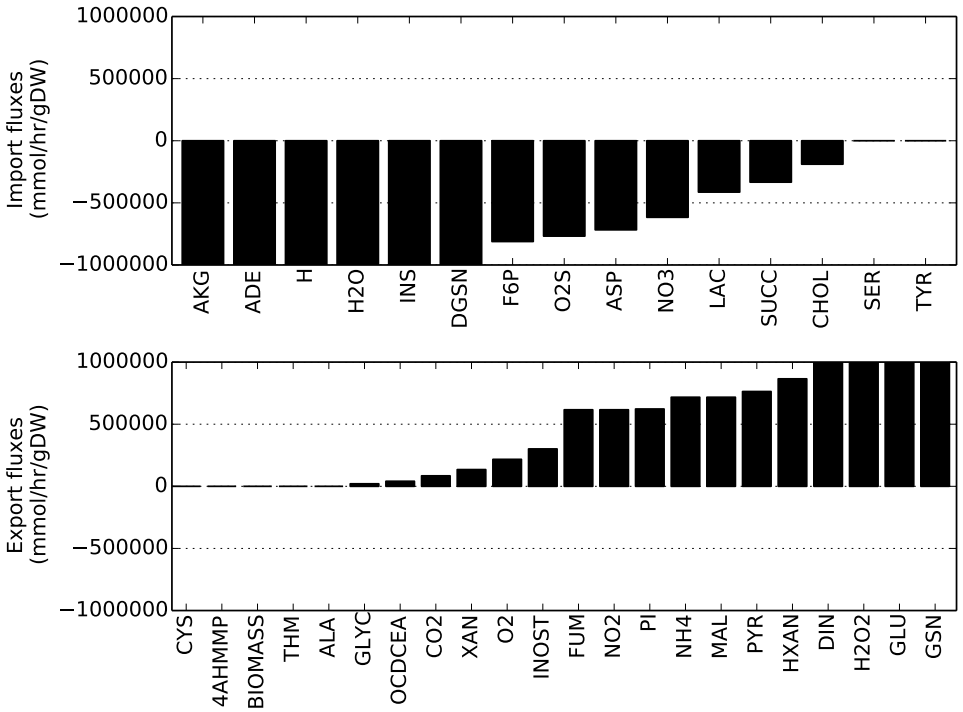


Fig.13 After applying the biological constraints to the core exchange reactions (haemoglobin digestion and biomass production), displayed are the flux values still above 0.05 mmol/h/gDW in magnitude. ADE = adenine; INS = inosine; O2S = superoxide anion; CHOL = cholesterol; 4AHMMP = 4-amino-5-hydroxymethyl-2-pyramidine; THM = thiamine; GLYC = glycerol; OCDCEA = octadecenoate; XAN = xanthine; INOST = myo-inositol; MAL = malate; PYR = pyruvate; DIN = deoxyinosine.

the product of the second reaction is the substrate of the first reaction, resulting in a metabolite being locked into an endless loop. These infinite loops are able to accumulate the entire pool of a metabolite and break model simulations. To prevent infinite looping, some reactions were set to either forward/reverse only, and in some cases turning off reactions completely.

We were still observing fluxes indicating extreme metabolic activity. Lactate production was at the upper bound of the model while the rate of glycolysis was occurring at a much lower, acceptable flux. This was initially observed with a mass import of glutamate and aspartate, and a concomitant export of glutamine and lactate (figure 14). The amino acid antiporter reaction "AKGMAL" was set to zero to prevent the looping of malate and α-ketoglutarate. This antiporter is part of a coupled-transporter known as the malate-aspartate (M-A) shuttle. The second transporter coupled in this M-A shuttle is a glutamate-aspartate

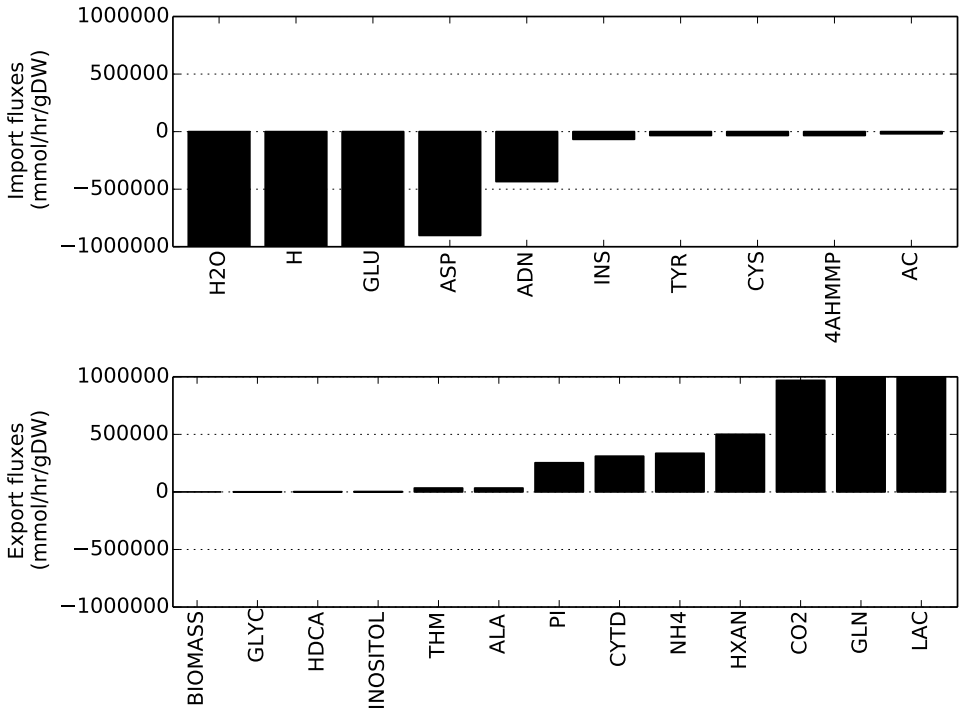


Fig.14 Exchange fluxes above 0.05 mmol/h/gDW had been addressed and the model was updated. Displayed in this figure are the unnaturally high exchange fluxes still simulated by the model. From this we addressed infinite loops, identified initially from the large flux outputs for glutamate, aspartate, and glutamine. AC = acetate; HDCA = hexadecanoate; CYTD = cytidine.

transporter. In our model, glutamate and aspartate were being mass-imported for the downstream transamination of alanine for pyruvate production, which would then be used for biomass production.

4.4.3 Model simulations

Initial simulations of our curated model proved to successfully make use of glucose as a primary carbon source, and haemoglobin as the primary amino acid source, to produce biomass, while importing any lacking (glutamine) or missing (isoleucine) biomass components. There were however some unexpected fluxes. The model seemed to highly favour a selection of amino acids for biomass generation, in order of rank: glutamine, asparagine, aspartate, glutamate. The asparagine and aspartate utilisation is odd as no literature supports this notion.

As the model would import as much glutamine as it could for growth, we had to establish an upper bound constraint on the glutamine exchange reaction ("EX_gln_L_e_").

Determining a maximum bound for exogenous glutamine uptake:

In standard RPMI medium, glutamine exists in concentration: 0.3 g/L. For standard 24 hour growth before medium is refreshed, 21% parasitemia cultures will grow healthily, and are assumed to not deplete exogenous glutamine. Therefore, in 24 hours, at most we would expect 0.3 g/L/24h for a 21% parasitemia culture, resulting in: 1.25×10^{-2} g/L/h. Using $MW_{gln} = 146.14$ g/mol, we get 8.55×10^{-5} mol/l/h. Model fluxes are in "mmol/gDW/h", therefore we scale the gDW of a 21% parasitemia culture to per 1 gDW:

gDW of parasite in a 21% parasitemia culture: 1.11×10^{-2} (calculated above)

Scaling to 1 gDW we get: 7.698 mmol/l/gDW/h of glutamine

Cultures are grown in 50 mL flasks, and the current flux is per litre, so conversion to 50mL volume yields: 0.385 mmol/gDW/h for a 50 mL volume (as used for experimental data). Thus, in a 9% parasitemia culture, the maximum possible uptake rate of glutamine could be 0.385 mmol/gDW/h for a 50 ml volume culture, and was set as our upper bound for exogenous glutamine import. From literature findings, and confirmation from our own experimentation, that parasite arginase plays a large role in free arginine being converted to ornithine, the arginase reaction flux was clamped on 3.30×10^{-1} mmol / gDW / h, based on arginase activity observed by Olszewski *et al.* [77]. In addition, the ornithine aminotransferase reaction was constrained to zero to prevent ornithine recycling back into the system.

4.5 Preliminary predictions of haemoglobin supply and biomass demand

Prior to conducting model simulations of the production of parasite biomass and the concomitant catabolism of host haemoglobin, a thought experiment would be to theoretically predict how much biomass could be formed per haemoglobin molecule, based on the compositions of biomass and haemoglobin as present in the model. This would also provide a strong motivation for the necessity

of a model based on a complicated network for amino acid metabolism in *P. falciparum*.

In the model there is a reaction responsible for the digestion of haemoglobin "HMGLB", supplying freed amino acids to the parasite for biomass production. The biomass objective function (BOF), named in the model as "biomass", is a manually created reaction that harvests a large variety of components created by the metabolic network necessary to create a "unit" of biomass. As this BOF is an artificial reaction, we represent biomass production as the rate of production of these units. By comparing the compositions of haemoglobin molecules and these units of biomass produced by the model, we can make preliminary predictions on which amino acids will likely be most limiting by haemoglobin catabolism (assuming no interconversions of amino acids occur). Further, assuming no availability of exogenous amino acids (besides isoleucine), we can predict how many units of biomass can be produced by the biomass reaction for the total haemoglobin provided within a single red blood cell before specific amino acid supply would start to run out.

The number of haemoglobin molecules required to generate 10 000 units of biomass was calculated (calculations based on 10 000 units for aesthetic appeal), and on the compositions of haemoglobin and biomass (table 2). Tryptophan was calculated to be the shortest in supply, requiring 350 haemoglobin molecules to produce enough tryptophan for the production of 10 000 units of biomass. This is followed by asparagine, which requires 313 haemoglobin molecules.

Table_2: Comparing biomass composition versus haemoglobin composition to illustrate the basic supply and demand of amino acids in the modeled system. "AA" = amino acid, "AA/Hb" = amino acid per haemoglobin molecule, "AA/10 000 biomass" = number of each amino acid required to produce 10 000 units of biomass, "Hb/10 000 biomass" = haemoglobin molecules required to supply enough of a specified amino acid to produce 10 000 units of biomass.

AA	AA/Hb	AA/10 000 biomass	Hb/10 000 biomass
ala	72	1364	19
arg	12	1349	113
asn	20	6245	313
asp	30	2994	100
cys	6	674	113
gln	8	1353	170
glu	24	3505	147
gly	40	1827	46
his	28	1063	39
iso	0	3566	N/A
leu	72	3405	48
lys	44	5124	117
met	6	1019	170
phe	30	1823	61
pro	28	1085	39
ser	32	3055	96
thr	32	2045	64
trp	6	2095	350
tyr	12	1379	115
val	62	2030	33

Using the average amount of haemoglobin in an erythrocyte $= 4.2 \times 10^{-13}$ mmoles (calculated previously), and assuming 75% of haemoglobin is digested during a 48 hour life-cycle of an intra-erythrocytic *P. falciparum*, we can use Avogadro's constant to calculate the number of molecules of haemoglobin present within a red blood cell:

75% of haemoglobin consumed $= 3.15 \times 10^{-13}$ mmoles

Avogadro's constant $= 6.022 \times 10^{23}$ molecules in one mole of substance

We can estimate there are 1.9×10^{8} molecules of haemoglobin digested in a 48 hour life-cycle by an intra-erythrocytic parasite.

Taking the number of digested haemoglobin molecules into account, the maximum units of biomass that can be generated using the released amino acid molecules can be calculated (table 3).

Table_3: Predicting the amount of biomass units that can be formed by the biomass objective function solely using amino acids supplied via the catabolism of haemoglobin within a single erythrocyte. "AA" = amino acid, "AA/Hb" = amino acid per haemoglobin molecule, "AA/total erythrocyte Hb" = Number of molecules of each amino acid for the total haemoglobin found in an average erythrocyte, "Biomass/erythrocyte" = The total potential biomass units produced using the haemoglobin within a single erythrocyte.

AA	AA/Hb	AA/total erythrocyte Hb	Biomass/erythrocyte
ala	72	1.4×10^{10}	1.00×10^{11}
arg	12	2.3×10^{9}	1.69×10^{10}
asn	20	3.8×10^{9}	6.08×10^{9}
asp	30	5.7×10^{9}	1.90×10^{10}
cys	6	1.1×10^{9}	1.69×10^{10}
gln	8	1.5×10^{9}	1.12×10^{10}
glu	24	4.6×10^{9}	1.30×10^{10}
gly	40	7.6×10^{9}	4.16×10^{10}
his	28	5.3×10^{9}	5.00×10^{10}
iso	0	N/A	N/A
leu	72	1.4×10^{9}	4.02×10^{10}
lys	44	8.4×10^{9}	1.63×10^{10}
met	6	1.1×10^{9}	1.12×10^{10}
phe	30	5.7×10^{9}	3.13×10^{10}
pro	28	5.3×10^{9}	4.90×10^{10}
ser	32	6.1×10^{9}	1.99×10^{10}
thr	32	6.1×10^{9}	2.97×10^{10}
trp	6	1.1×10^{9}	5.44×10^{9}
tyr	12	2.3×10^{9}	1.65×10^{10}
val	62	1.2×10^{10}	5.80×10^{10}

From table 2 and table 3 we can predict that modeled simulations will most likely indicate (in order of demand): tryptophan, asparagine, methionine, glutamine, glutamate, lysine, arginine, and cysteine, as the amino acids with the highest dependency on exogenous supply to allow for optimal biomass production. Using our model based on the complex metabolic network of amino acid metabolism for intra-erythrocytic *P. falciparum* we hope to provide further insights.

4.6 Network reconstruction simulations

4.6.1 Overview of *P. falciparum* flux profile, as simulated by the network reconstruction

Once we had applied all our constraints to the network, our simulations started to predict a flux distribution that was more in line with literature expectations. The largest fluxes through the modeled system are presented in figure 15, displaying vast conversion of glucose to lactate, indicating a functional glycolysis system. There is also a build up of urea and ammonia that is successfully exported by the parasite. Amino acid metabolism comprises a large part of metabolism in *Plasmodia*. Many transamination reactions occur in the parasite. For example, in figure 15 it can also be seen how exogenous glutamine has a large import flux and glutamate is exported out the cell. Glutamine can be processed by the parasite to α-ketoglutarate, by a set of transamination reactions, leading to the production of ammonia and urea.

In culturing conditions, exogenous arginine is taken up by the parasite at a high rate and converted to ornithine (arginine + H2O $\rightarrow$ ornithine + urea), before exporting out of the system. We manually clamped the arginase reaction flux to the rate of ornithine accumulation we determined experimentally (figure 11) of $0.330 mmol/h/gDW$. From this added constraint we observed how arginine import and ornithine export increased accordingly. We had set the upper limit of our BOF to 0.0550 $h^{-1}$and simulations displayed that maximal biomass formation rate was achieved.

4.6.2 Model-predicted flux profile for an asynchronous *P. falciparum* culture compared to an experimental data set

To test the accuracy of our model simulations we compared our simulation data to a set of experimental data obtained via UPLC-UV analysis of time-point samples obtained from asynchronous *P. falciparum* cultures grown at 9% parasitemia.

Experimental fluxes were calculated for each of the amino acids from the UPLC-UV observed changes in amino acid concentrations over a 48 hour period and converted from millimolar to mmol/h units for a 50mL volume to match the

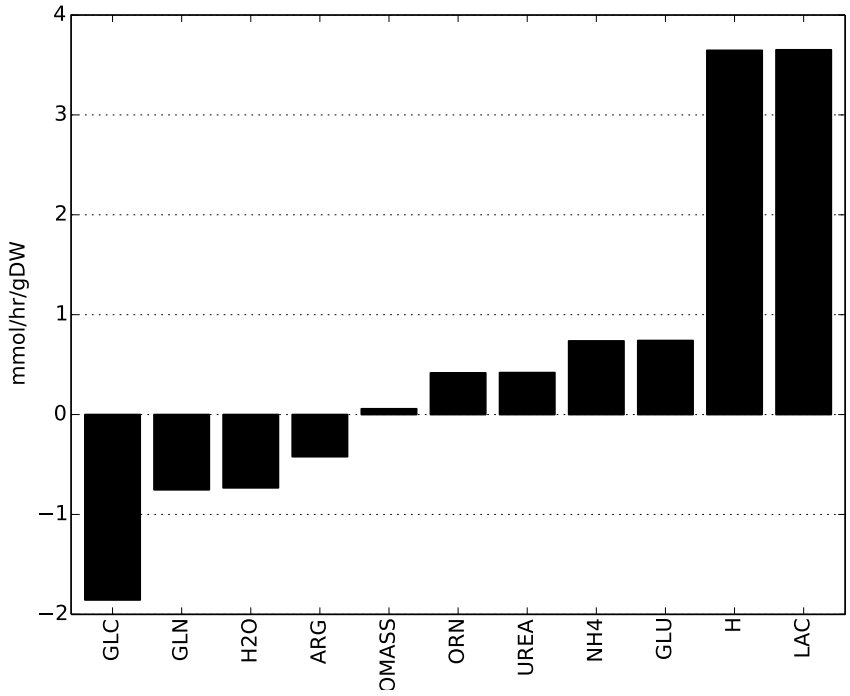


Fig.15 Overview of major fluxes predicted by the modeled system for *P. falciparum*. Negative values indicate a net import flux and positive values indicate a net export flux. OMASS = Biomass; ORN = ornithine; UREA = urea.

units used in the model. Additionally calculated fluxes were converted from parasite dry weight calculated for a 9% culture (4.76×10^{-3} gDW) to 1 gDW, as used in the model (mmol/h/gDW).

Qualitatively, fluxes predicted by the model compared well with our experimental data (figure 16). Both systems displayed a large favourability for exogenous glutamine as a primary amino acid source, with a resultant net export of glutamate and alanine, produced during glutaminolysis. The model indicated a slight over-utilization of glutamine as a result of us manually clamping the glutamine import flux at the maximum possible rate we could expect (0.385 mmoles/gDW/h). Upon closer inspection, in our experimental data we see that the net import flux of glutamine is substantially higher than the glutamate export flux. In the model, however, the glutamate export flux (0.374 mmoles/gDW/h) is almost as high as the import flux of glutamine (0.385 mmoles/gDW/h). Paired with this, in our simulations we see a smaller export

flux for alanine. In the process of glutaminolysis, glutamine is partially deaminated to glutamate. Thereafter three possible pathways are possible. Glutamate can either be further deaminated to produce α-ketoglutarate and ammonium, transfer the ammonium group to pyruvate to form alanine (transamination), or transfer the ammonium group to oxaloacetate to form aspartate (transamination). Our simulations indicate that perhaps glutamate is not being fully utilised in this glutaminolysis pathway of the model. Comparing model to experimental data, we see a lower glutamate utilisation and as a result a smaller export flux of alanine by the model. This could also be a possible reason for the unexpected import fluxes of exogenous aspartate and asparagine we see in our model.

Arginine is also largely converted to ornithine by arginase in experimental cultures as well as computational simulations. Ornithine export in the model was clamped to the flux we determined experimentally. Model simulations showed a 96% conversion of arginine to ornithine in the model. In experimental data that is slightly less and likely a result of an arginine deiminase converting some of the arginine into citrulline.

The amino acid flux distributions in figure 2 of Appendix C are preliminary simulations of amino acid flux distributions. Later in the development of the model, more accurate flux distributions can be observed once the correct constraints and boundary flux values were introduced (for biomass formation rate, haemoglobin digestion rate, and glucose consumption rate).

4.6.3 Investigation of model dependencies on various exogenous amino acids

Exchange reactions responsible for the import and export of metabolites between the parasite and cytosol of the host erythrocyte were systematically switched off and the effect thereof on parasite growth characteristics were examined. Removing exogenous isoleucine uptake prevented any biomass formation and energy demand by the parasite reflected this. When glutamine import was turned off only a slightly higher lactate production was observed, indicating that while glutamine is a desired source of amino acid, the model is able to make use of the various other amino acids in its absence, but causing a small increase in lactate production (Table 4). The increased transaminations occurring in the absence of glutamine increase the energetic demand of the parasite to produce biomass in the model. Analysis of upstream pathway fluxes revealed this increased lactate export flux to be as a result of glycerol being imported by the

cell to use as a carbon source. As a result of the additional reactions occurring in the absence of glutamine, the maximum glucose import rate was no longer sufficient to provide enough fuel for glycolysis.

The individual removal of exogenous arginine, methionine and tyrosine import resulted in lowered biomass production rates, indicating haemoglobin is not able to provide sufficient amounts of these amino acids, and that exogenous supply is necessary to sustain optimal growth by the parasite. The removal of exogenous proline and glutamate had no effect on parasite growth characteristics as the model suggests these amino acids are provided in sufficient quantity by catabolism of host haemoglobin (indicated by a net export of glutamate and proline in figure 16). When we simulated parasite growth on only haemoglobin and exogenous isoleucine, biomass production rates dropped to 52% of optimal rate. This is the same production rate observed as in the "no methionine" simulation, indicating that the absence of methionine alone is restricting growth rates to 52% regardless of availability of other amino acids. Adding exogenous methionine supply to simulated media containing only isoleucine led to a biomass production rate of 76%. However, when we supplied the other two exogenous amino acids indicated as necessary for optimal parasite growth, arginine and tyrosine, there was no further increase in biomass production rates. These model findings suggest that there is one or more amino acids, in addition to those currently mentioned in literature, required by *Plasmodium falciparum* to achieve maximal growth. Additionally, in comparison to the growth rates we measured for limited media cultures, we saw that restricting cultures to exogenous isoleucine and methionine only, we could achieve growth rates around 48% of optimal rate (table 1).

Table_4: Absence of exogenous amino acids and their effect on parasite growth (expressed as a percentage of standard growth simulations)

Model restriction	Hb catabolism	Glucose cons.	Lactate prod.	Biomass prod.
Standard growth	100.00	100.00	100.00	100.00
No isoleucine	100.00	0.03	0.42	0
No glutamine	100.00	100.00	113.97	100.00
No arginine	100.00	78.66	78.90	78.72
No glutamate	100.00	100.00	100.00	100.00
No methionine	100.00	52.05	52.33	52.08
No tyrosine	100.00	77.05	77.26	77.16
No proline	100.00	100.00	100.00	100.00
Only isoleucine	100.00	52.96	52.88	52.08
Only ile and met	100.00	78.45	79.18	76.12
Ile, met, tyr, arg	100.00	78.45	79.18	76.12

4.6.4 Model simulation of atovaquone activity

As atovaquone acts to restrict the activity of mitochondrial enzyme, dihydroorotate dehydrogenase, the activity of atovaquone can be simulated by applying constraints to the reaction in the model representing dihydroorotate dehydrogenase, named 'DHORD2_mt'. 'DHORD2_mt' has a normal flux of 0.0128 mmol/gDW/h when simulating standard growth conditions. Figure 18 illustrates the uniform decrease in biomass production as we simulate the effect of increasing atovaquone concentration. Model predictions were compared to a dose-response curve determined by Srivastava *et al.* [105] (figure 17) and observe a similar trend in activity. Srivastava *et al.* [105] showed that increasing concentrations of atovaquone would cause an increasing inhibition of respiration in *P. falciparum*. The rate of respiration in these parasites can be directly correlated to their rate of biomass production. By simulating the action of atovaquone to inhibit the activity of DHODH in the model, a linear correlation between the percentage inhibition of DHODH and the percent of optimal rate of biomass formation displays how the degree of functionality of the DHODH enzyme directly affects the capacity of *P. falciparum* to generate biomass.

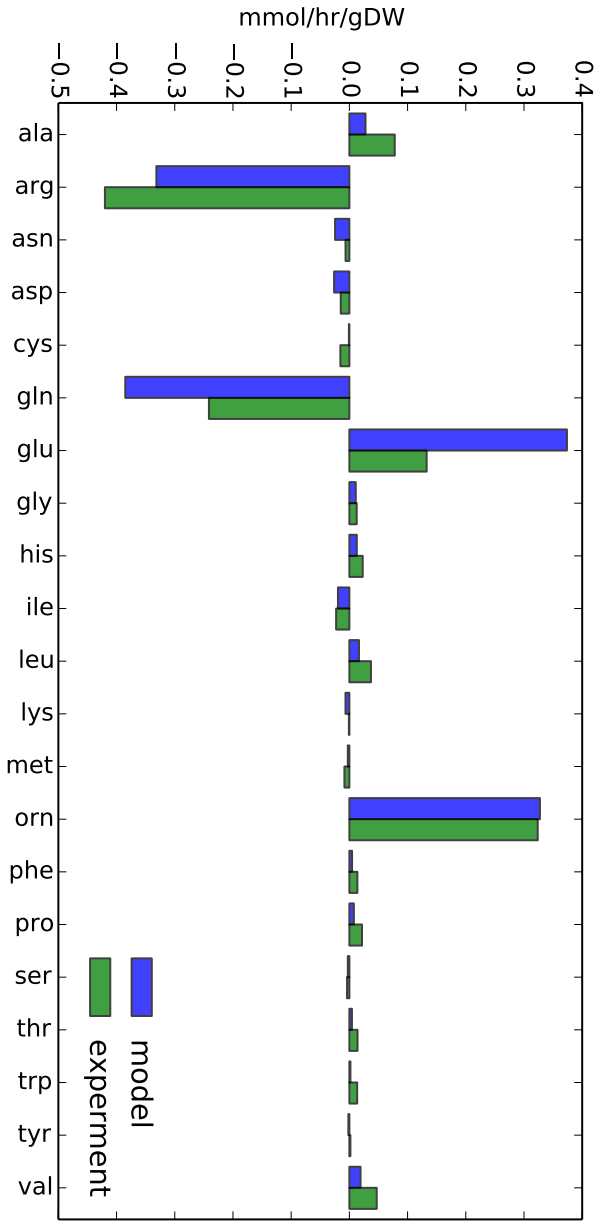


Fig.16 Model-predicted amino acid exchange fluxes obtained in comparison with fluxes calculated from data obtained from UPLC-UV analysed parasite culture growth. Fluxes for both sets of data are calculated based on a 9% parasitemia culture in a 50ml culture of standard RPMI at 4% haematocrit.

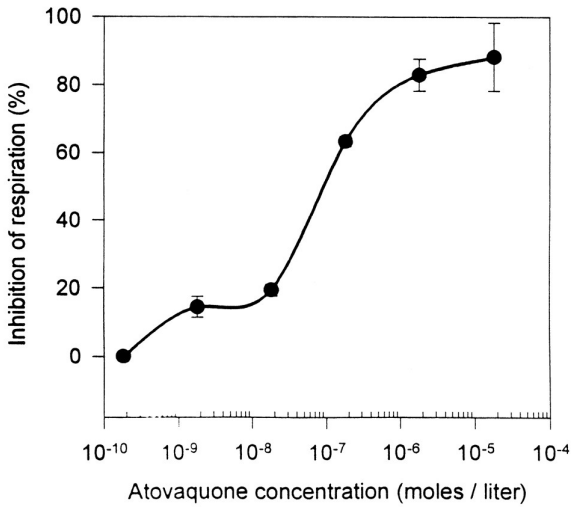


Fig.17 Dose-response curve of atovaquone effect on *Plasmodium falciparum* respiration, determined by Srivastava *et al.* [105]

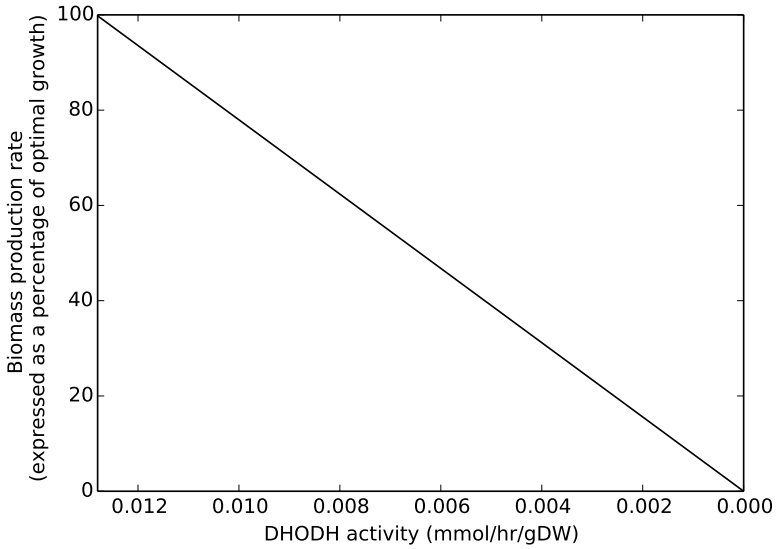


Fig.18 Model simulations show that constraining DHODH activity reduces the ability of *Plasmodium falciparum* to generate biomass

Chapter 5

Discussion

Network reconstructions are an ideal approach to capture large, intricate biochemical networks within a computational model. In this study we investigated the dependence of *P. falciparum* on exogenous amino acid supply in relation to haemoglobin-derived amino acids. *Plasmodia* have a very limited capacity for *de novo* amino acid biosynthesis, by aerobic CO_2 fixation and glucose dissimilation. This limited capacity for synthesis has been shown to not be primarily directed towards amino acid formation for *Plasmodial* protein synthesis, but rather for the replenishment of TCA-cycle intermediates [99]. *Plasmodia*, instead depend on the catabolism of host haemoglobin housed within erythrocytes, and free extracellular amino acid in the host's blood stream. Along with this, *Plasmodia* have developed a complex transamination system involved in a variety of metabolic processes [77, 12, 118]. Genome-scale network reconstruction-based modeling would therefore be a suitable approach for modeling amino acid metabolism of *P. falciparum*. Using a network reconstruction created by Plata *et al.* [89], we were able to create a functional model of *P. falciparum* metabolism. This model was able to reproduce the mass conversion of glucose to lactate (as observed experimentally), rapidly catabolise host haemoglobin and scavenge the freed amino acids, as well as those available extracellularly. When exposed to unfavourable conditions the model was able to adapt and find alternative solutions to try achieve optimal biomass production rates. Limiting the availability of extracellular amino acids caused the model to make use of transamination reactions to produce the missing amino acids required for biomass production. To validate the findings of the model, amino acid analysis was conducted on *P. falciparum* cultures grown for a full 48 hour life-cycle and fluxes obtained from this data were compared to the outputs of model simulations (figure 16). We also

displayed the ability of the model to mimic the effects of an active compound (atovaquone) in a popular antimalarial drug on the growth of *P. falciparum.*

Divo *et al.* [16] was able to achieve almost 90% of standard growth in a limited medium containing only 7 amino acids (isoleucine, methionine, cysteine, glutamine, glutamate, tyrosine, proline). Liu *et al.* [59] observed 86% of standard growth, showing at least a 1.8-fold increase in parasitemia over a single life cycle, using a growth medium supplemented with isoleucine as the sole-added amino acid. Based on these findings, we attempted to reproduce this high level of growth in amino-acid-limited media. Our model simulations were able to verify that we could achieve close to 80% of optimal biomass formation rates (table 4). Experimentally we were not as successful. When only adding isoleucine to an amino-acid-free custom RPMI culture medium, parasites were not able to grow well and we saw a net decrease in parasitemia over time. When adding both isoleucine and methionine, growth improved to about 50% of standard culturing conditions. The same was observed when adding the "core" 5 amino acids (ile; met; gln; glu; cys). This slower growth was largely due to sluggish progression of parasites through their life-cycle. In standard culturing, merozoites will transition to mature schizonts within 48 hours. However, in limited media cultures, parasites would take more than 72 hours to complete a life-cycle. Microscopically, it was observed that a large portion of parasites in limited media were undergoing metabolic stress, where intra-erythrocytic parasites would appear as solid, compact masses within a red blood cell (figure 5). Liu *et al.* [58] observed this same slower life-cycle progression in limited media. It is suspected that re-invasion efficiency and development of newly invading merozoites are significantly hampered due to a lack of specific amino acids that might be vital for parasite development. It is possible that *P. falciparum* we used in limited media experiments might have adapted over time to favour exogenous amino acids for growth over those made available during haemoglobin degradation. In future experimentation multiple strains should be considered for limited media culturing. It would also be recommended to reproduce the results obtained by Divo *et al.* [16] with his 7 amino acid limited medium.

In our theoretical calculations for specific growth rate, based on data presented by du Toit *et al.* [17], we calculated a maximal specific growth rate for P. falciparum of 0.0550 h^{-1} to be used in our model. However, in our standard cultures we observe less than 50% of that rate (table 1). This is primarily as a result of the specific growth rate used in the model to only represent the rate of growth for a single life cycle. In a system of multi-generational growth, such as continuous *in vitro* culturing there are a number of additional factors affecting the specific growth rate. Our theoretical specific growth does not account for

factors such as the time from red blood cell rupture and the accompanying release of daughter merozoites to the resultant re-invasion. Daughter merozoites also do not invade with a 100% re-invasion-to-schizogeny efficiency and as a result a portion of released new progeny will not reach schizogeny and release daughter merozoites of their own. Perhaps even more importantly, the volume increase from a merozoite to a schizont is by no means a perfect correlation to the increase in biomass, and the volume occupied by such a schizont divided by the volume of a merozoite does not accurately represent the number of new merozoites that will re-invade host erythrocytes. If this were the case, each schizont would result in ˜31 newly infected red blood cells, but in literature there is a reported average of 16-18 merozoites released per schizont [28], almost 50% less than the 31 merozoites we might otherwise presume just on the volume increase in a given life cycle.

There were some interesting findings from data obtained from amino acid analysis of 48 hour time-course samples of *P. falciparum* cultures. Most notably was the rapid depletion of arginine from the culture medium, accompanied by a large accumulation of ornithine (figure 11). In a 9% parasitemia culture it is predicted that arginine would be depleted after 48 hours of growth in standard RPMI 1640 medium. This is not a significant issue in standard practice, as culture medium is refreshed every 24 hours. However, in 21% parasitemia cultures, arginine is predicted to be depleted after about 30 hours of growth in the same medium (figure 8). Previous studies have concluded that arginine is not required for optimal growth, suggesting that sufficient quantities are obtained from haemoglobin digestion [59]. Olszewski *et al. [77]* also showed that the mass-conversion of arginine to ornithine is not for the use of ornithine as a substrate either, as arginase knockout strains were able to obtain enough ornithine from the plasma or activity of the host arginase. Instead they hypothesize that the parasite depletes the host arginine pool to prevent nitric oxide generation from arginine, suppressing the host immune response. Additionally the large amount of ornithine produced may compete with arginine for transporter-mediated uptake into macrophages. Also observed is a notable consumption of glutamine and resultant build up of glutamate and alanine, which can be attributed to the switch in parasite metabolism during the trophozoite-stage to glutamine metabolism as a source of α-ketoglutarate for use in the TCA-cycle. A small consumption of isoleucine and methionine were also noted.

Some initial calculations were conducted to predict which of the amino acids supplied by haemoglobin catabolism would be in short supply for biomass production in the model. From these calculations it was determined that tryptophan, asparagine, methionine, glutamine and glutamate were the top five amino

acids suspected to require exogenous supply (along with isoleucine, which does not occur in haemoglobin) for the optimal production of biomass by the model. Once we had produced an optimized model to simulate parasite metabolism, we could validate our calculated predictions on amino acid supply and demand. As expected we saw a demand on exogenous sources of glutamine, methionine and asparagine. However, the amino acid predicted to have the highest need for an exogenous supply, tryptophan, showed no need for import of extracellular supply. Investigation into the reaction network of the model showed that tryptophan was exclusively used as a building block in biomass formation. This finding is significant in showing that for the most part haemoglobin should be able to provide enough building blocks for biomass formation. The demand of other metabolic processes have a large impact on amino acid demand. Glutaminolysis is an example of this, whereby glutamine demand is increased as a result of glutamine additionally being converted to α-ketoglutarate. Simulations also indicated there was no demand for glutamate import. In fact, the model showed a net export of glutamate, further re-enforcing the belief that reactions other than the biomass producing reaction play a large role in amino acid metabolism.

The primary objective of our research was to investigate amino acid metabolism in intra-erythrocytic *Plasmodium falciparum*. To address this objective, we made use of computational modeling to simulate the metabolism of the parasite, and then validate the results of our model simulations with experimental data. In the first chapter we stated four key research questions we would attempt to answer:

1. Is the model able to reproduce experimental observations in terms of amino acid incorporation into biomass and amino acid export, using haemoglobin catabolism as a primary source, and extracellular amino acid scavenging to make up for any deficits?

2. Can the findings of Liu *et al.* [59] that *Plasmodium falciparum* is able to grow in culture medium containing isoleucine as the only supplemented amino acid be reproduced?

3. Can a steady-state flux distribution for the model system that is realistic to what we can observe in nature be identified?

4. From the flux distribution outputs, can any key reactions or metabolites that could be investigated further as potential drug targets be identified?

Using a genome-scale network reconstruction of *P. falciparum* and performing flux-balance analysis we were able to reproduce experimentally observed fluxes with good accuracy and display some of the key dependencies of the parasite on exogenous amino acid supply for optimal growth.The model successfully used haemoglobin catabolism as a primary source of amino acids for biomass formation and obtained any deficits from extracellular scavenging. Any excess amino acids were successfully exported from the system. Further, we were able to simulate the findings of Liu *et al.* that *P. falciparum* is able to grow on nutrient-restricted media at almost 80% of maximal biomass formation rates [59]. We were, however, not able to identify any novel metabolic targets for drug development.

Overall the model has provided some interesting insights into the amino acid metabolism of Plasmodium falciparum. The model has indicated that certain amino acids such as glutamine, are not vital for the generation of biomass, however, are more energetically favourable options for growth. Experimental studies [16] have also indicated a core 7 amino acids (isoleucine, methionine, cysteine, glutamine, glutamate, tyrosine, proline) required for the optimal rates of biomass formation, where the model reaffirms this finding in the case of isoleucine, methionine, tyrosine and arginine. However our model indicates that the absence of exogenous glutamine, glutamate and proline to not hamper optimal biomass formation. Adding to this, with the four amino acids indicated to be vital to optimal biomass formation only a 76% of optimal rate was simulated by the model. This perhaps indicates the necessity for another, unsuspected amino acid, or perhaps even combination of amino acids to actually reach optimal biomass formation rate (table 4).

Future improvements to the model will allow for more biologically accurate simulations, and a wider range of applications. Attaining a more accurate molecular representation of *Plasmodium* biomass for the creation of biomass objective functions will create a more realistic picture of metabolic requirements for parasite growth such as the discrepancies we observed when comparing model versus experimental amino acid exchange fluxes in figure 16. The addition of more biologically relevant constraints to reaction fluxes in the model, such as transamination reactions will vastly aid in the understanding of these complex reaction dynamics. Only partial success was achieved with minimal media culturing. Testing multiple parasite strains and media conditions are considerations that should be made when expanding on this work.

Bibliography

[1] Masamichi Aikawa, Louis H Miller, James Johnson, and John Rabbege. Erythrocyte entry by malarial parasites. a moving junction between erythrocyte and parasite. *The Journal of Cell Biology*, 77(1):72–82, 1978.

[2] E_ Almaas, B Kovacs, T Vicsek, ZN Oltvai, and A-L Barabási. Global organization of metabolic fluxes in the bacterium escherichia coli. *Nature*, 427(6977):839–843, 2004.

[3] Hani Atamna and Hagai Ginsburg. Origin of reactive oxygen species in erythrocytes infected with *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*, 61(2):231–241, 1993.

[4] Hani Atamna, Gianpiero Pascarmona, and Hagai Ginsburg. Hexose-monophosphate shunt activity in intact plasmodium falciparum-infected erythrocytes and in free parasites. *Molecular and Biochemical Parasitology*, 67(1):79–89, 1994.

[5] Eric G Ball, Ralph W McKee, Christian B Anfinsen, Walter O Cruz, Quentin M Geiman, et al. Studies on malarial parasites. ix. chemical and metabolic changes during growth and multiplication in vivo and in vitro. *Journal of Biological Chemistry*, 175(2):547–71, 1948.

[6] Katja Becker, Leann Tilley, Jonathan L Vennerstrom, David Roberts, Stephen Rogerson, and Hagai Ginsburg. Oxidative stress in malaria parasite-infected erythrocytes: host–parasite interactions. *International Journal for Parasitology*, 34(2):163–189, 2004.

[7] Scott A Becker, Adam M Feist, Monica L Mo, Gregory Hannum, Bernhard Ø Palsson, and Markus J Herrgard. Quantitative prediction of cellular metabolism with constraint-based models: the *COBRA* toolbox. *Nature Protocols*, 2(3):727–738, 2007.

[8] Joost Boele, Brett G Olivier, and Bas Teusink. Fame, the flux analysis and modeling environment. *BMC Systems Biology*, 6(1):8, 2012.

[9] Zbynek Bozdech, Manuel Llinás, Brian Lee Pulliam, Edith D Wong, Jingchun Zhu, and Joseph L DeRisi. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol*, 1(1):e5, 2003.

[10] Teresa Gil Carvalho and Robert Ménard. *Manipulating the Plasmodium genome*. Caister Academic Press: Wymondham, UK, 2004.

[11] Arvind K Chavali, Jeffrey D Whittemore, James A Eddy, Kyle T Williams, and Jason A Papin. Supplementary material i.

[12] Simon A Cobbold, Ashley M Vaughan, Ian A Lewis, Heather J Painter, Nelly Camargo, David H Perlman, Matthew Fishbaugher, Julie Healer, Alan F Cowman, Stefan HI Kappe, et al. Kinetic flux profiling elucidates two independent acetyl-coa biosynthetic pathways in plasmodium falciparum. *Journal of Biological Chemistry*, 288(51):36338–36350, 2013.

[13] Markus W Covert, Eric M Knight, Jennifer L Reed, Markus J Herrgard, and Bernhard O Palsson. Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, 429(6987):92–96, 2004.

[14] Alan F Cowman and Brendan S Crabb. Invasion of red blood cells by malaria parasites. *Cell*, 124(4):755–766, 2006.

[15] Peter H David, Marcel Hommel, Louis H Miller, Iroka J Udeinya, and Lynette D Oligino. Parasite sequestration in *Plasmodium falciparum* malaria: spleen and antibody modulation of cytoadherence of infected erythrocytes. *Proceedings of the National Academy of Sciences*, 80(16):5075–5079, 1983.

[16] Alan A Divo, Timothy G Geary, Nella L Davis, and James B Jensen. Nutritional requirements of *Plasmodium falciparum* in culture. i. exogenously supplied dialyzable components necessary for continuous growth. *Journal of Protozoology*, 32(1):59–64, 1985.

[17] Francois du Toit. *Modeling glycolysis in Plasmodium-infected erythrocytes*. PhD thesis, University of Stellenbosch, 2015.

[18] Natalie C Duarte, Scott A Becker, Neema Jamshidi, Ines Thiele, Monica L Mo, Thuy D Vo, Rohith Srivas, and Bernhard Ø Palsson. Global reconstruction of the human metabolic network based on genomic and bibliomic

data. *Proceedings of the National Academy of Sciences*, 104(6):1777–1782, 2007.

[19] Natalie C Duarte, Markus J Herrgård, and Bernhard Ø Palsson. Reconstruction and validation of *Saccharomyces cerevisiae* ind750, a fully compartmentalized genome-scale metabolic model. *Genome research*, 14(7):1298–1309, 2004.

[20] Jeremy S Edwards and Bernhard O Palsson. Systems properties of the haemophilus influenzaerd metabolic genotype. *Journal of Biological Chemistry*, 274(25):17410–17416, 1999.

[21] Adam M Feist, Christopher S Henry, Jennifer L Reed, Markus Krummenacker, Andrew R Joyce, Peter D Karp, Linda J Broadbelt, Vassily Hatzimanikatis, and Bernhard Ø Palsson. A genome-scale metabolic reconstruction for *Escherichia coli* k-12 mg1655 that accounts for 1260 orfs and thermodynamic information. *Molecular Systems Biology*, 3(1):121, 2007.

[22] Adam M Feist, Markus J Herrgård, Ines Thiele, Jennie L Reed, and Bernhard Ø Palsson. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, 7(2):129–143, 2009.

[23] Adam M Feist and Bernhard Ø Palsson. The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nature Biotechnology*, 26(6):659–667, 2008.

[24] Laurence Florens, Michael P Washburn, J Dale Raine, Robert M Anthony, Munira Grainger, J David Haynes, J Kathleen Moch, Nemone Muster, John B Sacci, David L Tabb, et al. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature*, 419(6906):520–526, 2002.

[25] Jochen Förster, Iman Famili, Patrick Fu, Bernhard Ø Palsson, and Jens Nielsen. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Research*, 13(2):244–253, 2003.

[26] Susan E Francis, Ilya Y Gluzman, Anna Oksman, Dolly Banerjee, and Daniel E Goldberg. Characterization of native falcipain, an enzyme involved in *Plasmodium falciparum* hemoglobin degradation. *Molecular and Biochemical Parasitology*, 83(2):189–200, 1996.

[27] Susan E Francis, IY Gluzman, A Oksman, A Knickerbocker, R Mueller, ML Bryant, DR Sherman, DG Russell, and DE Goldberg. Molecular characterization and inhibition of a *Plasmodium falciparum* aspartic hemoglobinase. *The EMBO Journal*, 13(2):306, 1994.

[28] Lynne S. Garcia. *Practical Guide to Diagnostic Parasitology*. American Society for Microbiology Press, 2009.

[29] Malcolm J Gardner, Neil Hall, Eula Fung, Owen White, Matthew Berriman, Richard W Hyman, Jane M Carlton, Arnab Pain, Karen E Nelson, Sharen Bowman, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, 419(6906):498–511, 2002.

[30] PCC Garnham, RG Bird, JR Baker, and R Killick-Kendrick. Electron microscope studies on the motile stages of malaria parasites vii. the fine structure of the merozoites of exoerythrocytic schizonts of *Plasmodium berghei yoelii*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 63(3):328IN3331–330IN9332, 1969.

[31] PCC Garnham, RS Bray, W Cooper, R Lainson, FI Awad, and J Williamson. The pre-erythrocytic stage of *Plasmodium ovals*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 49(2):158IN1165–164IN2167, 1955.

[32] H Ginsburg. Some reflections concerning host erythrocyte-malarial parasite interrelationships. *Blood cells*, 16(2-3):225–235, 1989.

[33] Hagai Ginsburg, Shirley Kutner, Miriam Krugliak, and Z Ioav Cabantchik. Characterization of permeation pathways appearing in the host membrane of *Plasmodium falciparum* infected red blood cells. *Molecular and Biochemical Parasitology*, 14(3):313–322, 1985.

[34] IY Gluzman, Susan E Francis, Anna Oksman, Christine E Smith, KL Duffin, and Daniel E Goldberg. Order and specificity of the *Plasmodium falciparum* hemoglobin degradation pathway. *Journal of Clinical Investigation*, 93(4):1602, 1994.

[35] Kasturi Haldar and Narla Mohandas. Malaria, erythrocytic infection, and anemia. *ASH Education Program Book*, 2009(1):87–93, 2009.

[36] Shortt HE. The pre-erythrocytic stage of *Plasmodium falciparum*; a preliminary note. *British Medical Journal*, 2(4635):1006–8, 1949.

[37] Markus J Herrgård, Baek-Seok Lee, Vasiliy Portnoy, and Bernhard Ø Palsson. Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*. *Genome Research*, 16(5):627–635, 2006.

[38] LL Hsiao, RJ Howard, M Aikawa, and TF Taraschi. Modification of host cell membrane lipid composition by the intra-erythrocytic human malaria

parasite *Plasmodium falciparum*. *Biochemical Journal*, 274(1):121–132, 1991.

[39] Rafael U Ibarra, Jeremy S Edwards, and Bernhard O Palsson. Escherichia coli k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, 420(6912):186–189, 2002.

[40] Isra Ittarat, Wanida Asawamahasakda, and Steven R Meshnick. The effects of antimalarials on the *Plasmodium falciparum* dihydroorotate dehydrogenase. *Experimental Parasitology*, 79(1):50–56, 1994.

[41] David L Vancer Jagt, Lucy A Hunsaker, and Noami M Campos. Characterization of a hemoglobin-degrading, low molecular weight protease from *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*, 18(3):389–400, 1986.

[42] Mona D Jensen, Margaret Conley, and Lucy Dale Helstowski. Culture of *Plasmodium falciparum*: the role of ph, glucose, and lactate. *The Journal of Parasitology*, pages 1060–1067, 1983.

[43] Oscar K Kai and David J Roberts. The pathophysiology of malarial anaemia: where have all the red cells gone? *BMC Medicine*, 6(1):24, 2008.

[44] Minoru Kanehisa and Susumu Goto. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27–30, 2000.

[45] Peter D Karp, Ingrid M Keseler, Alexander Shearer, Mario Latendresse, Markus Krummenacker, Suzanne M Paley, Ian Paulsen, Julio Collado-Vides, Socorro Gama-Castro, Martin Peralta-Gil, et al. Multidimensional annotation of the escherichia coli k-12 genome. *Nucleic Acids Research*, 35(22):7577–7590, 2007.

[46] Hangjun Ke, Ian A Lewis, Joanne M Morrisey, Kyle J McLean, Suresh M Ganesan, Heather J Painter, Michael W Mather, Marcelo Jacobs-Lorena, Manuel Llinás, and Akhil B Vaidya. Genetic investigation of tricarboxylic acid metabolism during the *Plasmodium falciparum* life cycle. *Cell Reports*, 11(1):164–174, 2015.

[47] Hyun Uk Kim, Soo Young Kim, Haeyoung Jeong, Tae Yong Kim, Jae Jong Kim, Hyon E Choy, Kyu Yang Yi, Joon Haeng Rhee, and Sang Yup Lee. Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. *Molecular Systems Biology*, 7(1):460, 2011.

[48] Kiaran Kirk. Membrane transport in the malaria-infected erythrocyte. *Physiological Reviews*, 81(2):495–537, 2001.

[49] Cynthia J Krieger, Peifen Zhang, Lukas A Mueller, Alfred Wang, Suzanne Paley, Martha Arnaud, John Pick, Seung Y Rhee, and Peter D Karp. Metacyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Research*, 32(suppl 1):D438–D442, 2004.

[50] S Krishna, DW Waller, F Ter Kuile, D Kwiatkowski, J Crawley, CFC Craddock, F Nosten, D Chapman, D Brewster, PA Holloway, et al. Lactic acidosis and hypoglycaemia in children with severe malaria: pathophysiological and prognostic significance. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 88(1):67–73, 1994.

[51] Miriam Krugliak, Jianmin Zhang, and Hagai Ginsburg. Intraerythrocytic *Plasmodium falciparum* utilizes only a fraction of the amino acids derived from the digestion of host cell cytosol for the biosynthesis of its proteins. *Molecular and Biochemical Parasitology*, 119(2):249–256, 2002.

[52] Chris Lambros and Jerome P Vanderberg. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *The Journal of Parasitology*, pages 418–420, 1979.

[53] Abigail A Lamikanra, Douglas Brown, Alexandre Potocnik, Climent Casals-Pascual, Jean Langhorne, and David J Roberts. Malarial anemia: of mice and men. *Blood*, 110(1):18–28, 2007.

[54] Edwin Lasonder, Chris J Janse, Geert-Jan van Gemert, Gunnar R Mair, Adriaan MW Vermunt, Bruno G Douradinha, Vera van Noort, Martijn A Huynen, Adrian JF Luty, Hans Kroeze, et al. Proteomic profiling of *Plasmodium* sporozoite maturation identifies new proteins essential for parasite development and infectivity. *PLoS Pathogens*, 4(10):e1000195, 2008.

[55] Jong Min Lee, Erwin P Gianchandani, James A Eddy, and Jason A Papin. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Computational Biology*, 4(5):e1000086, 2008.

[56] Joshua A Lerman, Daniel R Hyduke, Haythem Latif, Vasiliy A Portnoy, Nathan E Lewis, Jeffrey D Orth, Alexandra C Schrimpe-Rutledge, Richard D Smith, Joshua N Adkins, Karsten Zengler, et al. In silico method for modelling metabolism and gene product expression at genome scale. *Nature Communications*, 3:929, 2012.

[57] Rongshi Li, George L Kenyon, Fred E Cohen, Xiaowu Chen, Baoqing Gong, Jose N Dominguez, Eugene Davidson, Gary Kurzban, Robert E

Miller, Edwin O Nuzum, et al. In vitro antimalarial activity of chalcones and their derivatives. *Journal of Medicinal Chemistry*, 38(26):5031–5037, 1995.

[58] Jun Liu, Ilya Y Gluzman, Mark E Drew, and Daniel E Goldberg. The role of *Plasmodium falciparum* food vacuole plasmepsins. *Journal of Biological Chemistry*, 280(2):1432–1437, 2005.

[59] Jun Liu, Eva S Istvan, Ilya Y Gluzman, Julia Gross, and Daniel E Goldberg. *Plasmodium falciparum* ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. *Proceedings of the National Academy of Sciences*, 103(23):8840–8845, 2006.

[60] Shengfa Liu, Jianbing Mu, Hongying Jiang, and Xin-zhuan Su. Effects of *Plasmodium falciparum* mixed infections on in vitro antimalarial drug tests and genotyping. *The American Journal of Tropical Medicine and Hygiene*, 79(2):178–184, 2008.

[61] Manuel Llinas, Zbynek Bozdech, Edith D Wong, Alex T Adai, and Joseph L DeRisi. Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Research*, 34(4):1166–1173, 2006.

[62] Lior Lobel, Nadejda Sigal, Ilya Borovok, Eytan Ruppin, and Anat A Herskovits. Integrative genomic analysis identifies isoleucine and cody as regulators of *Listeria monocytogenes* virulence. *PLoS Genetics*, 8(9):e1002887, 2012.

[63] Paul Loria, Susanne Miller, Michael Foley, and Leann Tilley. Inhibition of the peroxidative degradation of haem as the basis of action of chloroquine and other quinoline antimalarials. *Biochemical Journal*, 339(2):363–370, 1999.

[64] RA Majewski and MM Domach. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnology and Bioengineering*, 35(7):732–738, 1990.

[65] Michael W Mather, Elisabeth Darrouzet, Maria Valkova-Valchanova, Jason W Cooley, Michael T McIntosh, Fevzi Daldal, and Akhil B Vaidya. Uncovering the molecular mode of action of the antimalarial drug atovaquone using a bacterial system. *Journal of Biological Chemistry*, 280(29):27458–27465, 2005.

[66] Michael W Mather, Karl W Henry, and Akhil B Vaidya. Mitochondrial drug targets in apicomplexan parasites. *Current Drug Targets*, 8(1):49–60, 2007.

[67] Jakob MA Mauritz, Alessandro Esposito, Hagai Ginsburg, Clemens F Kaminski, Teresa Tiffert, and Virgilio L Lew. The homeostasis of *Plasmodium falciparum*-infected red blood cells. *PLoS Computational Biology*, 5(4):e1000339, 2009.

[68] Louis H Miller et al. Distribution of mature trophozoites and schizonts of plasmodium falciparum in the organs of aotus trivirgatus, the night monkey. *American Journal of Tropical Medicine and Hygiene*, 18(6, Pt. 1):860–65, 1969.

[69] K Mohan, ML Dubey, NK Ganguly, and RC Mahajan. *Plasmodium falciparum*: role of activated blood monocytes in erythrocyte membrane damage and red cell loss during malaria. *Experimental Parasitology*, 80(1):54–63, 1995.

[70] Dempsie B Morrison, Harold A Jeskey, et al. Alterations in some constituents of the monkey erythrocyte infected with *Plasmodium knowlesi* as related to pigment formation. *Journal of the National Malaria Society*, 7(4):259–64, 1948.

[71] Kenji Nakahigashi, Yoshihiro Toya, Nobuyoshi Ishii, Tomoyoshi Soga, Miki Hasegawa, Hisami Watanabe, Yuki Takai, Masayuki Honma, Hirotada Mori, and Masaru Tomita. Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Molecular Systems Biology*, 5(1):306, 2009.

[72] Chris Newbold, Alister Craig, Sue Kyes, Alex Rowe, Delmiro Fernandez-Reyes, and Toby Fagan. Cytoadherence, pathogenesis and the infected red cell surface in *Plasmodium falciparum*. *International Journal for Parasitology*, 29(6):927–937, 1999.

[73] N.I.H. Rbc indices, accessed 15/07/2017. *Available: https://medlineplus.gov/ency/article/003648.htm*, 2016.

[74] Juan Nogales, Steinn Gudmundsson, Eric M Knight, Bernhard O Palsson, and Ines Thiele. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences*, 109(7):2678–2683, 2012.

[75] Matthew A Oberhardt, Bernhard Ø Palsson, and Jason A Papin. Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology*, 5(1):320, 2009.

[76] Kellen L Olszewski and Manuel Llinás. Central carbon metabolism of *Plasmodium* parasites. *Molecular and Biochemical Parasitology*, 175(2):95–103, 2011.

[77] Kellen L Olszewski, Joanne M Morrisey, Daniel Wilinski, James M Burns, Akhil B Vaidya, Joshua D Rabinowitz, and Manuel Llinás. Host-parasite interactions revealed by *Plasmodium falciparum* metabolomics. *Cell Host & Microbe*, 5(2):191–199, 2009.

[78] World Health Organization et al. World malaria report 2016. 2016. *Available: http://www.who.int/malaria/publications/world-malaria-report-2016/report/en/. Accessed 2017/02/10*, 2016.

[79] Augustine U Orjih and Coy D Fitch. Hemozoin production by *Plasmodium falciparum*: variation with strain and exposure to chloroquine. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1157(2):270–274, 1993.

[80] Jeffrey D Orth, Ines Thiele, and Bernhard Ø Palsson. What is flux balance analysis? *Nature Biotechnology*, 28(3):245–248, 2010.

[81] Heather J Painter, Joanne M Morrisey, Michael W Mather, and Akhil B Vaidya. Specific role of mitochondrial electron transport in blood-stage *Plasmodium falciparum*. *Nature*, 446(7131):88–91, 2007.

[82] Csaba Pál, Balázs Papp, and Martin J Lercher. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nature Genetics*, 37(12):1372–1375, 2005.

[83] Ana Pamplona, Ana Ferreira, József Balla, Viktória Jeney, György Balla, Sabrina Epiphanio, Ângelo Chora, Cristina D Rodrigues, Isabel Pombo Gregoire, Margarida Cunha-Rodrigues, et al. Heme oxygenase-1 and carbon monoxide suppress the pathogenesis of experimental cerebral malaria. *Nature Medicine*, 13(6):703–710, 2007.

[84] Eleftherios Terry Papoutsakis. Equations and calculations for fermentations of butyric acid bacteria. *Biotechnology and Bioengineering*, 26(2):174–187, 1984.

[85] Eleftherios Terry Papoutsakis and Charles L Meyer. Fermentation equations for propionic-acid bacteria and production of assorted oxychemicals from various sugars. *Biotechnology and Bioengineering*, 27(1):67–80, 1985.

[86] Balazs Papp, Csaba Pal, and Laurence D Hurst. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature*, 429(6992):661–664, 2004.

[87] Gerald Penkler, Francois du Toit, Waldo Adams, Marina Rautenbach, Daniel C Palm, David D van Niekerk, and Jacky L Snoep. Construction and validation of a detailed kinetic model of glycolysis in *Plasmodium falciparum*. *FEBS Journal*, 282(8):1481–1511, 2015.

[88] Kevin G Phillips, Steven L Jacques, and Owen JT McCarty. Measurement of single cell refractive index, dry mass, volume, and density using a transillumination microscope. *Physical Review Letters*, 109(11):118105, 2012.

[89] Germán Plata, Tzu-Lin Hsiao, Kellen L Olszewski, Manuel Llinás, and Dennis Vitkup. Reconstruction and flux-balance analysis of the *Plasmodium falciparum* metabolic network. *Molecular Systems Biology*, 6(1):408, 2010.

[90] Jennifer L Reed, Thuy D Vo, Christophe H Schilling, and Bernhard O Palsson. An expanded genome-scale model of *Escherichia coli* k-12 (i jr904 gsm/gpr). *Genome Biology*, 4(9):R54, 2003.

[91] Qinghu Ren, Kaixi Chen, and Ian T Paulsen. Transportdb: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Research*, 35(suppl 1):D274–D279, 2007.

[92] Philip J Rosenthal. *Plasmodium falciparum*: effects of proteinase inhibitors on globin hydrolysis by cultured malaria parasites. *Experimental Parasitology*, 80(2):272–281, 1995.

[93] Philip J Rosenthal, James H McKerrow, David Rasnick, and James H Leech. *Plasmodium falciparum*: inhibitors of lysosomal cysteine proteinases inhibit a trophozoite proteinase and block parasite development. *Molecular and Biochemical Parasitology*, 35(2):177–183, 1989.

[94] Philip J Rosenthal, Jed E Olson, Garson K Lee, James T Palmer, Jeffrey L Klaus, and David Rasnick. Antimalarial effects of vinyl sulfone cysteine proteinase inhibitors. *Antimicrobial Agents and Chemotherapy*, 40(7):1600–1603, 1996.

[95] E Roth Jr. Plasmodium falciparum carbohydrate metabolism: a connection between host cell and parasite. *Blood cells*, 16(2-3):453–60, 1989.

[96] MA Rudzinska, W Trager, and RS Bray. Pinocytotic uptake and the digestion of hemoglobin in malaria parasites. *The Journal of Protozoology*, 12(4):563–576, 1965.

[97] Jan Schellenberger, Richard Que, Ronan MT Fleming, Ines Thiele, Jeffrey D Orth, Adam M Feist, Daniel C Zielinski, Aarash Bordbar, Nathan E Lewis, Sorena Rahmanian, et al. Quantitative prediction of cellular metabolism with constraint-based models: the *COBRA* toolbox v2. 0. *Nature Protocols*, 6(9):1290–1307, 2011.

[98] Ida Schomburg, Antje Chang, Christian Ebeling, Marion Gremse, Christian Heldt, Gregor Huhn, and Dietmar Schomburg. Brenda, the enzyme database: updates and major new developments. *Nucleic Acids Research*, 32(suppl 1):D431–D433, 2004.

[99] IW Sherman. Amino acid metabolism and protein synthesis in malarial parasites. *Bulletin of the World Health Organization*, 55(2-3):265, 1977.

[100] HE Shortt, NH Fairley, G Covell, PG Shute, and PCC Garnham. The pre-erythrocytic stage of *Plasmodium falciparum*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 44(4):405–419, 1951.

[101] HE Shortt and PCC Garnham. The pre-erythrocytic development of *Plasmodium cynomolgi* and *Plasmodium vivax*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 41(6):785–795, 1948.

[102] Ksenija Slavic, Ursula Straschil, Luc Reininger, Christian Doerig, Christophe Morin, Rita Tewari, and Sanjeev Krishna. Life cycle studies of the hexose transporter of *Plasmodium* species and genetic validation of their essentiality. *Molecular Microbiology*, 75(6):1402–1413, 2010.

[103] Kieran Smallbone, Evangelos Simeonidis, David S Broomhead, and Douglas B Kell. Something from nothing- bridging the gap between constraint-based and kinetic modelling. *Febs Journal*, 274(21):5576–5585, 2007.

[104] Kieran Smallbone, Evangelos Simeonidis, Neil Swainston, and Pedro Mendes. Towards a genome-scale kinetic model of cellular metabolism. *BMC Systems Biology*, 4(1):6, 2010.

[105] Indresh K Srivastava, Hagai Rottenberg, and Akhil B Vaidya. Atovaquone, a broad spectrum antiparasitic drug, collapses mitochondrial membrane potential in a malarial parasite. *Journal of Biological Chemistry*, 272(7):3961–3966, 1997.

[106] Natalie J Stanford, Timo Lubitz, Kieran Smallbone, Edda Klipp, Pedro Mendes, and Wolfram Liebermeister. Systematic construction of kinetic models from genome-scale metabolic networks. *PloS one*, 8(11):e79195, 2013.

[107] Jörg Stelling, Steffen Klamt, Katja Bettenbrock, Stefan Schuster, and Ernst Dieter Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420(6912):190–193, 2002.

[108] Janet Storm, Sonal Sethia, Gavin J Blackburn, Achuthanunni Chokkathukalam, David G Watson, Rainer Breitling, Graham H Coombs, and Sylke Müller. Phosphoenol pyruvate carboxylase identified as a key enzyme in erythrocytic *Plasmodium falciparum* carbon metabolism. *PLoS Pathogens*, 10(1):e1003876, 2014.

[109] Angelika Sturm, Rogerio Amino, Claudia Van de Sand, Tommy Regen, Silke Retzlaff, Annika Rennenberg, Andreas Krueger, Jörg-Matthias Pollok, Robert Menard, and Volker T Heussler. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science*, 313(5791):1287–1290, 2006.

[110] Balázs Szappanos, Károly Kovács, Béla Szamecz, Frantisek Honti, Michael Costanzo, Anastasia Baryshnikova, Gabriel Gelius-Dietrich, Martin J Lercher, Márk Jelasity, Chad L Myers, et al. An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nature Genetics*, 43(7):656–662, 2011.

[111] AL Tappel. The mechanism of the oxidation of unsaturated fatty acids catalyzed by hematin compounds. *Archives of Biochemistry and Biophysics*, 44(2):378–395, 1953.

[112] Alice S Tarun, Kerstin Baer, Ronald F Dumpit, Sean Gray, Nicholas Lejarcegui, Ute Frevert, and Stefan HI Kappe. Quantitative isolation and in vivo imaging of malaria parasite liver stages. *International Journal for Parasitology*, 36(12):1283–1293, 2006.

[113] Ines Thiele, Neil Swainston, Ronan MT Fleming, Andreas Hoppe, Swagatika Sahoo, Maike K Aurich, Hulda Haraldsdottir, Monica L Mo, Ottar Rolfsson, Miranda D Stobbe, et al. A community-driven global reconstruction of human metabolism. *Nature Biotechnology*, 31(5):419–425, 2013.

[114] William Trager and James B Jensen. Human malaria parasites in continuous culture. *Science*, 193(4254):673–675, 1976.

[115] Giel G Van Dooren, Luciana M Stimmler, and Geoffrey I McFadden. Metabolic maps and functions of the *Plasmodium* mitochondrion. *FEMS Microbiology Reviews*, 30(4):596–630, 2006.

[116] Amit Varma, Brian W Boesch, and Bernhard O Palsson. Biochemical production capabilities of *Escherichia coli*. *Biotechnology and Bioengineering*, 42(1):59–73, 1993.

[117] Amit Varma and Bernhard O Palsson. Parametric sensitivity of stoichiometric flux balance models applied to wild-type *Escherichia coli* metabolism. *Biotechnology and Bioengineering*, 45(1):69–79, 1995.

[118] Carsten Wrenger, Ingrid B Müller, Anna J Schifferdecker, Rishabh Jain, Rositsa Jordanova, and Matthew R Groves. Specific inhibition of the aspartate aminotransferase of *Plasmodium falciparum*. *Journal of Molecular Biology*, 405(4):956–971, 2011.

[119] Avner Yayon, John A Vande Waa, Malka Yayon, Timothy G Geary, and James B Jensen. Stage-dependent effects of chloroquine on *Plasmodium falciparum* in vitro. *The Journal of Protozoology*, 30(4):642–647, 1983.

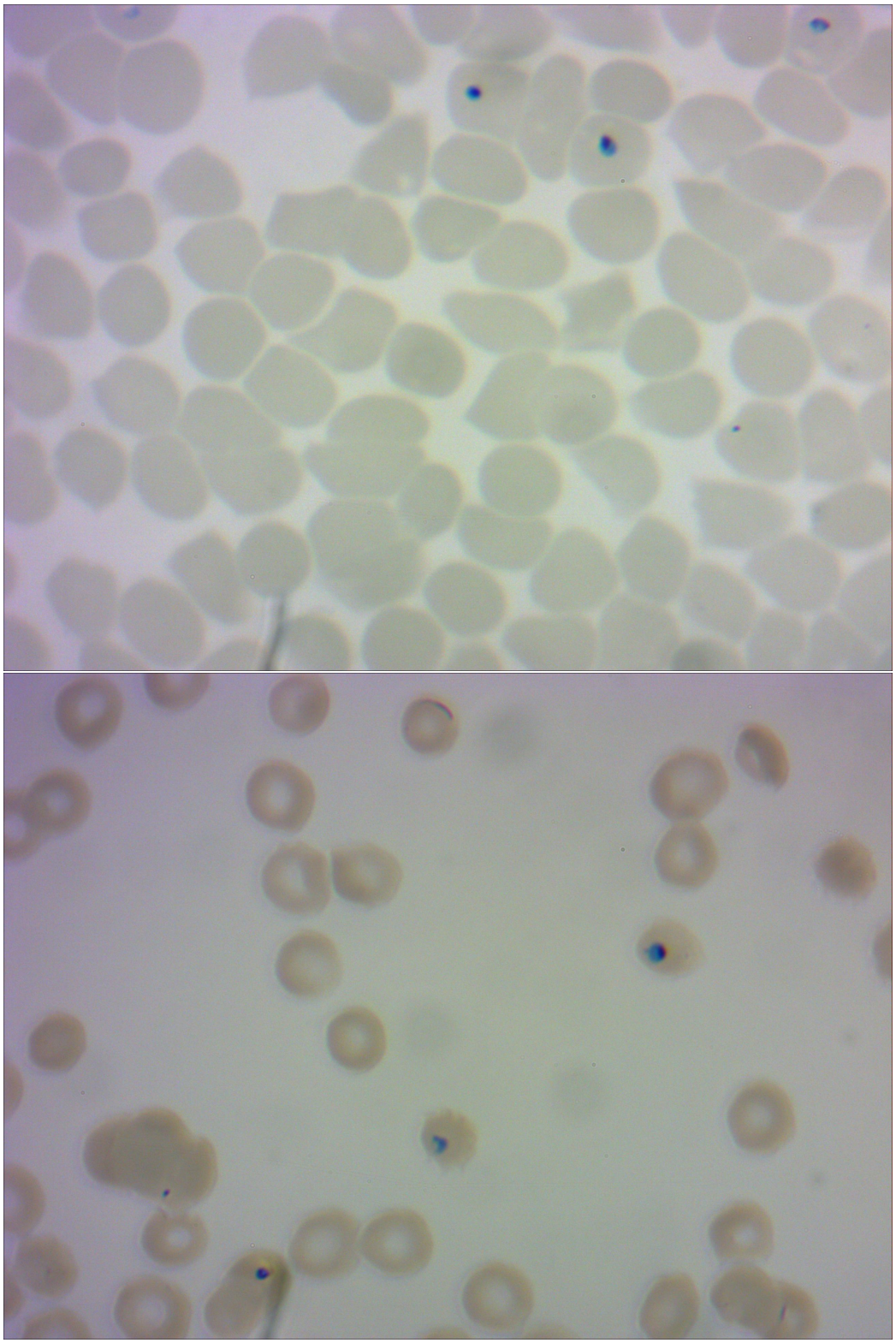
[120] Harry Yim, Robert Haselbeck, Wei Niu, Catherine Pujol-Baxley, Anthony Burgard, Jeff Boldt, Julia Khandurina, John D Trawick, Robin E Osterhout, Rosary Stephen, et al. Metabolic engineering of *Escherichia coli* for direct production of 1, 4-butanediol. *Nature Chemical Biology*, 7(7):445–452, 2011.

[121] Jorge J Yunis and Walid G Yasmineh. Glucose metabolism in human erythrocytes. In *Biochemical Methods in Red Cell Genetics*, pages 1–49. Academic Press New York, 1969.

[122] Sisi Zarchin, Miriam Krugliak, and Hagai Ginsburg. Digestion of the host erythrocyte by malaria parasites is the primary target for quinoline-containing antimalarials. *Biochemical Pharmacology*, 35(14):2435–2442, 1986.

Appendix A

Microscope blood slides showing unhealthy intraerythrocytic *P. falciparum* parasites



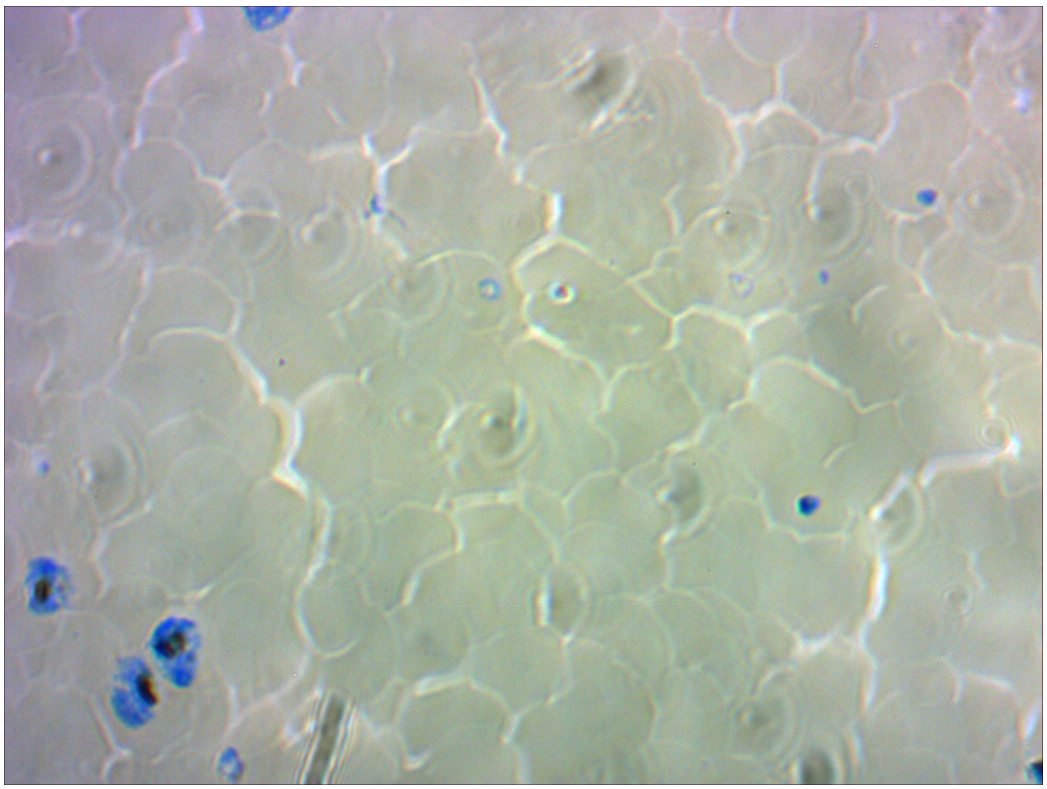


Fig.1 Microscope blood slides of limited media cultures where stressed parasite phenotypes were observed. In isoleucine-only supplemented media, within 48 hours of parasite culturing (top and middle) the majority of parasite phenotypes would present as small, dark„ compact masses that would persist over time and culture growth would stall. Compared to the growth of isoleucine- and methionine-supplemented media cultures, where growth conditions are more favourable for the parasite, it can be observed that there are fewer stressed parasites and a greater progression from rings to schizonts (bottom).

Appendix B

Model constraints

Note: The term "blim" is representative of the minimal/maximal extreme for a bound (1×10^6 mmol/gDW/h, effectively indicating boundless in a way that does not break the model.

'EX_gln_L_e_': [-0.385,blim],

'EX_hb': [-4.89e-4, 0],

'biomass': [0, 0.0550],

'EX_ile_L_e_': [-blim,0],

'EX_glc_e_': [-blim,0],

'EX_ins_e_': [0,0],

'EX_adn_e_': [0,0],

'EX_h2o2_e_': [0,0],

'EX_dgsn_e_': [0,0],

'EX_din_e_': [0,0],

'EX_nicotinamide2': [0,0],

'EX_gua_e_': [0,0],

'EX_ade_e_': [0,0],

'EX_inost_e_': [0,0],

'EX_fum_e_': [0,blim],

'EX_ura_e_': [0,0],

'EX_fru_e_': [0,0],

'EX_akg_e_': [0,0],

'EX_o2_e_': [-blim,0],

'EX_co2_e_': [0,blim],

'EX_f6p_e_': [0,0],

'EX_g6p_e_': [0,0],

'EX_man_e_': [0,0],

'EX_mal_L_e_': [0,0],

'EX_succ_e_': [0,blim],

'SERGLYexR' : [0,0]

'CYSGLYexR' : [0,0]

'CYSGLUexR' : [0,0]

'THRGLYexR' : [0,0]

'SERGLNexR' : [0,0]

'THRGLNexR' : [0,0]

'ALAGLNexR' : [0,0]

'ALAGLYexR' : [0,0]

'AKGMAL' : [0,0]

'ORNTA' : [0,0]

'NDPK1' : [0.02106,0.2106]

'ARGN' : [0.330,0.330]

'ASPTA' : [-blim,-0.001]

Appendix C

Publication

The paper in this Appendix has been accepted for publication in the Biochemical Society Transactions journal. Contributions made to this paper include the sections on "Genome-scale network analysis of metabolism" and "*Plasmodium* biomass production from haemoglobin", as well as the accompanying graphics in figure 2.

Quantitative analysis of drug effects at the whole-body level: a case study for glucose metabolism in malaria patients

Jacky L. Snoep*†‡[1], Kathleen Green*, Johann Eicher*, Daniel C. Palm*, Gerald Penkler*, Francois du Toit*, Nicolas Walters*, Robert Burger*, Hans V. Westerhoff†‡ and David D. van Niekerk*

*Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

†Molecular Cell Physiology, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands

‡MIB, University of Manchester, M1 7EN Manchester, U.K.

Abstract

We propose a hierarchical modelling approach to construct models for disease states at the whole-body level. Such models can simulate effects of drug-induced inhibition of reaction steps on the whole-body physiology. We illustrate the approach for glucose metabolism in malaria patients, by merging two detailed kinetic models for glucose metabolism in the parasite *Plasmodium falciparum* and the human red blood cell with a coarse-grained model for whole-body glucose metabolism. In addition we use a genome-scale metabolic model for the parasite to predict amino acid production profiles by the malaria parasite that can be used as a complex biomarker.

Multi-scale hierarchical modelling of disease states

Diseases manifest as phenotypic changes in whole-body physiology that are experienced as illness and can be caused by external factors (e.g. infectious diseases caused by bacteria) or by internal dysfunction (e.g. type II diabetes, cancer). Medical treatment of a disease can be in the form of medication, such as a pharmaceutical drug that affects specific reactions in target cells. In the case of a so-called 'magic bullet' drug, displaying complete specificity and complete inhibition of an essential reaction in a parasite or diseased cell, the outcome of the drug effect is direct and simple to predict and no detailed analysis is required. However, most drugs affect more than one reaction step, do not lead to complete inhibition and have side effects [1]. The response to such a drug is more complicated and a quantitative analysis of the combined effects on the whole-body physiology is necessary to evaluate the treatment. Analysis of whole-body responses to partial inhibition of one or more reaction steps is challenging; typically, mathematical models at the whole-body level are not fine-grained and do not include individual chemical reaction steps. We propose a hierarchical approach where parts of a system are resolved with sufficient detail to analyse drug effects at the individual reaction step, whereas for other parts of the system coarse-grained models are used. This approach is illustrated with a multi-scale hierarchical model for glucose metabolism in malaria patients (Figure 1).

Key words: glucose metabolism, malaria, multi-scale hierarchical model, *Plasmodium falciparum*.

[1] **To whom correspondence should be addressed (email jls@sun.ac.za).**

From diagnosis to drug

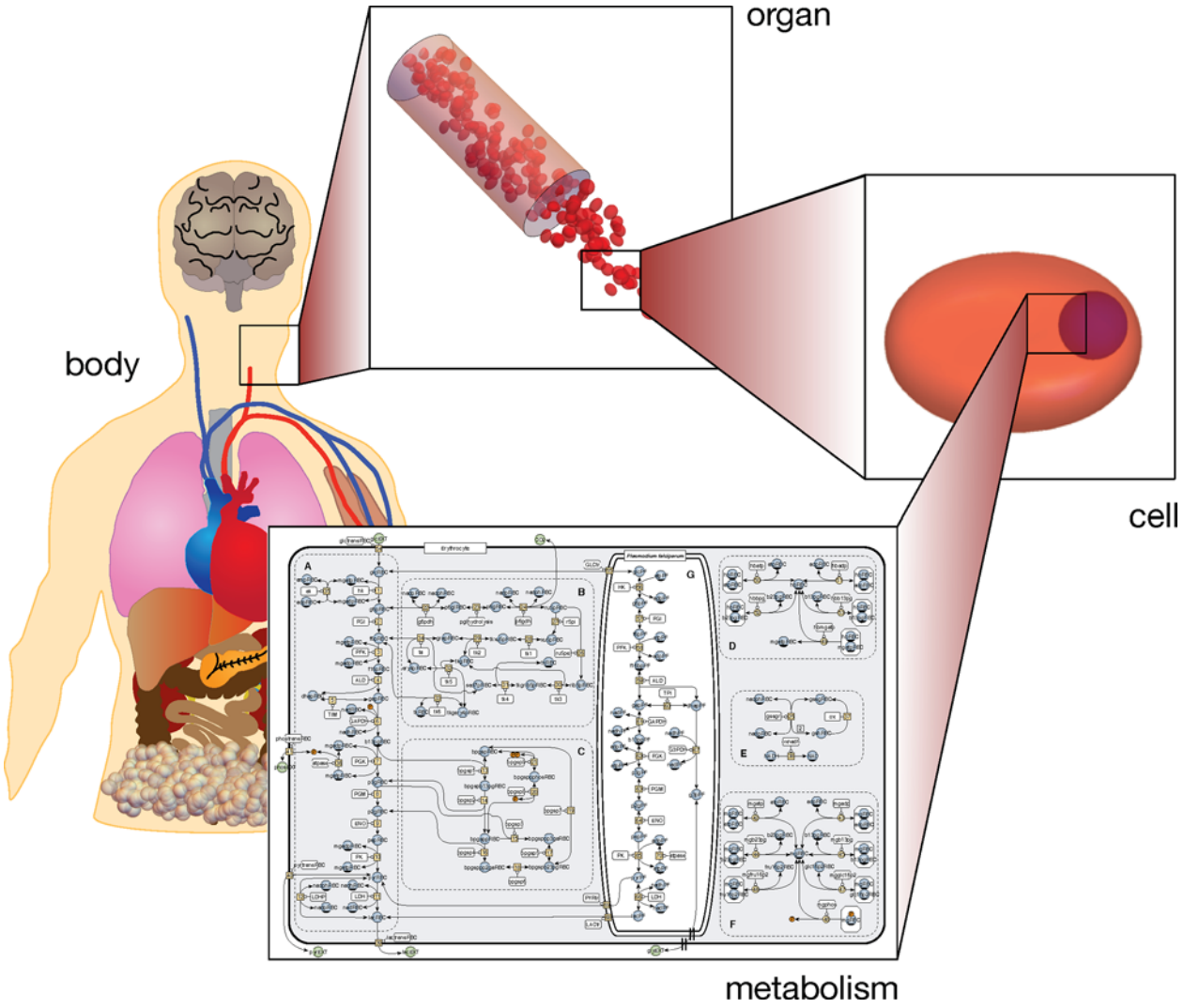
Medical treatment usually starts with a diagnosis and classification of a disease. Typically these are initially made on the basis of simple phenotypic descriptions. Relating these symptoms to a mechanistic interpretation of the disease is not a trivial task and will involve a systems based approach [2], especially when multiple molecular aetiologies with patient-dependent contributions are involved. Biomarkers play an important role in such a diagnosis and are likely to be found at the metabolomics level [3]. A mechanistic, systems level interpretation of a disease could point to multiple drug targets, moving away from the single drug single target paradigm [4]. Such systems approaches to pharmacology [5] and toxicology [6] could lead to more effective drug development strategies [7].

Genome-scale network analysis of metabolism

Metabolomics and metabolic modelling are important tools in following and predicting disease progress and understanding drug efficacy and mode of action [3,8]. In the last decade, enormous progress has been made in the genome-scale analysis of metabolic networks for a large number of species ranging from bacteria (e.g. *E. coli* [9]), to eukaryotes (e.g. *S. cerevisiae* [10]) to humans [11,12] and includes important human pathogens (e.g. *Haemophilus influenzae* [13], *Mycobacterium tuberculosis* [14] and *Plasmodium falciparum* [15–18]). These networks are very large, up to several thousands of reactions and the analyses are restricted to topological and constraint based modelling techniques,

Figure 1 | Hierarchical multi-scale model for malaria patients

The whole-body model consists of several modules at the organ level, each described with input-output functions. The red blood cell compartment is modelled at the cellular and detailed metabolic level, including *P. falciparum* metabolism.



such as flux balance analysis [19,20]. Such models have been very useful for calculating metabolic phenotypes [21], such as the prediction of changes in metabolite biomarkers for inborn errors of metabolism [12] or, for instance, in analysing medium composition requirements for bacterial growth [22]. The models are typically analysed for steady state conditions and optimization criteria (e.g. growth rate) are used to minimize the solution space. Choosing suitable constraints on exchange reactions, biomass composition and maintenance reactions can have important effects on model predictions and should be done carefully.

Why study metabolism in malaria patients?

Malaria is a dreaded disease that is widespread across tropical and sub-tropical regions and responsible for the death of between 500000 and 1000000 people per year, mostly children in sub-Saharan countries. One might not immediately think of malaria as a metabolic disease; the classic symptom of 48-h cyclical fever attacks and diagnosis via blood smears has no relation to metabolism. However, the key-diagnostics for poor chances of survival, lactic acidosis and hypoglycaemia [23] are clearly linked to metabolism. In addition, in malaria patients, blood concentrations of glycerol [24] and alanine are increased and arginine concentration is decreased [25], indicating more general metabolic changes [26].

To what extent can these metabolic changes be related to metabolic activity of the parasite? *Plasmodium* cannot synthesize its own amino acids and is dependent on the host's haemoglobin for protein biosynthesis and on the host's glucose for its free energy production. As such, the metabolic activity of the parasite has a direct effect on the host, but, in addition, the parasites cause indirect damage by lysis of red

blood cells and sequestration of parasitized red blood cells in the vasculature leading to reduced blood perfusion [27].

Although the pathophysiology used to be attributed to two main syndromes, cerebral malaria and severe anaemia malaria, it has become clear that severe malaria is complicated and involves several syndromes [27–29]. Ultimately, the goal is to delineate the individual contributions of these syndromes to the pathophysiology of malaria. Such an analysis would point at the best points of intervention to relieve the burden of the disease. In an attempt to estimate the direct contribution of *Plasmodium* activity we set out to analyse its amino acid and carbohydrate metabolism.

Plasmodium biomass production from haemoglobin

The genome of *P. falciparum* was sequenced in 2002 and several genome-scale metabolic maps have been reconstructed [30]. *Plasmodium* is severely limited in its biosynthetic reactions and is largely dependent on the host's supply of amino acids for protein synthesis, for which the parasite degrades almost all haemoglobin in the red blood cell during its 48-h growth cycle. The specific condition of *Plasmodium* growing in the red blood cell and using the available haemoglobin for protein synthesis, leads to an elegant set of constraints that can be used in a genome-scale analysis. We used an existing genome-scale model [18] with a set rate of haemoglobin consumption (assuming a 75 % consumption of total haemoglobin in 48-h [31]) and optimized for biomass production, only allowing uptake of the amino acids isoleucine and arginine. Isoleucine is not present in haemoglobin and must be taken up from the blood. *P. falciparum* is known to convert arginine to ornithine [32], leading to hypoargininaemia [25]. Under these conditions, a specific growth rate of 0.049 h^{-1} was calculated for *P. falciparum*, which is close to the expected value of 0.058 h^{-1} (calculated on the basis of 16 merozoites formed in 48-h). A glucose consumption rate of 1.6 $mmol \cdot h^{-1} \cdot gDW^{-1}$ was obtained which is somewhat lower than the experimentally measured value of 2.1 $mmol \cdot h^{-1} \cdot gDW^{-1}$ [33].

The complete set of reactions for the genome-scale network is shown in Figure 2(A), where the red lines indicate the fluxes through the reactions. In Figure 2(B), a subset of reactions involved in amino acid metabolism is shown. These reactions fall in three classes: (1) degradation of haemoglobin, (2) the synthesis of biomass and (3) export and inter-conversion of amino acids. The export fluxes for the amino acids are indicated in Figure 2(C).

The relative amino acid production rates compare well with rates observed in *P. falciparum* culture [32]. Note that the arginine to ornithine conversion was part of the objective and is therefore not a validation for the model. One cannot immediately compare these amino acid product formation rates to changes in blood amino acid concentrations in malaria patients since these concentrations are also dependent on the consumption rates in the body. However, the high capacity of the network to consume arginine and the high alanine production rates are in good agreement with the observed hypoargininaemia and high alanine blood concentrations in malaria patients. Interestingly, a high alanine blood concentration in malaria patients is usually attributed to a reduced alanine to glucose conversion in the liver [26], but our network analysis shows that a high alanine production by *Plasmodium* could contribute to this symptom. For an accurate prediction of blood concentration changes of amino acids a full body implementation of a model is required. However, high production rates (e.g. alanine) or consumption rates (e.g. arginine) due to *Plasmodium* activity can point to potential metabolic biomarkers for malaria progression. In addition, one can simulate the effect of a drug by setting a constraint on a metabolic flux in the network. If such an inhibition is complete then the network analysis can be accurate and the effect is dependent on whether the inhibited step is essential or not. If the inhibition is not complete, which is the likely scenario, it is better to analyse the effect in a kinetic model.

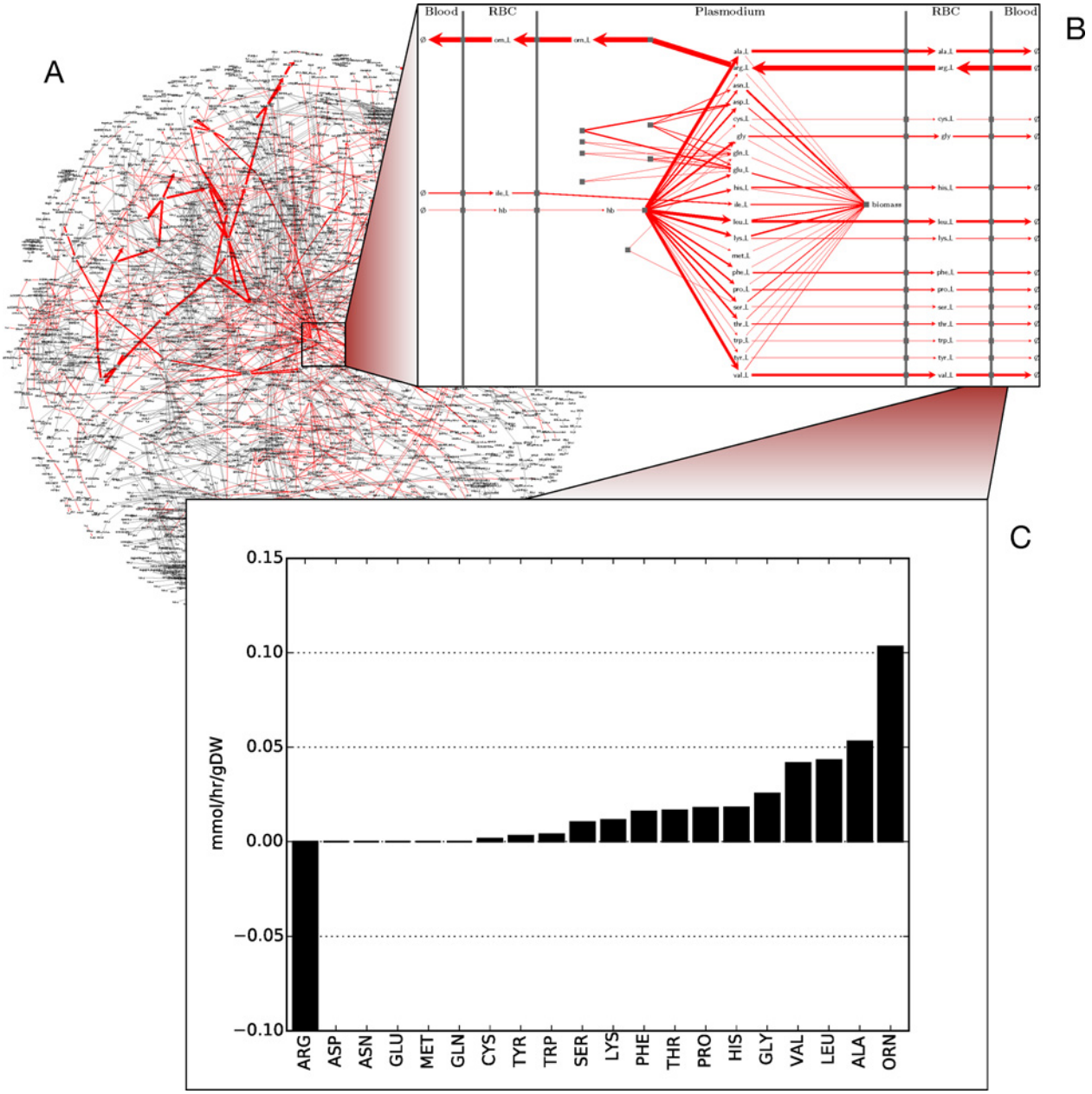
Kinetic modelling of *Plasmodium* glucose metabolism in malaria patients

Currently no detailed kinetic models exist for genome-scale networks, mostly due to limited kinetic information. Kinetic models do exist for smaller systems, such as central carbon metabolism and in a more coarse-grained form for organ and organism level metabolism. To analyse the effect of increased glycolytic activity of *Plasmodium* infected red blood cells in malaria patients, we merged three existing kinetic models: a detailed kinetic model for glycolysis of *P. falciparum* [34], a detailed kinetic model for central carbon metabolism of the red blood cell [35–37] and a coarse grained kinetic model for whole-body glucose metabolism [38]. The models were obtained from the JWS Online [39] and Biomodels [40] model repositories, corrected for units and shared-variable-names inconsistencies and integrated. No adaptations were made to the *P. falciparum* and red blood cell model and for the whole-body model only the fixed metabolites alanine and non-esterified fatty acids were changed. A detailed description of the merged model will be published elsewhere (K. Green, D.C. Palm, F. du Toit and D.D. van Niekerk and Snoep, manuscript in preparation).

Figure 3(A) shows a schema for the combined model with the compartmentalized whole-body model and the added *Plasmodium* infected red blood model. A simulation of the effect of increased parasitaemia on blood glucose concentration is given in Figure 3(B), together with patient data (and rat model data) obtained from the literature. The patient data show a lot of scatter, which is indicative of large intermittent variance (no longitudinal data for a patient followed over time was available). Most papers make reference to hypoglycaemia in severe malaria patients, but then do not report both parasitaemia and blood glucose levels. The model prediction does simulate the reference state and the few patient data with hypoglycaemia, for which data were

Figure 2 | ***P. falciparum*** **genome-scale network analysis**

The steady state solution space for the genome-scale metabolic network for *P. falciparum* [18] with a set influx rate of 0.83 μmol haemoglobin·h^{-1}·gDW^{-1} was optimized for biomass formation and ornithine production. Fluxes are indicated in red on the complete network structure in (**A**). A subset of reactions for amino acid metabolism, indicating the haemoglobin degradation, biomass formation and the amino acid inter-conversion and export are indicated in (**B**). The export fluxes of the different amino acids are indicated in (**C**).



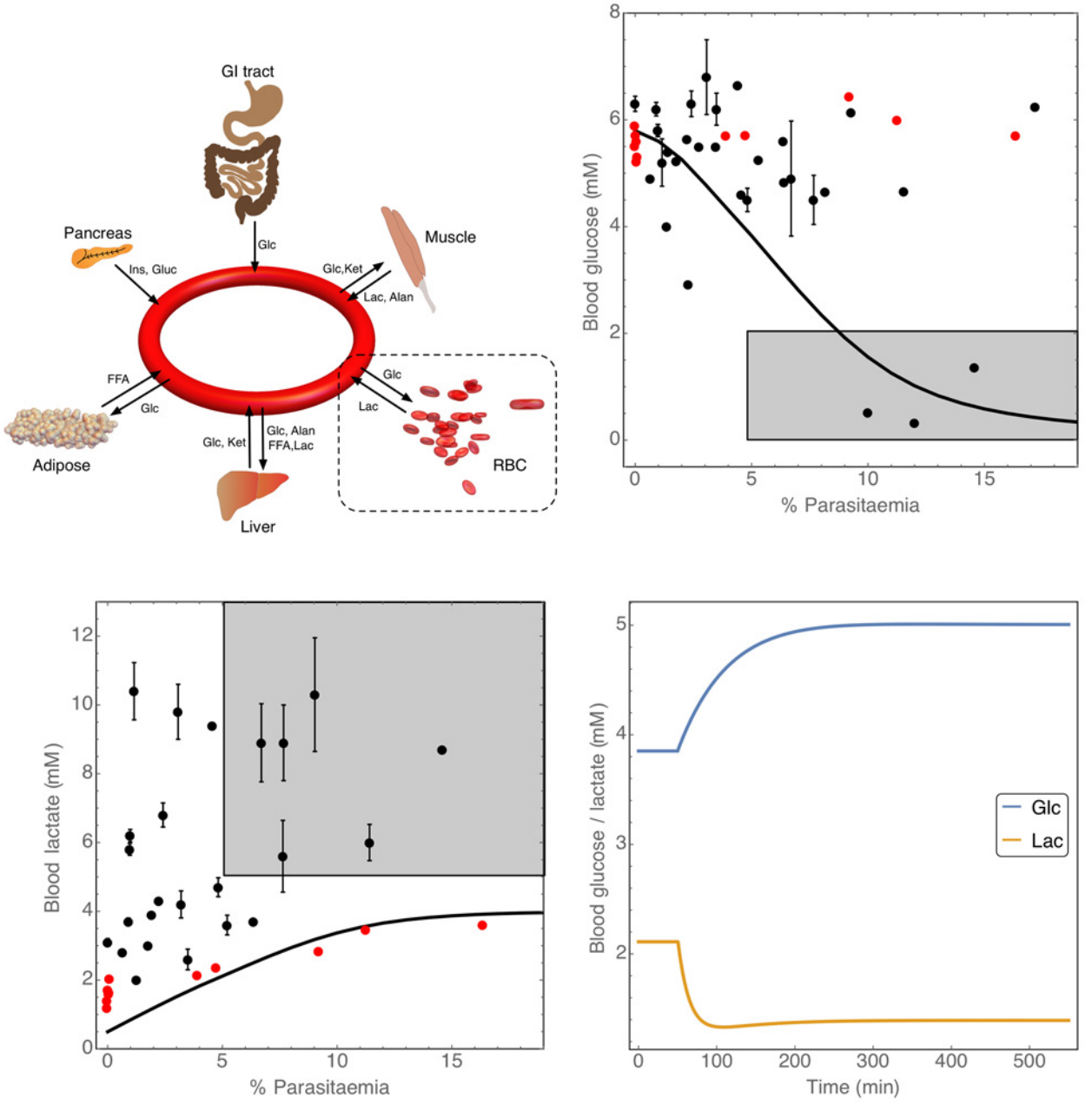
available, quite well. Similarly, the lactate data for malaria patients shows a lot of scatter (Figure 3C) and much more consistent data for the longitudinal rat study was obtained. The model prediction of lactate is low for the reference state but seems to follow the trend of lactate increase (and the rat data) reasonably well. In Figure 3(D), the results of an inhibition of the glucose transporter (to 50% of non-inhibited lactate flux) in *P. falciparum* are simulated.

Discussion and conclusion

To understand the pathophysiology of complex diseases, whole-body mathematical models can be strong tools to integrate and analyse the numerous effects that lead to the disease state. Specifically, when personal parameters can be added to such a model, they can be instrumental in choosing a correct treatment. Currently only very few molecularly informed models exist for the whole-body level

Figure 3 | Modelling glucose and lactate metabolism in malaria patients

(**A**) kinetic model for whole-body glucose metabolism [38] was merged with two detailed kinetic models for glucose metabolism in *P. falciparum* [34] and the red blood cell [35–37] (**A**). The effect of increasing levels of parasitaemia on steady state blood glucose and lactate concentrations was analysed and shown in (**B** and **C**) respectively, together with concentrations measured in malaria patients (black symbols, calculated from [41–52]) and rat data (red symbols, calculated from [53]). The shaded boxes indicate the severe malaria (> 5 % parasitaemia) and hypoglycaemia (<2 mM glucose, **B**) and lactic acidosis (>5 mM lactate, **C**) areas. (**D**) The effect of inhibition of the glucose transport step (starting at t = 50 min, resulting in 50 % reduction in glycolytic flux in the parasite) on blood glucose and lactate concentrations in a malaria patient (5 % parasitaemia) was simulated.



that are detailed enough to be useful in medical applications (e.g. http://www.entelos.com). Specifically for the simulation of pharmacological drug effects on the disease state there is a big challenge in terms of modelling at the correct level of detail. To simulate the drug effect at the reaction step, a high level of detail is needed at the drug target level, which cannot be sustained up to the whole-body level.

In the present paper, we illustrated how a hierarchical model, with a high level of detail at the drug target site and more coarse-grained for the whole-body level, can be used to simulate the effect of a pharmaceutical drug on blood glucose concentration. Clearly, the model is still in a very rudimentary stage; we only simulate the direct metabolic effect of parasite activity and have ignored any effects on blood perfusion or reduced red blood cell contents due to cell lysis or any of the large number of secondary effects caused by the malaria parasites. In addition, we simulated the drug effect by simply assuming a constant inhibitor concentration in the blood, a much more realistic simulation would have to include a full PK/PD (pharmacokinetics/pharmacodynamics) model to evaluate the drug efficacy [5,54].

The aim to mechanistically simulate drug effects (inhibiting a specific target reaction) at the whole-body level (physiological disease state) is very ambitious. However, we feel the time is right for this. Some large-scale projects have existed for quite some time and produced detailed kinetic models at the organ level that can be extended to the whole-body level (e.g. the physiome project, http://physiomeproject.org [55]; the virtual liver, http://www.virtual-liver.de [56]). In addition, a much stronger adherence to modelling standards, such as description in standard formats [systems biology markup language (SBML) and CellML)] and storage in curated model databases (JWS Online, Biomodels and CellML), makes model reuse much easier. For our initial model construction, we merged three existing models and this enabled us to make some preliminary simulations at different hierarchical levels. Of course, one cannot simple merge all existing models; they must be compatible and constructed for similar physiological conditions [57]. After merging of existing models, one can start a number of iterative cycles to improve the integral model and adapt it to specific disease states.

With the present concept paper, we hope to have illustrated the approach we follow towards whole-body modelling of blood glucose and lactate metabolism in malaria patients. Clearly, much work is still needed and especially at the whole-body level the model needs to be validated more thoroughly. For this, we will first work in a rat model system for which it is much easier to obtain longitudinal data. Although the specific model for malaria patients is still very preliminary, the suggested approach of a hierarchical model structure with a high level of detail at the drug target level and more coarse-grained models at the whole-body physiological level, is generic and could be applied to other complex metabolic diseases such as type II diabetes.

Acknowledgements

We thank Gunnar Cedersund and Brett Olivier, who were involved in the initial stages of the construction of the dynamic and structural model respectively and Barbara Bakker for discussing potential modelling strategies at the whole-body disease state.

Funding

This work was supported by the National Research Foundation, South Africa [grant numbers SARCHI 82813 (to J.L.S.) and TTK14051967526 (to D.D.v.N.)].

References

1 Xie, L., Xie, L., Kinnings, S. and Bourne, P. (2012) Novel computational approaches to polypharmacology as a means to define responses to individual drugs. Annu. Rev. Pharmacol. Toxicol. **52**, 361–379 CrossRef PubMed
2 Fryburg, D., Song, D., Laifenfeld, D. and de Graaf, D. (2014) Systems diagnostics: anticipating the next generation of diagnostic tests based on mechanistic insight into disease. Drug Discov. Today **19**, 108–112 CrossRef PubMed
3 Kell, D.B. and Goodacre, R. (2014) Metabolomics and systems pharmacology: why and how to model the human metabolic network for drug discovery. Drug Discov. Today **19**, 171–182 CrossRef PubMed
4 Hopkins, A. (2008) Network pharmacology: the next paradigm in drug discovery. Nat. Chem. Biol. **4**, 682–690 CrossRef PubMed
5 Zhao, S. and Iyengar, R. (2012) Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. Annu. Rev. Pharmacol. Toxicol. **52**, 505–521 CrossRef PubMed
6 Bai, J. and Abernethy, D. (2013) Systems pharmacology to predict drug toxicity: integration across levels of biological organization. Annu. Rev. Pharmacol. Toxicol. **53**, 451–473 CrossRef PubMed
7 Butcher, E., Berg, E. and Kunkel, E. (2004) Systems biology in drug discovery. Nat. Biotechnol. **22**, 1253–1259 CrossRef PubMed
8 Kell, D. (2006) Systems biology, metabolic modelling and metabolomics in drug discovery and development. Drug Discov. Today **11**, 1085–1092 CrossRef PubMed
9 Edwards, J. and Palsson, B. (2000) The *Escherichia coli* mg1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. Proc. Natl. Acad. Sci. U.S.A. **97**, 5528–5533 CrossRef PubMed
10 Forster, J., Famili, I., Fu, P., Palsson, B. and Nielsen, J. (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res. **13**, 244–253 CrossRef PubMed
11 Duarte, N.C., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M.L., Vo, T.D., Srivas, R. and Palsson, B.O. (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc. Natl. Acad. Sci. U.S.A. **104**, 1777–1782 CrossRef PubMed
12 Thiele, I., Swainston, N., Fleming, R.M., Hoppe, A., Sahoo, S., Aurich, M.K., Haraldsdottir, H., Mo, M.L., Rolfsson, O., Stobbe, M.D. et al. (2013) A community-driven global reconstruction of human metabolism. Nat. Biotechnol. **5**, 419–425 CrossRef
13 Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A. and Merrick, J.M. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269**, 496–512 CrossRef PubMed
14 Jamshidi, N. and Palsson, B. (2007) Investigating the metabolic capabilities of *Mycobacterium tuberculosis* h37rv using the *in silico* strain inj661 and proposing alternative drug targets. BMC Syst. Biol. **1**, 26 CrossRef PubMed
15 Yeh, I., Hanekamp, T., Tsoka, S., Karp, P. and Altman, R. (2004) Computational analysis of *Plasmodium falciparum* metabolism: organizing genomic information to facilitate drug discovery. Genome Res. **14**, 917–924 CrossRef PubMed
16 Fatumo, S., Plaimas, K., Mallm, J.P., Schramm, G., Adebiyi, E., Oswald, M., Eils, R. and Konig, R. (2009) Estimating novel potential drug targets of *Plasmodium falciparum* by analysing the metabolic network of knock-out strains *in silico*. Infect. Genet. Evol. **9**, 351–358 CrossRef PubMed
17 Huthmacher, C., Hoppe, A., Bulik, S. and Holzhutter, H.G. (2010) Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. BMC Syst. Biol. **4**, 120 CrossRef PubMed
18 Plata, G., Hsiao, T.L., Olszewski, K., Llinas, M. and Vitkup, D. (2010) Reconstruction and fluxbalance analysis of the *Plasmodium falciparum* metabolic network. Mol. Syst. Biol. **6**, 408 CrossRef PubMed
19 Fell, D. and Small, J. (1986) Fat synthesis in adipose tissue. An examination of stoichiometric constraints. Biochem. J. **238**, 781–786 CrossRef PubMed

20 Orth, J., Thiele, I. and Palsson, B. (2010) What is flux balance analysis? Nat. Biotechnol. **28**, 245–248 CrossRef PubMed
21 McCloskey, D., Palsson, B. and Feist, A. (2013) Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. Mol. Syst. Biol. **9**, 661 CrossRef PubMed
22 Teusink, B., Wiersma, A., Molenaar, D., Francke, C., de Vos, W.M., Siezen, R.J. and Smid, E.J (2006) Analysis of growth of *Lactobacillus plantarum* wcfs1 on a complex medium using a genome-scale metabolic model. J. Biol. Chem. **281**, 40041–40048 CrossRef PubMed
23 Krishna, S., Waller, D.W., ter Kuile, F., Kwiatkowski, D., Crawley, J., Craddock, C.F., Nosten, F., Chapman, D., Brewster, D., Holloway, P.A. and White, N.J. (1994) Lactic acidosis and hypoglycaemia in children with severe malaria: pathophysiological and prognostic significance. Trans. R. Soc. Trop. Med. Hyg. **88**, 67–73 CrossRef PubMed
24 Pukrittayakamee, S., White, N.J., Davis, T.M., Supanaranond, W., Crawley, J., Nagachinta, B. and Williamson, D.H. (1994) Glycerol metabolism in severe falciparum malaria. Metabolism **43**, 887–892 CrossRef PubMed
25 Lopansri, B.K., Anstey, N.M., Weinberg, J.B., Stoddard, G.J., Hobbs, M.R., Levesque, M.C., Mwaikambo, E.D. and Granger, D.L. (2003) Low plasma arginine concentrations in children with cerebral malaria and decreased nitric oxide production. Lancet **361**, 676–678 CrossRef PubMed
26 Planche, T., Dzeing, A., Ngou-Milama, E., Kombila, M. and Stacpoole, P. (2005) Metabolic complications of severe malaria. Curr. Top. Microbiol. Immunol. **295**, 105–136 PubMed
27 Mackintosh, C., Beeson, J. and Marsh, K. (2004) Clinical features and pathogenesis of severe malaria. Trends Parasitol **20**, 597–603 CrossRef PubMed
28 Clark, I. and Cowden, W. (2003) The pathophysiology of falciparum malaria. Pharmacol. Ther. **99**, 221–260 CrossRef PubMed
29 Maitland, K. and Marsh, M. (2004) Pathophysiology of severe malaria in children. Acta Trop. **90**, 131–140 CrossRef PubMed
30 Tymoshenko, S., Oppenheim, R., Soldati-Favre, D. and Hatzimanikatis, V. (2013) Functional genomics of *Plasmodium falciparum* using metabolic modelling and analysis. Brief Funct. Genomics **12**, 316–327 CrossRef PubMed
31 Esposito, A., Tiffert, T., Mauritz, J.M., Schlachter, S., Bannister, L.H., Kaminski, C.F. and Lew, V.L. (2008) Fret imaging of hemoglobin concentration in *Plasmodium falciparum*-infected red cells. PLoS One **3**, e3780 CrossRef PubMed
32 Olszewski, K.L., Morrisey, J.M., Wilinski, D., Burns, J.M., Vaidya, A.B., Rabinowitz, J.D. and Llinas, M. (2009) Host-parasite interactions revealed by *Plasmodium falciparum* metabolomics. Cell Host Microbe **5**, 191–199 CrossRef PubMed
33 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 CrossRef PubMed
34 Penkler, G., du Toit, F., Adams, W., Rautenbach, M., Palm, D.C., van Niekerk, D.D. and Snoep, J.L. (2015) Construction and validation of a detailed kinetic model of glycolysis in *Plasmodium falciparum*. FEBS J. **282**, 1481–1511 CrossRef PubMed
35 Mulquiney, P.J., Bubb, W.A. and Kuchel, P.W. (1999) Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: *in vivo* kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using 13C and 31P NMR. Biochem. J. **342**, 567–580 CrossRef PubMed
36 Mulquiney, P.J. and Kuchel, P.W. (1999) Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: computer simulation and metabolic control analysis. Biochem. J. **342**, 597–604 CrossRef PubMed
37 Mulquiney, P.J. and Kuchel, P.W. (1999) Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: equations and parameter refinement. Biochem. J. **342**, 581–596 CrossRef PubMed
38 Xu, K., Morgan, K.T., Gehris, A.T., Elston, T. and Gomez, S. (2011) A whole-body model for glycogen regulation reveals a critical role for substrate cycling in maintaining blood glucose homeostasis. PLoS Comput. Biol. **7**, e1002272 CrossRef PubMed
39 Olivier, B. and Snoep, J. (2004) Web-based kinetic modelling using JWS online. Bioinformatics **20**, 2143–2144 CrossRef PubMed
40 Le Novere, N., Bornstein, B., Broicher, A., Courtot, M., Donizelli, M., Dharuri, H., Li, L., Sauro, H., Schilstra, M., Shapiro, B. et al. (2006) Biomodels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Res. **34**, D689–D691 CrossRef PubMed
41 White, N.J., Marsh, K., Turner, R.C., Miller, K.D., Berry, C.D., Williamson, D.H. and Brown, J. (1987) Hypoglycaemia in African children with severe malaria. Lancet **329**, 708–711 CrossRef
42 Taylor, T.E., Molyneux, M.E., Wirima, J.J., Fletcher, K.A. and Morris, K. (1988) Blood glucose levels in Malawian children before and during the administration of intravenous quinine for severe falciparum malaria. N. Engl. J. Med. **319**, 1040–1047 CrossRef PubMed
43 Davis, T.M., Looareesuwan, S., Pukrittayakamee, S., Levy, J.C., Nagachinta, B. and White, N.J. (1993) Glucose turnover in severe falciparum malaria. Metabolism **42**, 334–340 CrossRef PubMed
44 Taylor, T.E., Borgstein, A. and Molyneux, M.E. (1993) Acid-base status in paediatric *Plasmodium falciparum* malaria. Q. J. Med. **86**, 99–109 PubMed
45 Waller, D., Krishna, S., Crawley, J., Miller, K., Nosten, F., Chapman, D., ter Kuile, F.O., Craddock, C., Berry, C., Holloway, P.A. et al. (1995) Clinical features and outcome of severe malaria in Gambian children. Clin. Infect. Dis. **21**, 577–587 CrossRef PubMed
46 Allen, S.J., O'Donnell, A., Alexander, N.D. and Clegg, J.B. (1996) Severe malaria in children in Papua New Guinea. QJM **89**, 779–788 CrossRef PubMed
47 Day, N.P., Phu, N.H., Mai, N.T., Chau, T.T., Loc, P.P., Chuong, L.V., Sinh, D.X., Holloway, P., Hien, T.T. and White, N.J. (2000) The pathophysiologic and prognostic significance of acidosis in severe adult malaria. Crit. Care Med. **28**, 1833–1840 CrossRef PubMed
48 Planche, T., Agbenyega, T., Bedu-Addo, G., Ansong, D., Owusu-Ofori, A., Micah, F., Anakwa, C., Asafo-Agyei, E., Hutson, A., Stacpoole, P.W. and Krishna, S. (2003) A prospective comparison of malaria with other severe diseases in African children: prognosis and optimization of management. Clin. Infect. Dis. **37**, 890–897 CrossRef PubMed
49 Dzeing-Ella, A., Nze Obiang, P.C., Tchoua, R., Planche, T., Mboza, B., Mbounja, M., Muller-Roemer, U., Jarvis, J., Kendjo, E., Ngou-Milama, E. et al. (2005) Severe falciparum malaria in Gabonese children: clinical and laboratory features. Malar. J. **4**, 1 CrossRef PubMed
50 Jarvis, J.N., Planche, T., Bicanic, T., Dzeing-Ella, A., Kombila, M., Issifou, S., Borrmann, S., Kremsner, P.G. and Krishna, S. (2006) Lactic acidosis in gabonese children with severe malaria is unrelated to dehydration. Clin. Infect. Dis. **42**, 1719–1725 CrossRef PubMed
51 Issifou, S., Kendjo, E., Missinou, M.A., Matsiegui, P.B., Dzeing-Ella, A., Dissanami, F.A., Kombila, M., Krishna, S. and Kremsner, P.G (2007) Differences in presentation of severe malaria in urban and rural Gabon. A. J. Trop. Med. Hyg. **77**, 1015–1019
52 Seydel, K.B., Kampondeni, S.D., Valim, C., Potchen, M.J., Milner, D.A., Muwalo, F.W., Birbeck, G.L., Bradley, W.G., Fox, L.L., Glover, S.J. et al. (2015) Brain swelling and death in children with cerebral malaria. N. Engl. J. Med. **372**, 1126–1137 CrossRef PubMed
53 Holloway, P.A., Knox, K., Bajaj, N., Chapman, D., White, N.J., O'Brien, R., Stacpoole, P.W. and Krishna, S. (1995) *Plasmodium berghei* infection: dichloroacetate improves survival in rats with lactic acidosis. Exp. Parasitol. **80**, 624–632 CrossRef PubMed
54 van der Graaf, P. and Benson, B. (2011) Systems pharmacology: bridging systems biology and pharmacokinetics-pharmacodynamics (PKPD) in drug discovery and development. Pharm. Res. **28**, 1460–1464 CrossRef PubMed
55 Hunter, P., Crampin, E. and Nielsen, P. (2008) Bioinformatics, multiscale modeling and the IUPS physiome project. Brief. Bioinform. **9**, 333–343 CrossRef PubMed
56 Holzhutter, H., Drasdo, D., Preusser, T., Lippert, J. and Henney, A. (2012) The virtual liver: a multidisciplinary, multilevel challenge for systems biology. Wiley Interdiscip. Rev. Syst. Biol. Med. **4**, 221–235 CrossRef PubMed
57 Snoep, J.L., Bruggeman, F., Olivier, B.G. and Westerhoff, H.V. (2006) Towards building the silicon cell: a modular approach. Biosystems **83**, 207–216 CrossRef PubMed

Received 23 June 2015
doi:10.1042/BST20150145