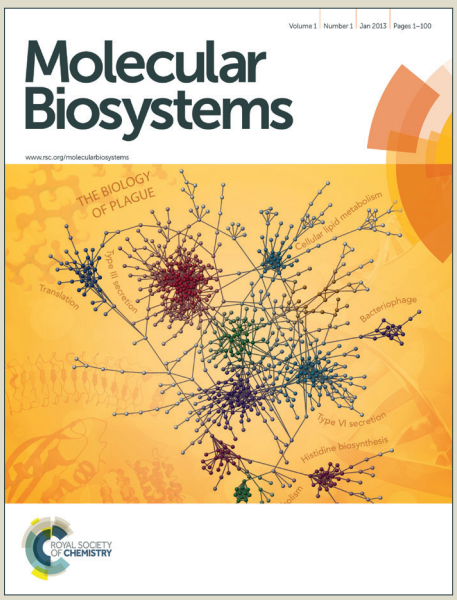
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Modeling metabolism and stage-specific growth of *Plasmodium falciparum* HB3 during the intraerythrocytic developmental cycle

Xin Fang, Jaques Reifman,* and Anders Wallqvist

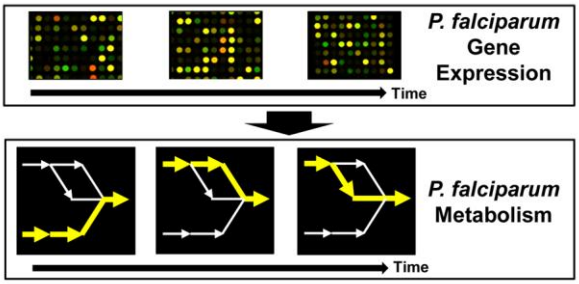
Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Ft. Detrick, MD 21702, USA

Tel.: +1-301-619-7915; Fax: +1-301-619-1983
E-mail: jaques.reifman.civ@mail.mil.

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Table of contents entry

We developed a metabolic network model that maps hourly gene expression to time-dependent metabolism and stage-specific growth, allowing us to link specific metabolites or pathways to specific physiological functions.



PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Abstract

The human malaria parasite *Plasmodium falciparum* goes through a complex life cycle, including a roughly 48-hour-long intraerythrocytic developmental cycle (IDC) in human red blood cells. A better understanding of the metabolic processes required during the asexual blood-stage reproduction will enhance our basic knowledge of *P. falciparum* and help identify critical metabolic reactions and pathways associated with blood-stage malaria. We developed a metabolic network model that mechanistically links time-dependent gene expression, metabolism, and stage-specific growth, allowing us to predict the metabolic fluxes, the biomass production rates, and the timing of production of the different biomass components during the IDC. We predicted time- and stage-specific production of precursors and macromolecules for *P. falciparum* (strain HB3), allowing us to link specific metabolites to specific physiological functions. For example, we hypothesized that coenzyme A might be involved in late-IDC DNA replication and cell division. Moreover, the predicted ATP metabolism indicated that energy was mainly produced from glycolysis and utilized for non-metabolic processes. Finally, we used the model to classify the entire tricarboxylic acid cycle into segments, each with a distinct function, such as superoxide detoxification, glutamate/glutamine processing, and metabolism of fumarate as a byproduct of purine biosynthesis. By capturing the normal metabolic and growth progression in *P. falciparum* during the IDC, our model provides a starting point for further elucidation of strain-specific metabolic activity, host-parasite interactions, stress-induced metabolic responses, and metabolic responses to antimalarial drugs and drug candidates.

Introduction

Malaria constitutes a major human health threat, with more than 207 million clinical cases and 627,000 deaths annually.[1] Despite the success of current anti-malarial efforts, including vector control, the improvement of diagnostic testing, and drug development, the complete control and elimination of this disease is hindered by an emerging resistance to existing drugs and lack of effective prophylactic vaccines.[1] Thus, efforts to develop truly novel treatment options for this disease warrant basic research into the different mechanisms used by *Plasmodium falciparum*, the most virulent causative agent of malaria, to survive and proliferate in host-based physiological environments.[2]

One important feature of this process is the organism's life cycle, during which it adopts distinct and radically different morphological stages.[3] The malaria parasite invades the host through the bite of an infected mosquito. The infective sporozoite rapidly moves to the liver, where it initially proliferates into merozoites, largely asymptomatically to the host. Subsequently, the merozoites enter the bloodstream. While some merozoites develop sexual forms that re-infect mosquitos, others begin a roughly 48-hour-long intraerythrocytic developmental cycle (IDC) of asexual reproduction. This allows the parasites to infect many more red blood cells and leads to the well-known malarial symptom of recurring fever. During this cycle, the parasites are susceptible to drug treatments, and efforts to elucidate essential biological activities such as metabolism, the focus of this study, provide a foundation for developing novel therapies.

Experimental studies of *P. falciparum* metabolism can be classified into two categories as focused either on single pathways or on more comprehensive "omics"-based studies. In the former category, studies have shed light on how *P. falciparum* during the IDC satisfies its energy

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

demand by oxidizing the bulk of the taken-up glucose into lactate through the glycolysis pathway.[4] Although the organism has all the enzymes to execute a fully functional tricarboxylic acid (TCA) cycle, its role is not fully elucidated, as only a fraction of the glucose is processed in this pathway.[5, 6] However, these studies could not comprehensively capture all aspects of *P. falciparum* metabolism during the IDC. Conversely, the latter studies include high-throughput "omics" profiling, *e.g.*, genome-scale transcriptomics profiling in *P. falciparum* during different stages of its life cycle.[7, 8] Of particular interest for IDC studies, the expression data for nearly all genes (including metabolic genes) have been collected at each hour during the IDC, providing a continual transcriptomics readout of stage-specific developments.[9, 10] Again, these data have primarily been analyzed in terms of individual metabolic pathways and not from a fully integrated metabolic analysis that takes into account the entire IDC.[11, 12] Here, we performed a system-level investigation of IDC metabolism using a combination of the high-throughput gene expression data with *in silico* metabolic network modeling to systematically connect altered genetic transcriptions to enzymes and metabolic activities.

Systems biology representations of metabolism use the constituent components (genes, metabolites, and reactions) to predict metabolic activity and growth phenotypes of organisms.[13] Genome-scale metabolic networks consist of interconnected biochemical reactions, each processing particular metabolites spontaneously or through enzyme(s) encoded by gene(s), which, when they were analyzed under certain constrained conditions, such as limited nutrient uptake, can predict cellular growth (biomass accumulation) and other phenotypic functions related to metabolism.[13] For instance, metabolic networks for *P. falciparum* have been developed and used to identify essential genes/reactions that represent candidates for target-based anti-malarial drug discovery.[14-19] Importantly, these networks were integrated with gene expression

data at different stages of the IDC, generating static models of stage-specific *P. falciparum* metabolic activities and correctly predicting metabolite exchanges between the parasite and the host.[16, 17] A limitation with this approach is that these models did not capture the parasite's stage-specific growth, failing to link genotypic alterations and phenotypic growth of *P. falciparum* during the IDC. These models used fixed biomass functions, and, thus, were not able to capture the distinct time-dependent synthesis of each biomass component and relate these components to stage-specific cellular activities.

To overcome these limitations, we created an integrated "global" *P. falciparum* metabolic network model that can predict metabolic fluxes, biomass production rates (representative of specific growth rates),[20] and the production of biomass components at each hour during the IDC. We validated our model by comparing the predicted production of the biomass and its components with the experimentally observed stage-specific growth and macromolecular syntheses. In addition, we analyzed the predicted net production of small metabolites, the ATP production and consumption, and the fluxes through the TCA cycle, providing a rationale for the stage-specific functions of these molecules, energy metabolism, and metabolic pathways. Thus, our model provides a framework to link gene expression, metabolic fluxes, and growth phenotypes that can be used to model and interpret *P. falciparum* metabolism under different conditions, including drug treatments and other physiological stress conditions.

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Results and Discussion

Metabolic model of *P. falciparum* during the IDC

We created a model that describes all metabolic activities in *P. falciparum* HB3 during asexual reproduction in red blood cells, by estimating *1*) the distribution of metabolite flows (or fluxes), *2*) an overall biomass production rate, and *3*) the production rate of each component of the organism's biomass at each hour of the IDC. We used a modified metabolic network reconstruction[16] (see Experimental Section for details) combined with hourly gene expression data from the highest time-resolution experiment available to date.[9] This approach used the time-series gene expression data[9] to alter a set of "nominal" metabolic fluxes, representing the average physiological reaction fluxes of *P. falciparum* in the IDC.

Despite the lack of a reliable correlation between gene transcription levels and enzyme activities, a number of methodologies have demonstrated the ability to capture condition-specific metabolic behaviors via the integration of a metabolic network with either absolute gene transcriptional data[21-23] or differential gene expressions compared to particular reference conditions.[24, 25] Here, we constructed a hybrid version of these approaches, as the gene expression data represent gene transcriptional differences between the sample at a particular time point and a mixed pool of samples from all time points.[9] Importantly, we used a semi-continuous approach to account for gradual changes in metabolite fluxes and growth phenotypes rather than considering metabolic reactions in a binary manner (active or inactive depending on predefined stage-specific metabolic states).[26, 27] This approach represents a data-driven methodology in which the time-dependent state of the metabolic network (and organism) is driven and determined by alterations in gene transcription.

Figure 1 schematically outlines the major steps we used to perform the integration of the metabolic network with time-series gene expression data. In this exemplar, we used a set of time-series gene expression data and a metabolic network containing five metabolites (A–E), two uptake reactions, three enzymatic reactions, and one biomass reaction. In *Step I*, we initially used the network to construct a nominal flux distribution that satisfied the mass balance of each metabolite and the average biomass composition. In *Step II*, we mapped the gene expression data into reactions. Given the nominal fluxes and relative expression data, we generated a set of time-series metabolic fluxes that fluctuated around their nominal value (*Step III*). For each reaction, the time-dependent pattern of these fluxes followed, as closely as possible, that of the corresponding expression levels. Using these fluxes, we could determine the syntheses and net production of all biomass components (B, C, and E in Figure 1), which when summed up yielded the overall time-series biomass production rate (*Step IV*).

The malaria model also incorporated mass conservation constraints to account for the growth from one organism to 16–32 organisms through four to five cell division cycles during the IDC. Furthermore, given that there is no one-to-one mapping between gene expression levels and enzyme activity, we introduced time shifts in the model to account for the time difference between when a gene is transcribed (the available data) to when the synthesized proteins appear and are active in the organism (metabolic activity as discussed in the paper). This procedure resulted in time-series metabolic states that indicated time-dependent alterations of reaction fluxes, the biomass production, and the synthesis of each biomass component. We provide a detailed description of these steps and their implementation in the Experimental Section.

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Prediction of the overall biomass production

Using the developed model, we predicted the overall biomass production rate μ of *P. falciparum* at each of the 48 hours during the IDC (see Figure 2). The prediction shows an overall qualitative agreement with the experimentally observed growth phenotypes of the parasite during the three stages (ring, trophozoite, and schizont) throughout the cycle. We predicted that the biomass production rate started from a very low level and increased during the ring stage (first 18 hours). The relatively low production rate is consistent with the experimental slow-growth phenotype in the ring stage,[28] during which the primary activity of the organism is remodeling its internal structures.[29] Further, we predicted sustained biomass production rates at relatively high levels during the trophozoite stage (18–30 hours), during which *P. falciparum* is observed to grow rapidly by consuming most of the cytoplasm of its host red blood cells.[30] Finally, the biomass production rates after 30 hours decreased, which is compatible with the observed schizont-stage *P. falciparum* shift in focus from growth to cell division.[29] In particular, the predicted biomass production rate at the end of the IDC was close to that at the beginning, suggesting that the metabolic program goes back to its original state after one round of the cycle and readies itself for another round of infection.[29]

Prediction of time-dependent macromolecule biomass-component synthesis

P. falciparum accumulates biomass by taking up nutrients and synthesizing required components, such as DNA, RNA, protein, and phospholipids. However, these processes are not uniformly active during the IDC and instead exhibit time- or stage-specific dependencies indicative of tightly controlled metabolic processes during merozoite asexual reproduction. We used our model to predict the net amounts of each of these macromolecules produced during different time intervals throughout the IDC, and compared the predicted results with the corresponding

HB3 strain-specific experimental data.[31, 32] This comparison was especially relevant as the measured levels from the experimental system of *P. falciparum*-infected erythrocytes are not obscured by difficulties in separating host and parasite components. Mature human red blood cells contain no nucleus to synthesize DNA and RNA,[33] are thought not to synthesize protein,[33] and have negligible phospholipid syntheses and metabolism.[34] Hence, the measured syntheses of these molecules are attributable to *P. falciparum* and not to the erythrocytes.

Figures 3A and 3B show that the predicted syntheses of RNA and protein, respectively, peaked during the mid-IDC. Specifically, both the prediction and experimental data showed that the net amount of RNA synthesized was relatively high between 12 and 36 hours after infection, with the highest amount produced at around 24 hours.[31] Figure 3B shows that our prediction results for the synthesized amount of protein were relatively high between 18 and 40 hours and at the highest level between 24 and 36 hours. Although no quantitative experimental measurements of protein synthesis were available, our predictions were qualitatively supported by the experimental data from polyacrylamide gel electrophoresis, which showed that the synthesized amounts of soluble proteins were relatively high between 24 and 36 hours, while those for antigens were high between 18 and 40 hours.[31]

Figures 3C and 3D show that our predictions and experimental results indicate that DNA and phospholipids, respectively, were mainly synthesized in *P. falciparum* during the late IDC. Figure 3C shows that the predicted DNA synthesis amount was the highest between 36 and 40 hours, compared to the maximum experimental amounts, which occurred a few hours earlier during the 30- to 36-hour time interval. Figure 3D shows that both the predicted and experimental synthesized amounts of phospholipids during the 30- to 43-hour time intervals were higher compared to the earlier intervals. Although the predicted and experimental amounts[32]

reached their maximums at different time intervals, the trends in both sets of data showed general increases from the early to the late IDC, confirming that phospholipids were mainly synthesized during the late stage of the cycle. The timing of the DNA and phospholipid syntheses prepares the organism for cell division during the schizont stage by generating the needed genomic materials and cellular membranes to support cell replication.

Despite the qualitative consistency in the timing of macromolecular production, the predicted and experimental results still exhibited quantitative discrepancies. For example, the predicted maximum in DNA synthesis occurred six hours later than experimentally observed. These discrepancies are partly a reflection of our limited consideration of post-transcriptional and post-translational regulations.[35] As outlined in the Experimental Section, we did approximately account for these processes by shifting the gene expression data by the average time delay derived from comparing a set of bi-hourly time-series proteomics data with the corresponding transcriptional levels.[36] However, time delays might vary non-linearly during the developmental stages, *e.g.*, due to synchronization switches or just a lack of available ribosomes. For ribosomal proteins, their transcription levels peaked during the ring and early trophozoite stages,[9] whereas the corresponding proteins were abundant only during the trophozoite and early schizont stages.[36] Thus, the variation in the number of ribosomal proteins may also contribute to changes in protein synthesis rates and in the time delay of the proteomics data compared to the corresponding transcriptomics data throughout the IDC.

Prediction of time-dependent precursor synthesis

In addition to the production of macromolecules, *P. falciparum* synthesizes a number of critical small-molecule precursors which function as cofactors in different biological processes.[16] Figure 4 shows the results of our time-dependent prediction of the precursor synthesis classified into

four general groups. Groups I, II, and III include the metabolites mainly produced during the early (ring stage), middle (trophozoite and early schizont stages), and late (schizont stage) periods of the IDC, respectively, whereas Group IV includes the metabolites for which the production/uptake rate was roughly constant throughout the IDC.

Group I: The early IDC synthesis included pyridoxal 5-phosphate, the active form of vitamin B6, and NAD, a ubiquitous redox intermediate. The timing of these syntheses suggests that they play a key biological role during this period, although both molecules are known to be important in a variety of biological processes. Pyridoxal 5-phosphate is a cofactor for more than 100 known enzymes,[37] including 13 *P. falciparum* enzymes suggested as potential druggable targets by Kronenberger et al.[38] Given the predicted high production levels of pyridoxal 5-phosphate in the ring stage, we examined the stage-specific expressions of 11 out of the 13 enzymes for which the related gene expression data were available.[9] We could categorize all enzymes into stage-specific activities, except for lysine decarboxylase (PFD0285c), which peaked during 26-36 hours, *i.e.*, overlapping with the late trophozoite and early schizont stages. Thus, we found one, five, and four enzymes linked to the ring, trophozoite, and schizont stage, respectively. Supplemental Table S1 provides information for each enzyme. In particular, we could stage-match the production of pyridoxal 5-phosphate with a cysteine desulfurase (encoded by MAL7P1.150 [IscS]) at the late ring stage. This suggests an initial functional link between pyridoxal 5-phosphate and the iron-sulfur complex in the mitochondria, which potentially affects electron transfer, catalysis, and regulatory processes.[39]

Group II: Concomitant with the RNA and protein syntheses during the trophozoite and early schizont stages, a number of precursors were also preferentially produced during this time, suggesting a role for them in protein production or other activities during the mid-IDC. For

example, *de-novo*-synthesized heme was demonstrated to be a requirement for optimal protein syntheses.[40, 41] Figure 4 also shows that the polyamines (spermidine and putrescine) were predicted to achieve peak production levels during mid-IDC, yet they are involved in a wide range of functions, including the stabilization of DNA, RNA, and proteins.[42] However, our prediction of their preferential synthesis during mid-IDC is supported by the experimental observation that inhibition of polyamine syntheses had no visible effect during the ring stage but caused morphological growth arrest in the trophozoite stage.[42] Finally, our prediction of the time-dependent synthesis of 10-formyltetrahydrofolate coincided with protein synthesis, tied to its function of donating its formyl moiety to produce formylmethionine, the initiator of mRNA translation.[43]

Group III: Figure 4 shows that the peak production of the precursors included in this group roughly coincided with the enhanced production of DNA and phospholipids during the schizont stage. In this group, the predicted synthesis of ubiquinone was confirmed by noting that the highest experimental ubiquinone concentration occurred in the schizont stage.[44]

The observed time dependency among this group of molecules also allowed us to hypothesize about their biological role. For example, S-adenosyl L-methionine involves both the trophozoite-stage polyamine syntheses[42] and DNA methylation, an important step in cell division during the schizont stage.[45] Based on our predictions of an increased production of S-adenosyl L-methionine during the late IDC, we hypothesized that the primary function of this molecule might be involved in DNA methylation. A further example was coenzyme A, whose inclusion in *Group III* suggests that its function is related to schizont-stage activities. Coenzyme A is reported to be necessary for DNA replication and cell division in fission yeast,[46] but its hypothesized role in DNA replication of *P. falciparum* remains to be confirmed. Conversely,

coenzyme A biosynthesis is a valid antimicrobial drug target and, based on its essentiality for *P. falciparum* in red blood cells and different biochemical characteristics from its counterpart in humans,[47] may be considered as a potential anti-malarial drug-development target.

Group IV: This group includes molecules for which we predicted roughly constant production/uptake rates throughout the IDC. Because the production/uptake rates were driven by altered gene transcription profiles, our model did not capture other time- and stage-specific processes. Similarly, any lack of a direct gene-reaction relationship in our model would prevent us from modeling such a reaction. For example, in the network model, riboflavin was directly taken up by the parasite and used for biomass production.[16] However, as no gene was associated with this process, we did not model any time-dependent variation in riboflavin accumulation.

Prediction of energy production and consumption

It has long been proposed that *P. falciparum* obtains energy by anaerobically metabolizing glucose into lactate through the glycolysis pathway.[4] However, it remains unclear whether other metabolic pathways contribute appreciably to energy production and what biological processes use the produced energy. Using our model, we calculated the time-dependent ATP production from glycolysis, the production and consumption by other metabolic pathways, and the consumption by non-metabolic activities (*i.e.*, conversion of ATP to ADP). Figure 5 shows that compared to glycolysis, other metabolic pathways produced negligible amounts of ATP, and that these non-glycolytic pathways consumed much less ATP than non-metabolic activities. In addition, these ATP production and consumption pathways reached their highest levels during the trophozoite and late schizont stages, indicating that *P. falciparum* relied almost exclusively on the glycolysis pathway to generate energy during the IDC, most of which was used for non-metabolic activities, such as RNA production, protein syntheses, and cellular growth.

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Prediction and analysis of metabolite fluxes in the TCA cycle

P. falciparum possesses a mitochondrial compartment and encodes all of the enzymes in the TCA cycle,[5] indicating the presence of a fully functional TCA cycle. However, the function of the cycle still remains unclear, because *P. falciparum* does not generate the bulk of its energy by processing glucose through the TCA cycle.[5, 6] Instead, *P. falciparum* metabolizes glucose into lactate, which is subsequently secreted.[5, 6] Here, we predicted the time-series fluxes for all the reactions in *P. falciparum* during the IDC to clarify the metabolic role of the TCA cycle.

The predicted fluxes through the TCA reactions shared some common features. For example, Figure 6A shows that all reaction fluxes were in the oxidative direction, which was consistent with isotope profiling experiments.[5, 6] In addition, Figures 6B and 6C show that the predicted flux profiles for all the reactions in the TCA cycle peaked during the schizont stage (30–48 hours), indicating the importance of these reactions for this stage of the IDC.

The TCA reaction fluxes also showed different time-dependent profiles, suggesting that we could identify individual segments with different biological functions. Figure 6B shows that the fluxes through the reactions catalyzed by citrate synthase (CS), aconitate hydratase (ACONT), and isocitrate dehydrogenase (ICDH) were much lower than other reactions. This segment of the TCA cycle involves the detoxification of superoxide leaked from the mitochondrial respiration chain where the leaked superoxide is converted by dismutase into hydrogen peroxide, which is subsequently reduced to water.[48] This reduction of hydrogen peroxide oxidizes NADPH into NADP, which, in turn, is reduced by ICDH.[48, 49] Furthermore, the similarity in the flux profiles of α-ketoglutarate dehydrogenase, succinate-coenzyme A ligase, and succinate dehydrogenase (AKGDH, SUCOAS, and SUCD in Figure 6, respectively) enabled us to group these reactions into one segment, which we linked to oxidation of 2-oxoglutarate generated from glutamate and

glutamine.[5] This oxidation is continued by fumarase (FUM in Figure 6),[5] which also metabolizes fumarate from the purine pathway.[50] Finally, we predicted that part of the produced L-malate was transported out of mitochondria while the remaining L-malate was converted in the TCA cycle into oxaloacetate by L-malate dehydrogenase (MDH in Figure 6).

Conclusion

The IDC represents the major growth-through-replication phase of malaria parasites in humans, and its roughly 48-hour cycle corresponds to the well-known recurring host-fever attacks. During this phase, the merozoite undergoes three distinct transformations that allow it to absorb nutrients from the infected blood cell and undergoes four or five cycles of replication to produce 16 or 32 new merozoites, respectively. To examine how *P. falciparum* executes and adjusts its metabolism during this replication phase, we created a genome-scale metabolic network model that predicted metabolic fluxes, the biomass production rate, and the production of each biomass metabolite at each hour during the IDC. This work builds on and extends the capabilities of the existing metabolic models of blood-stage *P. falciparum*.[16, 17] In particular, we have extended the metabolic network models to *1*) predict time-dependent biomass production and *2*) allow for a time-variable biomass function, so that we could correctly model and capture the stage-specific synthesis of each biomass metabolite and capture stage-specific growth phenotypes. Our results indicate that the model provides detailed mechanistic links between transcriptomics, metabolism, and phenotypically distinct organism development. This lays the foundation for the future interpretation and exploration of *P. falciparum* metabolism under a wide variety of stress conditions.

Experimental

Metabolic network of *P. falciparum*

From the two recently developed metabolic networks of *P. falciparum*,[16, 17] we started our model development using the *i*TH366 gene-protein-reaction formulations, stoichiometric representations of all reactions, and biomass objective functions.[16] Although equivalent in scope, the *i*TH366 network differentiates the reactions catalyzed by isozymes from those by multiple-unit protein complexes,[16] while the PlasmoNet network assumes that all of these reactions are catalyzed by isozymes.[17] The more extensive gene-protein-reaction formulation theoretically allows for a finer distinction when integrating the gene expression data with the individual metabolic reactions.

Modifications

We made modifications to the existing *i*TH366 set of reactions by changing the compartment of glutathione reductase from mitochondria to cytosol,[48, 51] adding the genes encoding phosphoethanolamine methyltransferase,[52] and the enzymes associated with the hemoglobin degradation.[30, 53] We also added a new set of reactions, including serine decarboxylase,[54] methylenetetrahydrofolate reductase,[55] cardiolipin synthase,[56] the polymerization of hemoglobin-generated heme, the degradation of the heme by glutathione, mitochondrial thioredoxin reductase, and mitochondrial peroxiredoxin.[48, 51, 57] We only allowed uptake and secretion for the set of metabolites that were found to be transported by *P. falciparum* during the IDC (see the complete list in Supplemental Table S2 in Electronic Supplementary Information).[8, 30, 48, 58-70] We also assumed free secretion of CO_2 and free transports of H_2O and H^+, and we added sulfate uptake and urea secretion to avoid a zero biomass production rate.

We redefined the original biomass objective function into a new set of biomass functions (denoted by *B*) to allow for monitoring biomass metabolite production at different times during the IDC. Thus, the new formulation *B* of the biomass function describes the biological activities necessary for the biomass production in *P. falciparum*, through the following equations:

$$AA_a + tRNA_a \rightarrow AA\text{-}tRNA_a,\ a = 1,\ldots,20 \quad (1)$$

$$c_{AA1}AA\text{-}tRNA_1 + \ldots + c_{AA20}AA\text{-}tRNA_{20} \rightarrow c_{AA1}tRNA_1 + \ldots + c_{AA20}tRNA_{20} + Protein \quad (2)$$

$$c_{ATP}ATP + c_{CTP}CTP + c_{GTP}GTP + c_{UTP}UTP \rightarrow RNA + (c_{ATP}+c_{CTP}+c_{GTP}+c_{UTP})Diphosphate \quad (3)$$

$$c_{dATP}dATP + c_{dCTP}dCTP + c_{dGTP}dGTP + c_{dTTP}dTTP \rightarrow DNA + (c_{dATP}+c_{dCTP}+c_{dGTP}+c_{dTTP})Diphosphate \quad (4)$$

$$c_{PC}PC + c_{PE}PE + c_{PS}PS + c_{PI}PI + c_{SPH}SPH + c_{CL}CL \rightarrow \phi \quad (5)$$

$$c_iM_i \rightarrow \phi \text{ (for each metabolite } i \text{ that was a polyamine, a cofactor, or an inorganic ion)} \quad (6)$$

$$c_{energy}\ H_2O + c_{energy}\ ATP \rightarrow c_{energy}\ ADP + c_{energy}\ Phosphate + c_{energy}\ H^+ \quad (7)$$

Equations 1 and 2 represent the addition of amino acids for protein syntheses, where AA denotes an amino acid, *a* represents its index, $tRNA_a$ denotes tRNA used to transport amino acid *a*, AA-$tRNA_a$ indicates the complex of amino acid *a* and its corresponding tRNA, and c_{AAa} is the corresponding stoichiometric coefficient for amino acid *a*. Similarly, Equations 3 and 4 represent the use of nucleotides for the syntheses of RNA and DNA, respectively. Equation 5 accounts for the addition of the phospholipids phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, sphingomyelin, and cardiolipin (denoted as PC, PE, PS, PI, SPH, and CL in Equation 5, respectively), which are all necessary components of cellular membranes.[34, 71] Equations 6 and 7 denote the utilization of other metabolites (M_i) and energy consumption for the biomass production, respectively. In the above equations, *c* represents the coefficient of a metabolite in the original biomass objective function. One should note that

carbohydrates are not part of the original biomass objective function, primarily because there are no experimental data available. This omission should not be a major limitation for the metabolic network formulation for blood-stage malaria, as the parasites do not store energy reserves during this stage, but instead rely heavily on glucose uptake (G. Plata, personal communication). Once experimental data are available for other metabolites, such as carbohydrates, one could add their compositions into the biomass objective function.

We have provided the constructed model representing *P. falciparum* metabolism during the IDC (in MATLAB format) in the Supplemental Protocol S1 in Electronic Supplementary Information.

Gene expression data for *P. falciparum* during the IDC

To capture time-dependent metabolism as accurately as possible, we used the time-series gene expression data at the highest time resolution, collected hourly from synchronized populations of *P. falciparum* HB3 during the IDC.[9, 10] This data set is measured using two-channel microarrays in which the treatment channel contained the sample at one time point, while the control channel contained a mixed pool of samples from all time points.[9, 10] Basing our model construction and calculations on data for the HB3 strain[9] enabled us to validate our results using experimentally measured synthesis rates of DNA, RNA, and protein that are only available for this strain.[31]

Calculation of metabolic fluxes in *P. falciparum* at each hour during the IDC

Figure 1 shows the overall scheme for calculating metabolic fluxes of *P. falciparum* at each time point of the IDC. We used a stepwise approach to construct a set of time-series reaction fluxes that fluctuated around their nominal values based on corresponding time-series gene expression data.

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Step I. We calculated a set of nominal fluxes that were equal to the most parsimonious metabolite flow through the metabolic network, based on the assumption of the most efficient utilization of nutrients.[72] This was done by solving two minimization problems, the first of which was the minimization of the overall nutrient uptake rate,

$$\min \sum_{j \in T} v_j / \alpha_j \qquad (8)$$

$$\text{s.t.} \quad \mathbf{S} \cdot \mathbf{v} = \mathbf{0}$$

$$\mathbf{lb} \le \mathbf{v} \le \mathbf{ub}$$

$$v_j = \mu^N \qquad \text{for each } j \in B,$$

where **v** denotes the flux vector with element v_j indicating the flux through reaction *j* in units of mmol/h/gDW, or millimole per hour per gram dry weight of the original merozoite (the asexual form of *P. falciparum* capable of infecting erythrocytes); *T* represents the set of nutrient uptake reactions; α_j denotes the human serum concentration[73] of the nutrient taken up by reaction *j;* **S** indicates the stoichiometric matrix; **lb** and **ub** denote the lower and upper bounds of the fluxes, respectively; and μ^N represents the average growth rate of *P. falciparum* during the IDC.

The coefficients ($1/\alpha_j$) in the objective function ensure a preference to uptake nutrients associated with high concentrations. We set the average growth rate based on the observation that during the IDC, the organism undergoes an average of four to five cell divisions.[30] Because the total number of divisions is not controlled under experimental conditions, the total mass at the end of the IDC is ½×(2^4 + 2^5) or 24 times the initial merozoite mass. This sets the value for μ^N to (24-1)/48 or 0.48 gram new biomass per hour per gram dry weight of the original merozoite weight (g/h/gDW). By setting the constraints $v_j = \mu^N$ (for $j \in B$), we fixed a certain flux through each biomass function, which ensures that the mass of each biomass component at the end of the IDC

is 24 times its initial mass, corresponding to the observation that, on average, one merozoite generates 24 copies of itself during the IDC.

Given minimum nutrient uptake, we subsequently minimized the sum of intracellular fluxes by solving the following problem:

$$\min \sum_{j \notin T} |v_j| \tag{9}$$

$$\text{s.t.} \quad \mathbf{S} \cdot \mathbf{v} = \mathbf{0}$$

$$\mathbf{lb} \le \mathbf{v} \le \mathbf{ub}$$

$$v_j = \mu^N \quad \text{for each reaction } j \in B,$$

$$v_j \le v_j^* \quad \text{for each reaction } i \in T,$$

where v_j^* represents the minimum nutrient uptake rates obtained in Equation (8), and the inequalities $v_j \le v_j^*$ (for $j \in T$) indicate that the nutrient uptakes are constrained by the corresponding minimum values v_j^*.

After the two minimizations, we obtained a set of nominal fluxes $\mathbf{v}^N$ representative of typical physiological fluxes through the reactions. For example, fluxes through the glycolysis pathway (including lactate production and secretion) were among the largest, consistent with the well-established fermentative glucose utilization by the parasite.[4]

Step II. Next, we processed the time-series gene expression data of *P. falciparum* during the IDC[9] to obtain an expression level r_j^t for each reaction *j* at each time point *t*. After first smoothing the original gene expression data using the locally weighted scatterplot smoothing (LOWESS) method with a span parameter of 30%,[9] we subsequently mapped the resultant gene expression data to reactions by established methods.[21, 22] If a single enzyme catalyzed a reaction,

we assigned the corresponding expression level of that gene to that reaction. If a reaction was catalyzed by multiple isozymes, we assigned their maximum gene expression level to that reaction. If a reaction was catalyzed by a protein complex composed of multiple subunits, we assigned their minimum gene expression level to that reaction.

Time shift between gene transcription and protein translation

In this article, we have discussed metabolism in terms of when enzymatic reactions occur and when the model predicts synthesis of biomaterials. As the available experimental data of gene transcription levels do not directly give us protein levels, we time-shifted the available transcription levels by a time difference equal to the difference between peak expression levels and peak protein abundance levels.[36] For reactions associated with directly measured time-dependent protein abundances, the time delays were taken as the differences between the peak times of protein abundances and those of the corresponding mRNA levels. If these reactions were part of the same metabolic pathway and proteomic data were unavailable for some enzymes in the same pathway, the time delay for the reactions with missing data was assigned based on the average peak-time differences of the reactions with experimentally available data. For all remaining reactions, the time delays were 11 hours, corresponding to the median of peak time differences between all the measured protein abundances and the corresponding transcriptomics data.[36]

Linear normalization of expression data

Finally, we normalized the expression data linearly, so that the mean and minimum of each reaction's time-series expression values were one and zero, respectively. This assumes that the minimum expression value corresponded to complete inactivation of the reaction, and the mean value corresponded to the nominal flux through the reaction.

Step III. Given the nominal fluxes v_j^N (from *Step I*) and the expression level r_j^t for each reaction j at each time point t (from *Step II*), we calculated the reaction fluxes at each time point by first minimizing the difference J^t between the reaction fluxes and the product of their nominal fluxes and expression values as follows:

$$\min \; J^t = \sum_{j \in G} \left| v_j^t - r_j^t \cdot v_j^N \right| \qquad (10)$$

$$\text{s.t.} \quad \mathbf{S} \cdot \mathbf{v}^t = \mathbf{0}$$

$$\mathbf{lb} \le \mathbf{v}^t \le \mathbf{ub},$$

where v_j^t represents the flux through reaction j at time point t, and G represents the set of intracellular irreversible reactions that can be associated with gene expression data. Because the expression data of transport (uptake or secretion) reactions are mapped to transport fluxes for molecules involved in both metabolic (subject to constraints in our model) and non-metabolic processes (not subject to constraints), G does not include these reactions. Examples of transport reactions for non-metabolic processes include water for regulation of osmotic pressure[74, 75] and H^+ transport for maintaining intracellular pH levels.[76, 77] G also excludes reactions that can carry fluxes in both directions, because we cannot unambiguously determine the directions of their fluxes. In total, we excluded eight such reactions among the 317 intracellular reactions that were associated with gene expression data and not in dead-end pathways.

Given that the problem formulated in Equation 10 may have multiple solutions, we selected the solution closest to the nominal flux distribution, based on the assumption that reaction fluxes minimally deviated from their nominal values. We did this by solving the following additional optimization problem:

$$\min \sum_j \left| v_j^t - v_j^N \right| \tag{11}$$

$$\text{s.t.} \quad \mathbf{S} \cdot \mathbf{v}^t = \mathbf{0}$$

$$\mathbf{lb} \le \mathbf{v}^t \le \mathbf{ub}$$

$$\sum_{j \in G} \left| v_j^t - r_j^t \cdot v_i^N \right| \le J_t^* \ ,$$

where J_t^* is the optimal value for the objective function from the previous optimization problem defined by Equation 10, and the last constraint ensures that this solution is one of the optimal solutions for Equation 10.

After solving Equations 10 and 11, we obtained metabolic flux distributions for all time points during the IDC. These time-series fluxes varied around their nominal values in time-dependent patterns similar to those of the corresponding expression data. If no gene expression data were available, we assigned the reaction the time-independent value of the nominal flux distribution.

The intent of the above approach was to be able to use the relative gene expression level as a semi-quantitative indicator that the cell is readying itself to transcribe the produced mRNA into proteins to execute or affect some biological/metabolic process. Thus, the metabolic model captures time-dependent aspects of when the gene transcripts became available, as well as the relative magnitudes of the expression levels. We further accounted for the mRNA/protein mismatch by using the experimentally measured time difference between when a gene was transcribed and when the synthesized proteins appeared in the organism. The use of relative expression levels around the average levels allows the optimization procedure flexibility in determining metabolic fluxes consistent with, but not equal to, gene expression levels.

Step IV. We determined the overall biomass production level μ^t at each time point t from

$$\mu^t = \frac{\sum_{j \in B} v_j^t w_j}{\sum_{j \in B} w_j}, \quad (12)$$

where w_j indicates the biomass fraction of the metabolite(s) associated with the biomass function j. We defined w_j as

$$w_j = \sum_i \frac{c_{ij} W_i}{1{,}000}, \quad (13)$$

where c_{ij} represents the coefficient of metabolite i in biomass function j, W_i denotes the molecular weight of the metabolite, and the factor 1,000 converts mol into mmol.

Simulation environment

We constructed the model and ran the simulations in MATLAB (2012a, MathWorks, Natick, MA) using the COBRA toolbox.[78] The metabolic model of *P. falciparum* during the IDC (in MATLAB format) is provided as Supplemental Protocol S1 in Electronic Supplementary Information.

Acknowledgments

The authors thank Dr. F. Vital-Lopez and Dr. H.-S. Song for valuable discussions. The authors were supported by the U.S. Army Medical Research and Materiel Command (Ft. Detrick, MD), as part of the U.S. Army's Network Science Initiative. The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

Electronic Supplementary Information

Electronic Supplementary Information includes Supplemental Table S1 for the list of pyridoxal 5-phosphate-dependent enzymes in *P. falciparum*, Supplemental Table S2 for the list of metabolite uptakes and secretions that were allowed by the metabolic network of *P. falciparum*, and Supplemental Protocol S1 for the MATLAB-formatted models of metabolism for *P. falciparum* during the IDC.

References

1. World Health Organization, *World Malaria Report*, 2013.
2. B. M. Greenwood, D. A. Fidock, D. E. Kyle, S. H. Kappe, P. L. Alonso, F. H. Collins and P. E. Duffy, *J Clin Invest*, 2008, 118, 1266-1276.
3. R. Tuteja, *FEBS J*, 2007, 274, 4670-4679.
4. K. L. Olszewski and M. Llinas, *Mol Biochem Parasitol*, 2011, 175, 95-103.
5. J. I. Macrae, M. W. Dixon, M. K. Dearnley, H. H. Chua, J. M. Chambers, S. Kenny, I. Bottova, L. Tilley and M. J. McConville, *BMC Biology*, 2013, 11, 67.
6. S. A. Cobbold, A. M. Vaughan, I. A. Lewis, H. J. Painter, N. Camargo, D. H. Perlman, M. Fishbaugher, J. Healer, A. F. Cowman, S. H. Kappe and M. Llinas, *J Biol Chem*, 2013, 288, 36338-36350.
7. K. G. Le Roch, Y. Zhou, P. L. Blair, M. Grainger, J. K. Moch, J. D. Haynes, P. De La Vega, A. A. Holder, S. Batalov, D. J. Carucci and E. A. Winzeler, *Science*, 2003, 301, 1503-1508.

8. K. L. Olszewski, J. M. Morrisey, D. Wilinski, J. M. Burns, A. B. Vaidya, J. D. Rabinowitz and M. Llinas, *Cell Host Microbe*, 2009, 5, 191-199.
9. Z. Bozdech, M. Llinas, B. L. Pulliam, E. D. Wong, J. Zhu and J. L. DeRisi, *PLOS Biology*, 2003, 1, E5.
10. M. Llinas, Z. Bozdech, E. D. Wong, A. T. Adai and J. L. DeRisi, *Nucleic Acids Res*, 2006, 34, 1166-1173.
11. Z. Bozdech and H. Ginsburg, *Malar J*, 2004, 3, 23.
12. Z. Bozdech and H. Ginsburg, *Malar J*, 2005, 4, 17.
13. A. M. Feist, M. J. Herrgard, I. Thiele, J. L. Reed and B. O. Palsson, *Nat Rev Microbiol*, 2009, 7, 129-143.
14. I. Yeh, T. Hanekamp, S. Tsoka, P. D. Karp and R. B. Altman, *Genome Res*, 2004, 14, 917-924.
15. S. Fatumo, K. Plaimas, J. P. Mallm, G. Schramm, E. Adebiyi, M. Oswald, R. Eils and R. Konig, *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 2009, 9, 351-358.
16. G. Plata, T. L. Hsiao, K. L. Olszewski, M. Llinas and D. Vitkup, *Mol Syst Biol*, 2010, 6, 408.
17. C. Huthmacher, A. Hoppe, S. Bulik and H. G. Holzhutter, *BMC Syst Biol*, 2010, 4, 120.
18. S. Bazzani, A. Hoppe and H. G. Holzhutter, *BMC Syst Biol*, 2012, 6, 118.
19. K. Jensen, D. Plichta, G. Panagiotou and I. Kouskoumvekaki, *Mol Biosyst*, 2012, 8, 1678-1685.
20. J. D. Orth, I. Thiele and B. O. Palsson, *Nat Biotechnol*, 2010, 28, 245-248.
21. S. A. Becker and B. O. Palsson, *PLOS Comput Biol*, 2008, 4, e1000082.
22. T. Shlomi, M. N. Cabili, M. J. Herrgard, B. O. Palsson and E. Ruppin, *Nat Biotechnol*, 2008, 26, 1003-1010.
23. C. Colijn, A. Brandes, J. Zucker, D. S. Lun, B. Weiner, M. R. Farhat, T. Y. Cheng, D. B. Moody, M. Murray and J. E. Galagan, *PLOS Comput Biol*, 2009, 5, e1000489.
24. X. Fang, A. Wallqvist and J. Reifman, *PLOS Comput Biol*, 2012, 8, e1002688.
25. F. G. Vital-Lopez, A. Wallqvist and J. Reifman, *BMC Syst Biol*, 2013, 7, 63.
26. P. A. Jensen and J. A. Papin, *Bioinformatics*, 2011, 27, 541-547.
27. N. Topfer, S. Jozefczuk and Z. Nikoloski, *BMC Syst Biol*, 2012, 6, 148.
28. T. Spielmann, P. L. Hawthorne, M. W. Dixon, M. Hannemann, K. Klotz, D. J. Kemp, N. Klonis, L. Tilley, K. R. Trenholme and D. L. Gardiner, *Molecular Biology of the Cell*, 2006, 17, 3613-3624.
29. L. H. Bannister, J. M. Hopkins, R. E. Fowler, S. Krishna and G. H. Mitchell, *Parasitology Today*, 2000, 16, 427-433.
30. S. E. Francis, D. J. Sullivan, Jr. and D. E. Goldberg, *Annu Rev Microbiol*, 1997, 51, 97-123.
31. C. A. Gritzmacher and R. T. Reese, *J Bacteriol*, 1984, 160, 1165-1167.
32. H. J. Vial, M. J. Thuet and J. R. Philippot, *The Journal of Protozoology*, 1982, 29, 258-263.
33. S. Kabanova, P. Kleinbongard, J. Volkmer, B. Andree, M. Kelm and T. W. Jax, *Int J Med Sci*, 2009, 6, 156-159.
34. L. L. Hsiao, R. J. Howard, M. Aikawa and T. F. Taraschi, *Biochem J*, 1991, 274 (Pt 1), 121-132.

35. E. M. Bunnik, D. W. Chung, M. Hamilton, N. Ponts, A. Saraf, J. Prudhomme, L. Florens and K. G. Le Roch, *Genome Biol*, 2013, 14, R128.
36. B. J. Foth, N. Zhang, B. K. Chaal, S. K. Sze, P. R. Preiser and Z. Bozdech, *Mol Cell Proteomics*, 2011, 10, M110 006411.
37. C. Wrenger, M. L. Eschbach, I. B. Muller, D. Warnecke and R. D. Walter, *J Biol Chem*, 2005, 280, 5242-5248.
38. T. Kronenberger, J. Lindner, K. A. Meissner, F. M. Zimbres, M. A. Coronado, F. M. Sauer, I. Schettert and C. Wrenger, *BioMed Research International*, 2014, 2014, 108516.
39. T. A. Dellibovi-Ragheb, J. E. Gisselberg and S. T. Prigge, *PLOS Pathog*, 2013, 9, e1003227.
40. N. Surolia and G. Padmanaban, *Proc Natl Acad Sci U S A*, 1991, 88, 4786-4790.
41. Z. Q. Bonday, S. Taketani, P. D. Gupta and G. Padmanaban, *J Biol Chem*, 1997, 272, 21839-21846.
42. A. C. van Brummelen, K. L. Olszewski, D. Wilinski, M. Llinas, A. I. Louw and L. M. Birkholtz, *J Biol Chem*, 2009, 284, 4635-4646.
43. A. Nzila, S. A. Ward, K. Marsh, P. F. Sims and J. E. Hyde, *Trends in Parasitology*, 2005, 21, 292-298.
44. M. B. Cassera, E. F. Merino, V. J. Peres, E. A. Kimura, G. Wunderlich and A. M. Katzin, *Memorias do Instituto Oswaldo Cruz*, 2007, 102, 377-383.
45. T. H. Bestor and G. L. Verdine, *Current Opinion in Cell Biology*, 1994, 6, 380-389.
46. T. Nakamura, T. Pluskal, Y. Nakaseko and M. Yanagida, *Open Biology*, 2012, 2, 120117.
47. C. Spry, K. Kirk and K. J. Saliba, *FEMS Microbiology Reviews*, 2008, 32, 56-106.
48. S. Muller, *Mol Microbiol*, 2004, 53, 1291-1305.
49. C. Wrenger and S. Muller, *European Journal of Biochemistry/FEBS*, 2003, 270, 1775-1783.
50. V. Bulusu, V. Jayaraman and H. Balaram, *J Biol Chem*, 2011, 286, 9236-9245.
51. S. Kehr, N. Sturm, S. Rahlfs, J. M. Przyborski and K. Becker, *PLOS Pathog*, 2010, 6, e1001242.
52. W. H. Witola, K. El Bissati, G. Pessi, C. Xie, P. D. Roepe and C. B. Mamoun, *J Biol Chem*, 2008, 283, 27636-27643.
53. S. Subramanian, M. Hardt, Y. Choe, R. K. Niles, E. B. Johansen, J. Legac, J. Gut, I. D. Kerr, C. S. Craik and P. J. Rosenthal, *PLOS ONE*, 2009, 4, e5156.
54. N. Elabbadi, M. L. Ancelin and H. J. Vial, *Biochem J*, 1997, 324 (Pt 2), 435-445.
55. W. Asawamahasakda and Y. Yuthavong, *Parasitology*, 1993, 107 (Pt 1), 1-10.
56. C. Aurrecoechea, J. Brestelli, B. P. Brunk, J. Dommer, S. Fischer, B. Gajria, X. Gao, A. Gingle, G. Grant, O. S. Harb, M. Heiges, F. Innamorato, J. Iodice, J. C. Kissinger, E. Kraemer, W. Li, J. A. Miller, V. Nayak, C. Pennington, D. F. Pinney, D. S. Roos, C. Ross, C. J. Stoeckert, Jr., C. Treatman and H. Wang, *Nucleic Acids Res*, 2009, 37, D539-543.
57. J. Sun and B. L. Trumpower, *Archives of Biochemistry and Biophysics*, 2003, 419, 198-206.
58. E. Roth, Jr., *Blood Cells*, 1990, 16, 453-460.
59. S. E. Babbitt, L. Altenhofen, S. A. Cobbold, E. S. Istvan, C. Fennell, C. Doerig, M. Llinas and D. E. Goldberg, *Proc Natl Acad Sci U S A*, 2012, 109, E3278-3287.
60. M. B. Cassera, Y. Zhang, K. Z. Hazleton and V. L. Schramm, *Current Topics in Medicinal Chemistry*, 2011, 11, 2103-2115.

61. A. M. Lehane, K. J. Saliba, R. J. Allen and K. Kirk, *Biochem Biophys Res Commun*, 2004, 320, 311-317.
62. K. J. Saliba, R. E. Martin, A. Broer, R. I. Henry, C. S. McCarthy, M. J. Downie, R. J. Allen, K. A. Mullin, G. I. McFadden, S. Broer and K. Kirk, *Nature*, 2006, 443, 582-585.
63. K. Kirk and K. J. Saliba, *Current Drug Targets*, 2007, 8, 75-88.
64. P. Wang, Q. Wang, P. F. Sims and J. E. Hyde, *Mol Biochem Parasitol*, 2007, 154, 40-51.
65. S. Pollack and J. Fleming, *Br J Haematol*, 1984, 58, 289-293.
66. A. M. Vaughan, M. T. O'Neill, A. S. Tarun, N. Camargo, T. M. Phuong, A. S. Aly, A. F. Cowman and S. H. Kappe, *Cell Microbiol*, 2009, 11, 506-520.
67. A. D. Murphy, J. E. Doeller, B. Hearn and N. Lang-Unnasch, *Experimental Parasitology*, 1997, 87, 112-120.
68. R. E. Martin, H. Ginsburg and K. Kirk, *Mol Microbiol*, 2009, 74, 519-528.
69. P. Dutta, *The Journal of Protozoology*, 1991, 38, 479-483.
70. C. Wrenger, J. Knockel, R. D. Walter and I. B. Muller, *Brazilian Journal of Medical and Biological Research*, 2008, 41, 82-88.
71. G. A. Thompson, Jr. and Y. Nozawa, *Annu Rev Microbiol*, 1972, 26, 249-278.
72. X. Fang, A. Wallqvist and J. Reifman, *BMC Syst Biol*, 2009, 3, 92.
73. N. Psychogios, D. D. Hau, J. Peng, A. C. Guo, R. Mandal, S. Bouatra, I. Sinelnikov, R. Krishnamurthy, R. Eisner, B. Gautam, N. Young, J. Xia, C. Knox, E. Dong, P. Huang, Z. Hollander, T. L. Pedersen, S. R. Smith, F. Bamforth, R. Greiner, B. McManus, J. W. Newman, T. Goodfriend and D. S. Wishart, *PLOS ONE*, 2011, 6, e16957.
74. M. A. Zanner, W. R. Galey, J. V. Scaletti, J. Brahm and D. L. Vander Jagt, *Mol Biochem Parasitol*, 1990, 40, 269-278.
75. P. Gervais and L. Beney, *Cellular and Molecular Biology*, 2001, 47, 831-839.
76. S. Wunsch, C. P. Sanchez, M. Gekle, L. Grosse-Wortmann, J. Wiesner and M. Lanzer, *The Journal of Cell Biology*, 1998, 140, 335-345.
77. A. Mobasheri, R. Mobasheri, M. J. Francis, E. Trujillo, D. Alvarez de la Rosa and P. Martin-Vasallo, *Histology and Histopathology*, 1998, 13, 893-910.
78. J. Schellenberger, R. Que, R. M. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke and B. O. Palsson, *Nat Protoc*, 2011, 6, 1290-1307.

PAPER www.rsc.org/molecularbiosystems | Molecular BioSystems

Figure Legends

Figure 1: Schematic description of integrating a metabolic network with time-series gene expression data.

We constructed a set of metabolic states over a time course based on time-series gene expression data by altering nominal fluxes obtained from a metabolic network that represent average and typical fluxes through each reaction. The example network contains five metabolites (A–E), two uptake reactions, three enzymatic reactions, and one biomass reaction. In *Step I*, we obtained the set of nominal fluxes that satisfy the mass balance of each metabolite and the typical biomass composition. In *Step II*, we mapped the time-series gene expression data to their corresponding reactions. In *Step III*, given the nominal fluxes and gene expression data, we calculated a set of time-series fluxes. For each reaction, the time-dependent pattern of these fluxes followed that of the corresponding gene expression data, while the average of the fluxes was equal to (or as close as possible to) the nominal flux. In *Step IV*, we determined the time-series biomass production rate, by adding the production rates of all biomass components. We finally constructed time-series metabolic states that showed time-dependent alterations of reaction fluxes, biomass production, and net production of each biomass component.

Figure 2: The predicted overall biomass production rates μ of *Plasmodium falciparum* during the intraerythrocytic developmental cycle.

The whole intraerythrocytic developmental cycle was classified into ring, trophozoite, and schizont stages.[10] The unit of the biomass production rate μ is g/h/gDW, denoting gram biomass per hour per gram dry weight of the *P. falciparum* merozoite at the beginning of the cycle.

Figure 3: Macromolecule syntheses in *Plasmodium falciparum* during the intraerythrocytic developmental cycle.

The synthesized amount of RNA (**A**), Protein (**B**), DNA (**C**), and phospholipids (**D**) during a given time interval was equal to the amount of material synthesized during the interval normalized by the maximum amount produced during the entire cycle. The predicted amounts were compared with the corresponding experimental data for DNA and RNA,[31] and for phospholipids.[32] The horizontal bars indicate the length of the time intervals. The colors of these bars represent the simulation results (blue) and experimental data (green).

Figure 4: The predicted time-dependent production of biomass metabolites of *Plasmodium falciparum*.

The heat map denotes the predicted time-dependent production levels of each biomass metabolite of *P. falciparum*, in which orange, grey, and blue colors represent high, normal, and low production levels, respectively. Based on the time-dependent production, we classified these metabolites into four groups. Groups I, II, and III include the metabolites mainly produced during the early (ring stage), middle (trophozoite and early schizont stages), and late (schizont stage) periods of the intraerythrocytic developmental cycle, respectively, while Group IV includes the metabolites for which the production levels were basically constant throughout the intraerythrocytic developmental cycle. The vertical dashed lines indicate the boundaries between ring, trophozoite, and schizont stages.

Figure 5: Predicted energy production and consumption.

(**A**) Schematic description of energy production and consumption. Energy (in form of ATP) was produced from glycolysis and other metabolic pathways, and consumed by non-glycolytic metabolism and non-metabolic activity. (**B**) Predicted time-dependent ATP production and consumption with respect to metabolic and non-metabolic processes (excluding ATP used for RNA synthesis). The unit of the production or consumption is mmol/h/gDW, denoting millimole per hour per gram dry weight of the *P. falciparum* merozoite at the beginning of the intraerythrocytic developmental cycle.

Figure 6: Predicted metabolic fluxes through the tricarboxylic acid cycle.

(**A**) Reactions in the *Plasmodium falciparum* tricarboxylic acid (TCA) cycle. The arrows represent the directions of the predicted fluxes and different line styles indicated our classification of the TCA cycle into several segments, each of which had a distinct function. (**B**) The predicted time-dependent flux profiles of these reactions during the intraerythrocytic developmental cycle (IDC). (**C**) The rescaled flux profile of aconitate hydratase (ACONT), citrate synthase (CS), and isocitrate dehydrogenase (ICDH) during the IDC. The unit of the fluxes is mmol/h/gDW, denoting millimole per hour per gram dry weight of the *P. falciparum* merozoite at the beginning of the IDC. AKGDH, α-ketoglutarate dehydrogenase; FUM, fumarase; MDH, L-malate dehydrogenase; SUCD, succinate dehydrogenase; and SUCOAS, succinate-coenzyme A ligase.

Figure 1

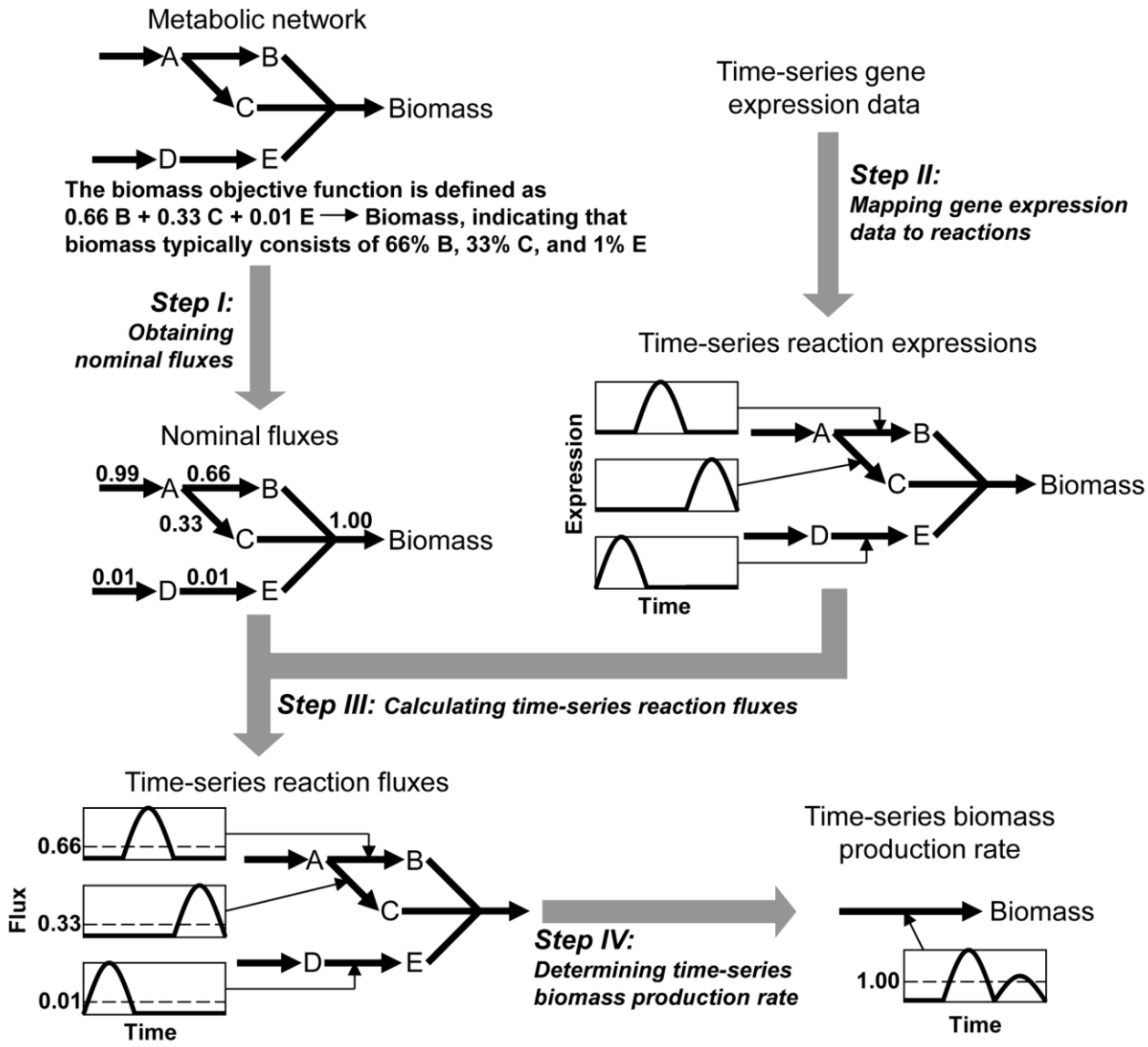


Figure 2

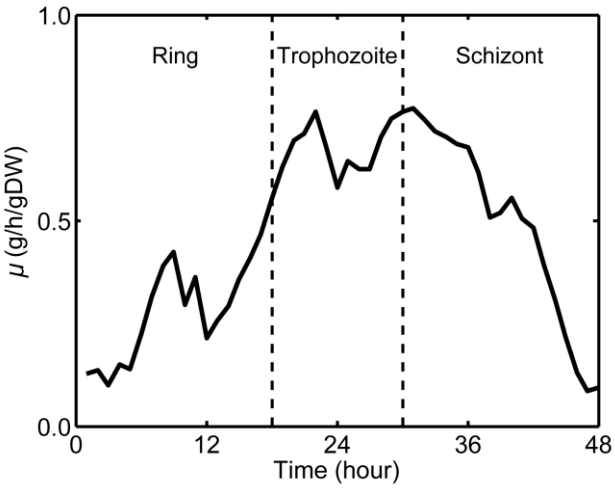


Figure 3

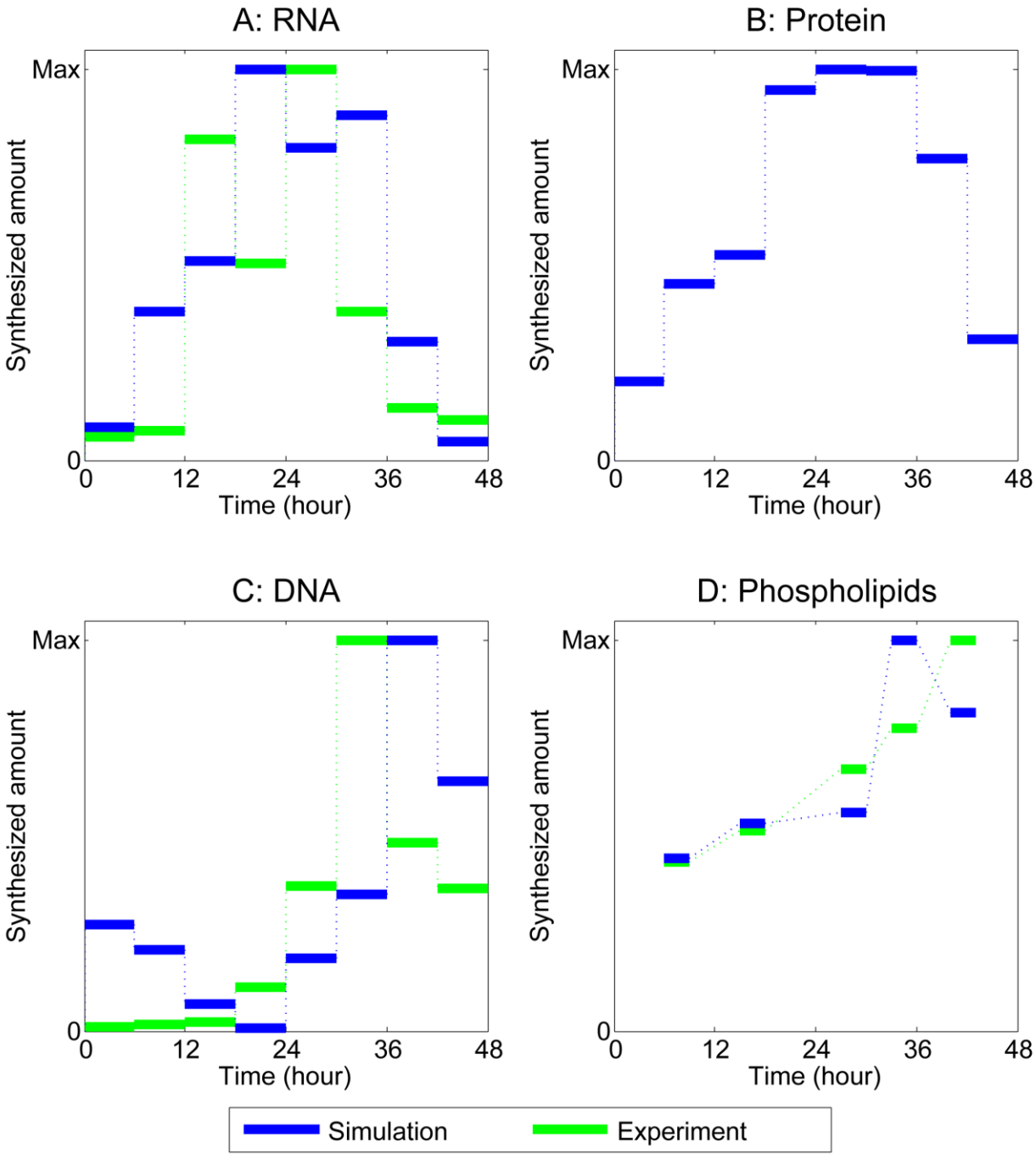


Figure 4

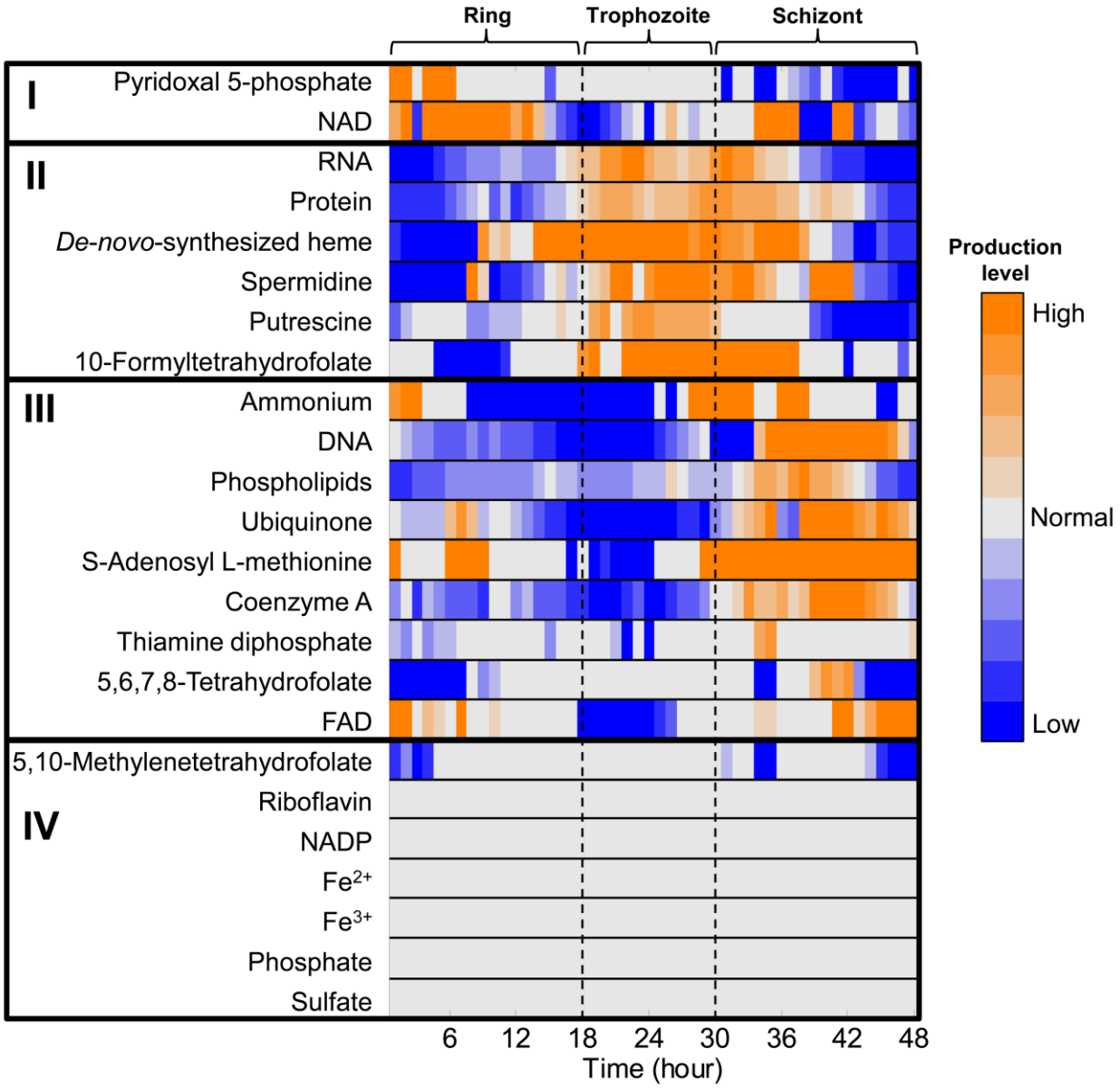


Figure 5

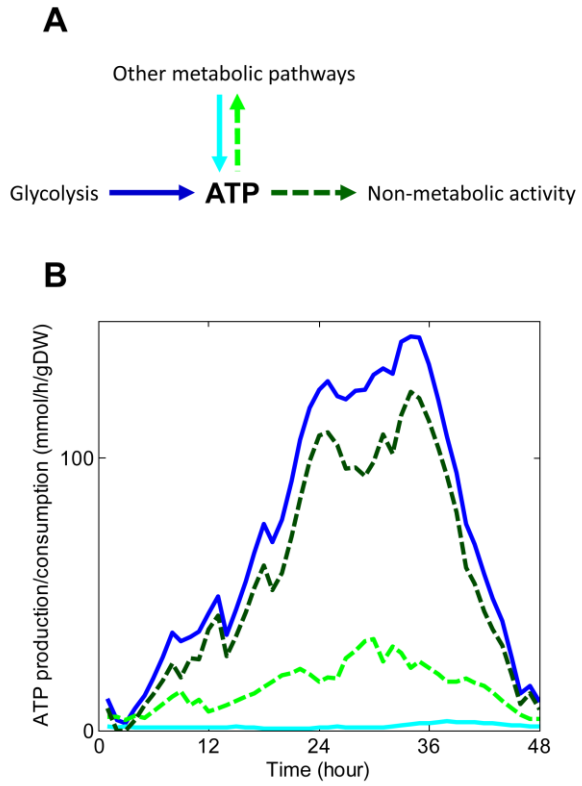


Figure 6