

Bioenergetics-based modeling of *Plasmodium falciparum* metabolism reveals its essential genes, nutritional requirements, and thermodynamic bottlenecks

CHIAPPINO-PEPE, Anush, *et al.*

Reference

CHIAPPINO-PEPE, Anush, *et al.* Bioenergetics-based modeling of *Plasmodium falciparum* metabolism reveals its essential genes, nutritional requirements, and thermodynamic bottlenecks. *PLoS computational biology*, 2017, vol. 13, no. 3, p. e1005397

DOI : 10.1371/journal.pcbi.1005397

Available at:

<http://archive-ouverte.unige.ch/unige:93270>

Disclaimer: layout of this document may differ from the published version.



UNIVERSITÉ
DE GENÈVE

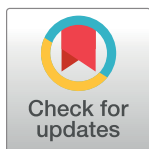
RESEARCH ARTICLE

Bioenergetics-based modeling of *Plasmodium falciparum* metabolism reveals its essential genes, nutritional requirements, and thermodynamic bottlenecks

Anush Chiappino-Pepe¹, Stepan Tymoshenko^{1,2}, Meriç Ataman¹, Dominique Soldati-Favre², Vassily Hatzimanikatis^{1*}

1 Laboratory of Computational Systems Biotechnology, École Polytechnique Fédérale de Lausanne, EPFL, Lausanne, Switzerland, **2** Department of Microbiology and Molecular Medicine, Faculty of Medicine, University of Geneva, CMU, Geneva, Switzerland

* vassily.hatzimanikatis@epfl.ch



OPEN ACCESS

Citation: Chiappino-Pepe A, Tymoshenko S, Ataman M, Soldati-Favre D, Hatzimanikatis V (2017) Bioenergetics-based modeling of *Plasmodium falciparum* metabolism reveals its essential genes, nutritional requirements, and thermodynamic bottlenecks. PLoS Comput Biol 13 (3): e1005397. doi:10.1371/journal.pcbi.1005397

Editor: Costas D. Maranas, The Pennsylvania State University, UNITED STATES

Received: September 7, 2016

Accepted: February 4, 2017

Published: March 23, 2017

Copyright: © 2017 Chiappino-Pepe et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files, except the genome-scale model iPfa, which can be downloaded from <http://lcsb-databases.epfl.ch/GEMs/iPfa>.

Funding: VH, DSF, ACP, ST, and MA are supported by the RTD grants MalarX and MicroScapesX within SystemsX.ch, the Swiss Initiative for Systems Biology evaluated by the Swiss National Science Foundation: <http://www.systemsx.ch/>

Abstract

Novel antimalarial therapies are urgently needed for the fight against drug-resistant parasites. The metabolism of malaria parasites in infected cells is an attractive source of drug targets but is rather complex. Computational methods can handle this complexity and allow integrative analyses of cell metabolism. In this study, we present a genome-scale metabolic model (iPfa) of the deadliest malaria parasite, *Plasmodium falciparum*, and its thermodynamics-based flux analysis (TFA). Using previous absolute concentration data of the intraerythrocytic parasite, we applied TFA to iPfa and predicted up to 63 essential genes and 26 essential pairs of genes. Of the 63 genes, 35 have been experimentally validated and reported in the literature, and 28 have not been experimentally tested and include previously hypothesized or novel predictions of essential metabolic capabilities. Without metabolomics data, four of the genes would have been incorrectly predicted to be non-essential. TFA also indicated that substrate channeling should exist in two metabolic pathways to ensure the thermodynamic feasibility of the flux. Finally, analysis of the metabolic capabilities of *P. falciparum* led to the identification of both the minimal nutritional requirements and the genes that can become indispensable upon substrate inaccessibility. This model provides novel insight into the metabolic needs and capabilities of the malaria parasite and highlights metabolites and pathways that should be measured and characterized to identify potential thermodynamic bottlenecks and substrate channeling. The hypotheses presented seek to guide experimental studies to facilitate a better understanding of the parasite metabolism and the identification of targets for more efficient intervention.

Author summary

Almost half of the world population is at risk of infection by malaria parasites. The rise in drug-resistant parasites requires better understanding and targeting of their metabolism.

index.php?id=276&L=3 and <http://www.systemsx.ch/index.php?id=277&L=3>. VH, ACP, ST and MA are supported by the École Polytechnique Fédérale de Lausanne. DSF is supported by the University of Geneva. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

In this study, we present a genome-scale metabolic reconstruction (iPfa) of the deadliest malaria parasite, *Plasmodium falciparum*, and its thermodynamics-based flux analysis (TFA). Our results support and extend the available experimental evidence on the essential genes and nutritional requirements of this organism. Besides, we identify metabolites that give rise to thermodynamic bottlenecks and suggest substrate channeling. Overall, these results provide novel insight into the metabolism of *P. falciparum* and may guide experimental studies to develop a better characterization of the parasite metabolism and the identification of antimalarial drug targets.

Introduction

Malaria remains a major global health care concern, with almost half of the world population at risk of infection that ultimately results in over half a million deaths each year [1]. Of the five *Plasmodium* species capable of infecting humans, *P. falciparum* is responsible for most malaria-related deaths. The current increase in parasites with resistance to most of the clinically used antimalarial drugs, including artemisinin, renders the treatment of this disease more challenging [1]. The development of more efficient antimalarial treatments is, therefore, a highly pressing need. Because it is essential for cell development, metabolism represents a potential source for identifying novel targets. Computational methods can handle its complexity and thus facilitate the discovery of drug targets (as demonstrated for other pathogens [2, 3]) that are particularly interesting for malaria research.

Genome-scale metabolic models (GEMs) represent an invaluable platform for the integrative analysis of cell metabolism [4]. Currently, two lineages of independently developed GEMs exist for *P. falciparum*, iTH366 [5] and PlasmoNet [6]. Since 2010, these GEMs have been slightly modified to study the metabolism of the parasite in the liver stage [7] or in the blood stages [8]. However, no study, to our knowledge, has developed an independent reconstruction of *P. falciparum* metabolism that updates, among other important features, the functional annotation of the genome, the localization of the enzymes and the definition of the available substrates according to the currently existing data.

The standard approach for analyzing different phenotypes using GEMs is flux balance analysis (FBA) [9, 10]. FBA predictions provide a good understanding of the metabolism at a systems level [11], and they can be further enhanced by integrating context-specific information in the form of constraints. FBA considers mass balance constraints for each metabolite in the metabolic network [9, 10]. Thermodynamics-based flux analysis (TFA) [12, 13] further accounts for thermodynamic constraints and provides a framework for the integration of metabolomics data in GEMs [12–17]. Thermodynamic constraints determine the feasible direction under which the reaction can operate, defined as reaction directionality [14–16].

Thermodynamic bottlenecks arise when alternative metabolic pathways are thermodynamically impeded although, based on network topology, they could have theoretically served a metabolic function. Unfavorable thermodynamics imposed by bulk-phase metabolite concentrations can be circumvented with substrate channeling. Substrate channeling involves the coupling of two or more reactions, and the common intermediate is transferred from the first enzyme to the second without escaping into the bulk phase. Such process has a major effect on the thermodynamics and kinetics of the involved catalytic functions and might determine specific responses to regulatory mechanisms. Techniques such as isotope dilution or enrichment, competing reaction or enzyme buffering have been traditionally used to detect and characterize substrate channeling for an enzyme pair or larger metabolon [18]. Such methods require

systematic approaches, such as TFA, that generate and test thermodynamically consistent hypotheses to ultimately enhance our understanding of metabolism.

The primary goal of this study is to provide new insight into the essential metabolic capabilities and the nutritional requirements of *P. falciparum* that can reveal potential targets in its metabolism for efficient intervention. We also seek to identify metabolites whose intracellular concentrations give rise to thermodynamic bottlenecks and pathways where substrate channeling may exist. These analyses can guide metabolomics and biochemical studies on the metabolism of the parasite. For this purpose, we developed a GEM of *P. falciparum* (iPfa) and performed thermodynamically consistent studies using TFA and integrating the metabolite concentration ranges previously measured in intraerythrocytic *P. falciparum* [19–22]. We present here the TFA results for iPfa that suggest the essential genes, bottleneck metabolites, pathways with substrate channeling, and nutritional requirements of *P. falciparum*.

Results

Reconstruction of iPfa

We combined semi-automated approaches with a manual curation process based on the available literature on *P. falciparum* metabolism to reconstruct iPfa. A detailed description of the steps followed for the reconstruction process is provided in the [S1 Methods](#). This process was performed in agreement with the standard protocol defined for the high-quality component-by-component (bottom-up) reconstruction of GEMs [23].

iPfa includes 325 genes and 670 metabolic reactions localized within five intracellular compartments: the cytosol, the mitochondrion, the apicoplast, the endoplasmic reticulum and the nucleus ([Table 1](#)). Nearly 13% of all metabolic reactions (excluding transport reactions) are orphans, i.e., enzymatic reactions that are not associated with any particular gene in the genome ([S1 Methods](#)). Orphan reactions in iPfa render feasible 24 of the 73 total metabolic tasks. Metabolic tasks are defined here as the production of biomass building blocks [3, 24]. The identified orphan reactions and alternatives might serve as a reference for further biochemical characterization of non-annotated genes in *Plasmodium* species ([S1 Table](#)).

iPfa also accounts for transport reactions: 236 potential uptakes, i.e. transports from the medium (host cell cytosol or blood serum) to the parasite's cytosol, and 155 transports between intracellular compartments ([Table 1](#)). Due to the high uncertainty in the exact type and number of proteins that serve as transporters in *Plasmodium* spp., we included transport reactions based on a set of assumptions about the metabolite transportability ([S1 Methods](#)). These assumptions prevent the introduction of *ad hoc* constraints in iPfa and serve as an upper bound on the metabolite transportability. Such an approach leads to an underestimated number of essential genes. Nearly 98% of the transport reactions in iPfa remain orphan ([Fig 1B](#)).

The metabolic enzymes included in iPfa are distributed in six classes according to the E.C. identifiers. The majority of the metabolic enzymes in iPfa are transferases (43%), followed by oxidoreductases (24%), ligases (12%), hydrolases (12%), lyases (7%) and isomerases (2%) ([Fig 1A](#)). Regarding the metabolic subsystems, a significant percentage of metabolic reactions in iPfa can be classified as part of lipid metabolism (31%), followed by a broad and diverse group involving reactions related to the biosynthesis of cofactors, vitamins and redox molecules (21%). A similar number of metabolic reactions are involved in amino acid metabolism (17%), nucleotide metabolism (16%) and carbohydrate metabolism (15%) ([Fig 1B](#)).

The GEM iPfa presents an unbiased bottom-up reconstruction and an updated database of *P. falciparum* metabolism. For example, based on recent findings [25], the pyruvate dehydrogenase activity of the mitochondrial BCKDH complex was included in iPfa, whereas it was absent in the previous GEMs of the malaria parasite, iTH366 [5], PlasmoNet [6] and their modifications [7,

Table 1. Description of iPfa.

Metabolites		1258
Intracellular metabolites		1017
	Cytoplasm	499
	Mitochondrion	171
	Apicoplast	209
	Endoplasmic reticulum	119
	Nucleus	19
Extracellular metabolites		241
<i>Unique metabolites</i>		673
Reactions		1066
Metabolic reactions		670
	Cytoplasm	374
	Mitochondrion	97
	Apicoplast	125
	Endoplasmic reticulum	67
	Nucleus	7
Transport reactions		396
	Cytosol—extracellular	241
	Cytosol—mitochondrion	58
	Cytosol—apicoplast	58
	Cytosol—endoplasmic reticulum	34
	Cytosol—nucleus	5
<i>Unique metabolic reactions</i>		558
Gene-protein-reaction associations		586
	Cytoplasm	328
	Mitochondrion	83
	Apicoplast	113
	Endoplasmic reticulum	50
	Nucleus	7
	Transport	5
<i>Unique genes</i>		325

doi:10.1371/journal.pcbi.1005397.t001

8] (see [S1 Methods](#) for more details). The coverage of metabolic enzymes ([Fig 1A](#)) is consistent with the ones reported in iTH366 and in the model of the related apicomplexan parasite *Toxoplasma gondii*, ToxoNet1 [3]. In contrast to the earlier GEMs of *P. falciparum*, iPfa is capable of growing on an *in silico* rich medium composed of 236 substrates that allow less restrictive conditions in the analyses. iPfa includes 245 common genes with iTH366 and 80 additional genes. Regarding metabolic tasks, iPfa accounts for 73 biomass building blocks of which 40 are common with iTH366 [5] and PlasmoNet [6], and nine are unique to iPfa. For example, some cofactors and the nucleotide sugars that are monomers of the essential GPI-anchor proteins [26] were not present in previous GEMs of *P. falciparum*. A detailed comparison of iPfa with the previous reconstructions iTH366 [5] and PlasmoNet [6] is provided in the [S1 Methods](#).

We integrated various omics data in iPfa, which include available uptake and secretion rates of important metabolites, such as lactate [27], glucose [27] and L-isoleucine [28] ([S1 Methods](#)). Within the TFA framework, iPfa integrated the ranges of metabolite concentration measured at the blood stage of the malaria infection [19–22] to study the effect of the concentration ranges on the reaction directionality and gene essentiality. The results of these analyses validate that iPfa serves as a scaffold for the integration of context-specific information.

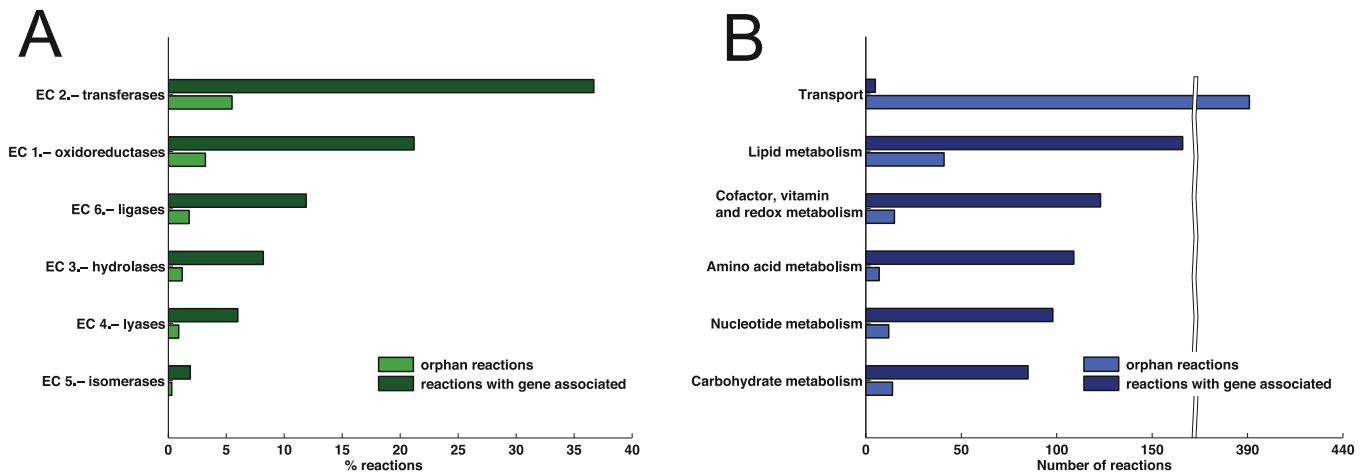


Fig 1. Metabolic capabilities (A) and subsystems (B) defined in iPfa.

doi:10.1371/journal.pcbi.1005397.g001

Gene and reaction essentiality predictions per metabolic task with iPfa

We used iPfa to study single and double gene essentiality *in silico* (S2 Table and S3 Table). To understand the effect of the simulated knockouts on specific metabolic tasks of *P. falciparum*, we designed an analysis that identifies which metabolic tasks cannot be fulfilled in iPfa upon disruption of each *in silico* gene (Materials and Methods).

We first performed essentiality studies without accounting for thermodynamic constraints and identified 55 genes that are essential for growth (S2 Table). The genes are involved in 26 of the 73 total metabolic tasks that are defined in iPfa. The essentiality of 31 of the 55 genes has been previously validated experimentally through gene knockouts, suppressed transcription or inhibition of enzyme activity by drugs in *P. falciparum*, whereas the remaining 24 genes have not been examined empirically, to our knowledge. We grouped the 24 non-validated genes in two classes: the first class is composed of 14 genes that participate in metabolic pathways where other genes have been experimentally defined as essential, and the second class contains 10 genes that are associated with functions or pathways where other genes have not been reported as indispensable. Therefore, we could not hypothesize the latter class of genes to be essential based on previous context-dependent information. For example, the MEP/DOXP pathway in the apicoplast is essential in *P. falciparum* for the production of isopentenyl diphosphate and isoprenoid derivatives [29]. There are nine genes in iPfa that are involved in the MEP/DOXP pathway, and all nine genes are predicted as essential for the synthesis of three biomass building blocks: isopentenyl diphosphate, geranylgeranyl diphosphate and ubiquinone-8 (S2 Table). Experimental evidence validates the predicted essentiality of four of the nine genes [30–33], while the other five genes of the pathway remain to be tested and, based on our classification of non-validated genes, are allocated to the first class. The results of such studies could be used for further model validation and refinement. The second class involves genes that are essential for the synthesis of cardiolipin, ubiquinone-8 (downstream enzymes from the MEP/DOXP pathway), sugar nucleotides and nucleotides (S2 Table).

Next, we performed essentiality analysis using TFA, which includes information about the Gibbs free energy of reaction ($\Delta_r G'$) and allows the integration of metabolomics data (Materials and Methods). We applied two different approaches within the TFA framework. In the first approach, we allowed the concentration of every intracellular metabolite to vary between 1 μ M and 50 mM. Besides, the concentration of extracellular metabolites (e.g., present in the

parasitophorous vacuole, red blood cell or blood serum) varied between 0.01 μM and 100 mM. Such concentration limits are similar to the physiological ranges used in previous TFA studies [12, 13]. The constraints in the first TFA approach set 57 reactions in iPfa to be unidirectional (S4 Table), which allowed the identification of three additional essential genes (S2 Table). The three genes are involved in 12 tasks, including two tasks not identified with FBA, i.e., the production of GTP and phosphatidylserine. The essentiality of one of the genes, the gene encoding a serine palmitoyl transferase (E.C. 2.3.1.50), has been validated experimentally [34]. The other two genes encode the ATP:UMP and ATP:GMP phosphotransferases (E.C. 2.7.4.4 and E.C. 2.7.4.8), which were predicted to be essential for the tasks involving nucleotides and sugar nucleotides.

The study of the tasks using TFA provided additional insight into the effect of simulated knockouts on the growth. Earlier studies on the pyrimidine biosynthetic pathway have targeted the gene encoding the orotidine 5'-monophosphate decarboxylase (E.C. 4.1.1.23) with antimalarial inhibitors [35]. Despite the suggestions that the gene is only involved in the synthesis of the pyrimidine nucleotides CTP and UTP [35], TFA proposed an additional explanation for its essentiality: the disrupted synthesis of dCTP, dTTP, and the nucleotide sugars UDP-N-acetyl-D-glucosamine, UDP-glucose, and UDP-D-galactose (S2 Table). The identification of a high number of biomass building blocks impacted upon knockout of a gene suggests that the inhibitory effect of the antimalarial drug on the parasite's growth is higher than believed. This result indicates a thermodynamic dependency between the production of phosphorylated and sugar nucleotides. The actual drug action mechanism should be further investigated with experiments and kinetic analysis.

In the second approach within the TFA framework, we integrated ten available metabolomics data sets, i.e., absolute concentrations measured in mature *P. falciparum* trophozoites [19–22] (Materials and Methods and S1 Methods). After the metabolomics data had been simultaneously integrated, 31 additional reactions in iPfa became unidirectional, and five additional genes were then identified to be essential (S1 Methods, S2 Table, and S4 Table). Interestingly, four of the genes had been assessed experimentally; three were found to be essential, and one was dispensable but growth reducing during the blood stages. The four genes encode a phosphatidyl serine carboxylase (E.C. 4.1.1.65) [36], a thioredoxin and glutathione oxidoreductase (E.C. 1.8.1.7) [37, 38] and a glycerol kinase (E.C. 2.7.1.30) [39]. The fifth gene, whose essentiality has not been assessed, encodes a glycerol 3-phosphate oxidoreductase (E.C. 1.1.1.21). All genes are associated with the production of phospholipids in the endoplasmic reticulum (ER) and are required to maintain the energy and redox balance in this compartment (see the next result section for further discussion). Although FBA identified alternative metabolic routes in iPfa for the production of phospholipids, TFA suggested that the alternative pathways were not thermodynamically feasible and that the five genes were essential. We also performed TFA with the simultaneous integration of the concentration data sets and identified 26 synthetic lethal pairs (S3 Table). Overall, the accuracy score of the gene essentiality predictions with iPfa (as defined in the Materials and Methods section) is 59%.

TFA identifies bottleneck metabolites and suggests substrate channeling in iPfa

The number of reactions whose directionality was affected by thermodynamic constraints and the number of essential genes identified by TFA varied depending on the metabolite concentration ranges, and on the metabolomics data set used. With the generic concentration range, TFA predicted three additional essential genes compared with the FBA prediction. When each data set was independently integrated, three to five additional essential genes were identified

(S2 Table). We evaluated the impact of the metabolite concentrations on the identification of the eight essential genes using TFA, and we identified *bottleneck metabolites* in iPfa. The *bottleneck metabolites* are the metabolites whose concentrations give rise to thermodynamic bottlenecks and are responsible for constraining the directionality of sets of reactions and for rendering any of these eight genes essential.

We first studied the effect of the generic concentration bounds on the prediction of the three essential genes. The lower and upper bound of the generic concentration range (1 μM —50 mM) were allowed to vary from 0.1 fM to 500 mM (Materials and Methods). While increasing the upper limit of the concentration range did not impact the essentiality of the three genes, decreasing the lower bound was critical for their essential function (Table 2).

When the intracellular concentrations were decreased and allowed to vary between 0.01 μM and 50 mM, the gene encoding a serine palmitoyl transferase (E.C. 2.3.1.50) was incorrectly predicted as non-essential [34]. The essential function of the ATP:UMP and ATP:GMP phosphotransferases (E.C. 2.7.4.4 and E.C. 2.7.4.8) was only lost when the lower bound of the intracellular concentrations was reduced four and ten orders of magnitudes, respectively, with reference to 1 μM . The available kinetic data, such as the Michaelis constant K_M , provides a measure of the substrate concentration required for effective catalysis to occur. The K_M of the enzymes usually varies between 10^{-1} and 10^{-7} M. The K_M values found in the literature indicate that metabolite concentrations of 1 fM are not physiologically relevant, which further supports the essential function of the *P. falciparum* genes predicted with TFA and generic concentration ranges.

We next sought the bottleneck metabolites and the associated reactions in iPfa whose directionalities were constrained and led to the essential function of the genes (Materials and Methods). The analysis of bottleneck metabolites suggests that AMP and ATP constitute a single minimal set of metabolites that is responsible for rendering three genes essential (Table 2). The three genes associated with the bottleneck metabolites AMP and ATP are essential for maintaining the AMP/ATP ratio in the ER, which varies between 2.4 and 1.9 mol/L_{cell} / mol/L_{cell} based on the absolute minimum and maximum experimentally measured concentrations [19–22] (S1 Methods).

We also found four alternative minimal sets of metabolites that are responsible for the essential function of one gene of the five genes that become essential with TFA and metabolomics data. Each of these four sets includes four metabolites from an overall set containing CMP, choline, AMP, ATP, ADP and CTP. The bottleneck analysis allowed the identification of metabolites involved in additional alternative sets of larger size (Materials and Methods, Table 2).

Overall, we identified 18 bottleneck metabolites whose concentration ranges played a critical role in the directionality of reactions in the ER and cytosol (Table 2). The bottleneck metabolites are phosphorylated nucleotides, sugar nucleotides, and intermediates in the production of phospholipids that link the pyrimidine, aminosugar and phospholipid metabolism in *P. falciparum*.

Two bottleneck metabolites are CTP and ethanolamine, which are involved in the synthesis of phosphatidylethanolamine (PE) from ethanolamine through the Kennedy pathway (Fig 2). The bottleneck analysis indicates that the directionality of the reactions and the essential function of the gene products in the ER and the Kennedy pathway are very sensitive to the concentration values of CTP and ethanolamine. These metabolites impact the metabolism with generic (broad) and experimentally measured (narrow) concentration values (Table 2).

TFA with metabolomics data rendered the Kennedy pathway thermodynamically infeasible. However, the genes of *P. falciparum* involved in the Kennedy pathway are expressed in the blood stages [40], and evidence reported in the literature indicates that the Kennedy pathway is essential in the rodent malaria parasite *Plasmodium berghei* [41], which further suggests that

Table 2. Bottleneck metabolites and affected reactions that determine the essential function of eight genes in iPfa.

Essential Gene (E. C.)	Min. set of bottleneck metabolites	Reaction impacted after integration of the bottleneck metabolites concentration ranges ¹ (Reaction directionalities ²)	All bottleneck metabolites among alternatives	Data set description
PF3D7_1415700 (2.3.1.50)	(i) CTP[r], Ethanolamine[c]	(Table SIII A and Table SIII D in S1 Methods)	CTP[r], Ethanolamine[c], Ethanolamine[r]	Required lower bound of 0.1 μM
	(i) CTP[r], Ethanolamine[r]			
PF3D7_0111500 (2.7.4.-)	(i) UDP-N-acetyl-D-glucosamine[c], UTP[c], UDP-glucose[c], UDP-D-galactose[c]	(Table SIII B and Table SIII D in S1 Methods)	UDP-N-acetyl-D-glucosamine[c], UTP[c], UDP-glucose[c], UDP-D-galactose[c]	Required lower bound of 1 nM
PF3D7_0928900 (2.7.4.8)	(i) GTP[c]	(Table SIII C and Table SIII D in S1 Methods)	Diphosphate[c], GTP[c], 2,5-Diaminopyrimidine nucleoside triphosphate[c], GDP-mannose[c]	Required lower bound of 1 fM
	(ii) 2,5-Diaminopyrimidine nucleoside triphosphate[c]			
PF3D7_0927900 (4.1.1.65)	(i) CDP[r], CDP-ethanolamine [r], Ethanolamine phosphate[r]	R02057_r (B)	CDP[r], CDP-ethanolamine[r], Ethanolamine phosphate[r]	MS 2012 [21]
		R01468_r (B)		
		R02038_r (B)		
		T_c_to_r_C00189 (B)		
		R02055_r (F)		
		T_c_to_r_C00065 (F)		
PF3D7_1351600 (2.7.1.30)	(i) CMP[r], Choline[r], AMP[r], ADP[r]	R01021_r (F)	CMP[r], Choline[r], AMP[r], CTP [r], ADP[r], ATP[r]	MS 2012 [21] , NMR 2009 [19] , NMR 2014 only strains 3D7, 7G8, C2 (GC03), C4(Dd2), C6 (7G8) [20]
		R01890_r (F)		
		R01321_r (F)		
	(ii) CMP[r], Choline[r], AMP[r], ATP[r]	R01280_r (F)		
		R00127_r (F)		
		R00094_r (F)		
	(iii) CMP[r], Choline[r], AMP[r], CTP[r]	T_c_to_e_C00116 (F)		
		T_c_to_r_C00111 (F)		
		R00842_r (R)		
	(iv) CMP[r], Choline[r], CTP[r], ADP[r]	R00847_r (R)		
		T_c_to_r_C00013 (R)		
		T_c_to_r_C00116 (R)		
PF3D7_1419800.1 (1.8.1.7)	(i) AMP[r], ATP[r]	R00094_r (F)	ADP[r], AMP[r], CDP-choline[c], ATP[r], CTP[r], CTP[c], CMP[r]	All the data [19–21]
		R01280_r (F)		
		R00127_r (F)		
PF3D7_0923800.2 (1.8.1.7)	(i) AMP[r], ATP[r]	T_c_to_r_C00111 (F)		
PF3D7_1216200 (1.1.1.21)		R00842_r (R)		
		T_c_to_r_C00013 (R)		

¹Reaction names as defined in iPfa. Metabolic reactions are defined with their R-5 digit identifier as obtained from the KEGG database. Transport reactions are marked with T_. Cellular compartments are defined with the reaction name: _r, endoplasmic reticulum; _c, cytosol; _e, extracellular (outside the parasite's cell)

²Reaction directionalities obtained with Thermodynamic Variability Analysis (TVA): (B) blocked, (F) forwards, (R) reverse. See the [S4 Table](#) for the reaction description.

doi:10.1371/journal.pcbi.1005397.t002

the pathway is functional. The analysis of thermodynamic bottlenecks identified CDP-ethanolamine, CDP and ethanolamine phosphate as the bottleneck metabolites that block the reaction from CDP-ethanolamine to PE in the Kennedy pathway. Previous studies have also identified a thermodynamic impediment in the Kennedy pathway based on metabolite concentrations and have suggested channeling between ethanolamine phosphate and PE through the enzymes E.C. 2.7.7.14 and E.C. 2.7.8.1 [\[42\]](#). When we simulated the presence of such channeling in TFA

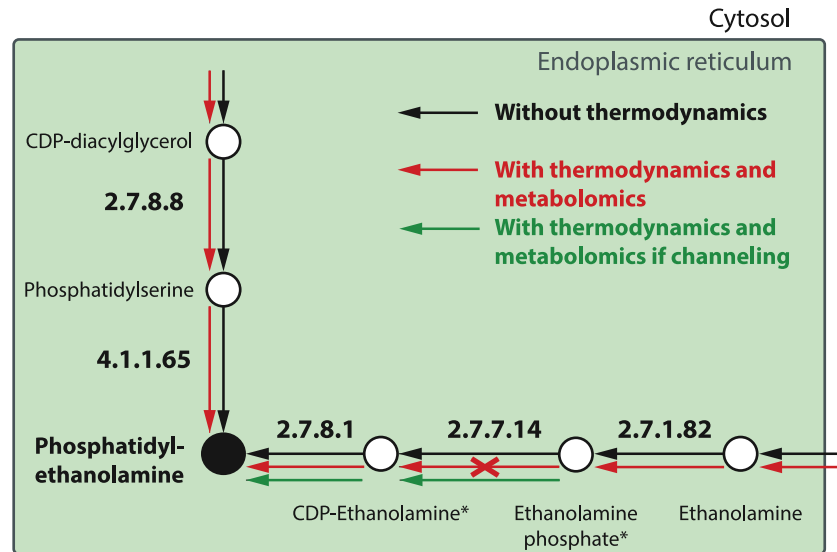


Fig 2. Bottleneck metabolites and substrate channeling in the production of PE. Arrows denote reaction fluxes. Circles define metabolites. Reactions are described by their 4-digit E.C. identifier. (*) are bottleneck metabolites.

doi:10.1371/journal.pcbi.1005397.g002

(Materials and Methods and [S1 Methods](#)), the Kennedy pathway became feasible ([Fig 2](#)) and the phosphatidylserine carboxylase (E.C. 4.1.1.65) is predicted as non-essential (further discussion in [S1 Methods](#)). These results demonstrate how TFA and the bottleneck metabolite analysis can be used to integrate biochemical studies and hypothesis testing.

Study of nutritional requirements with iPfa

Obligate intracellular parasites, such as *P. falciparum*, depend on the availability of a broad array of metabolites provided by the host cell and its surroundings. To characterize the nutritional requirements of *P. falciparum*, we followed and extended the approach previously applied to the related pathogen *T. gondii* to study the *in silico* minimal medium (IMM) [3].

The *in silico* minimal medium (IMM) in iPfa. We used iPfa and an *in silico* rich medium composed of 236 substrates ([S1 Methods](#)) and searched for the IMM, defined as the minimum number of substrates required for growth [3]. The minimum number of substrates was 23, and there are 10,032 such IMM sets, which are generated by the alternative combination of only 52 substrates ([S1 Methods](#), [Table 3](#)). We further identified 16 constitutive substrates, i.e., metabolites that were present in all alternative IMM, and 36 non-constitutive substrates, i.e., metabolites that varied in each alternative IMM. In theory, one would expect $36!/((36-7)! \cdot 7!) = 8,347,680$ combinations ([S1 Methods](#)); the relatively small number of alternatives suggests a limited ability of the 36 substrates to substitute for each other.

Next, we investigated the reasons for the limited number of alternatives. Given the requirement of minimal utilization of substrates in the IMM studies, one would expect that some of the non-constitutive substrates that do not appear in the same IMM could substitute for each other. We created groups of non-constitutive substrates that did not belong to the same IMM and found that 20 of the 36 components were substitutable, suggesting that each of these components serve a specific biosynthetic requirement. However, ten can be substituted by one or more than one substrate in the IMM, which indicates that they serve multiple biosynthetic needs. For example, S-adenosyl-L-methionine can be substituted by S-adenosyl-L-homocysteine alone, or by two substrates simultaneously. In the latter case, one substrate is a source of the purines, such

Table 3. Composition of the *in silico* minimal media (IMMs) that allow growth of iPfa and essentiality.

Constitutive metabolites of the IMM		Essential ¹
Source of amino acids	L-Isoleucine; Oxyhemoglobin	E; NE
Source of inorganic sulfur	Sulfate; Sulfur donor	E; E
Source of inorganic iron	Heme	NE ²
Cofactors and others not synthesized <i>de novo</i>	Biotin; Methylcobalamin; Thiamine; Cholesterol	E; E; E; E
Precursors for lipoylation	Lipoate; Apoprotein	E; E
Precursors of lipid components	Choline; N-Acylsphingosine	E; NE ²
Precursor of FMN/FAD	Riboflavin	E
Precursors of isoprenoids	4-Hydroxybenzoate; HCO ³⁻	E; E
Non-constitutive metabolites of the IMM ³		
Source of inorganic phosphate	Orthophosphate or Diphosphate	NE ²
Precursors of CoA	Pantothenate or Pantetheine or N-((R)-Pantothenoyl)-L-cysteine	E
Source of pyridine ring (* and source of carbon)	Nicotinate D-ribonucleoside* or N-Ribosylnicotinamide* or Nicotinate or Nicotinamide	E
Source of DNA nucleotides (* and source of carbon)	S-adenosyl-L-methionine* or S-adenosyl-L-homocysteine* or S-adenosylmethioninamine* or Se-adenosyl-L-selenohomocysteine* or Adenosine* or Inosine* or Guanosine* or Xanthosine* or Adenine or Guanine or Hypoxanthine or Xanthine	NE ²
Source of folate and derivatives	Tetrahydrofolate (THF) or 5,10-MethyleneTHF or 10-FormylTHF or 5-MethylTHF or 5,10-MethenylTHF or Dihydrofolate or Folate or Dihydropterolate or 4-Amino-4-deoxychorismate or Chorismate or 4-Aminobenzoate	E
Source of pyrimidine ring (precursors of UMP)	Orotate or (S)-Dihydroorotate	E
Source of C2/C4 for acetyl-CoA (precursors of nucleotide sugars)	Acetate or L-2-Amino-3-oxobutanoic acid	E

¹Phenotype observed in the simulation after one-by-one depletion of the substrate or group of substrates in the *in silico* rich medium of 236 substrates: E, essential; NE, non-essential.

²See the next [result](#) section for explanation.

³The 10,032 alternative IMMs are generated through the combination of the constitutive metabolites and one non-constitutive metabolite from each group reported in [Table 3](#). Note that the following constraints in the combinations should be considered: (a) the IMM should provide a ribose-containing molecule, which serves as source of carbon (marked with * in [Table 3](#)) and (b) the presence of orthophosphate and S-adenosylmethioninamine in the same IMM is not allowed ([S1 Methods](#) for more details).

doi:10.1371/journal.pcbi.1005397.t003

as adenine, guanine, hypoxanthine or xanthine, and the other serves as a source of carbon in the form of a ribose, such as nicotinate D-ribonucleoside or N-ribosylnicotinamide (Materials and Methods). This result demonstrated that the non-appearance of two substrates in the same IMM is a necessary but not sufficient condition to define the substrate substitutability.

We further searched for a measure that could determine whether two substrates are substitutable and found that the molecular structure of the 36 non-constitutive substrates could provide information on this criterion. Specifically, we identified eight common molecular substructures (here referred as *backbone moieties*) among the non-constitutive substrates, which allowed us to cluster the 36 non-constitutive substrates of the IMM ([Table 3](#)). Twenty-six substrates contain a unique backbone moiety out of five backbone moieties ([Table 3](#)). Importantly, substrates that contain the same unique backbone moiety are substitutable. Ten substrates contain two out of three backbone moieties ([Table 3](#)). These substrates can be substituted by one or more than one substrate in the IMM, as in the case of S-adenosyl-L-methionine discussed above.

The identification of the backbone moieties sheds light on the metabolic capabilities of *P. falciparum*. Analysis of the metabolic fate of the backbone moieties suggested that the cell may

not be able to synthesize these backbone moieties *de novo* and must scavenge them from the substrates provided by the host cell. For example, the IMM analysis correctly identified the following nutritional requirements: the substrates pantothenate [43], hypoxanthine [44] or nicotinate [45], which are provided by default *in vitro* to maintain the cultures of *P. falciparum* [46]. However, our analysis suggested that each of these molecules could be substituted with other derivatives (Table 3). Nicotinate belongs to a group of substitutable substrates, which includes nicotinamide, nicotinate D-ribonucleoside, and N-ribosylnicotinamide. The four substrates contain the pyridine ring as a backbone moiety and this ring is also preserved in the final NAD⁺ molecule and its derivatives. Similar observations were obtained for the remaining backbone moieties (see next result section). Additionally, intraerythrocytic parasites rely on the uptake of exogenous niacin to synthesize NAD⁺ via the canonical Preiss-Handler salvage pathway [45]; however, to our knowledge, no one has studied the possibility of its substitution with D-ribonucleoside or N-ribosylnicotinamide, or with other molecules that provide the corresponding backbone moiety required for growth. The malaria parasites might be able to use a medium with fewer metabolites that contain the essential backbone moieties. Hence, the knowledge of nutritional requirements based on molecular structures will be used to simplify to its maximum extent the medium formulation, design synthetic media [47], and study the uptake mechanism and the metabolic fate of the backbone moieties.

Essentiality of IMM components. Although the components of the IMM serve all the essential functions in iPfa, they may not be essential when a rich medium of 236 substrates is available because other substrates could substitute for them. We removed one constitutive substrate at a time from the rich medium and found that 13 of the 16 constitutive substrates were essential, i.e., they are also required in a rich medium (Table 3). The other three substrates are heme, oxyhemoglobin, and N-acylsphingosine. They need more than one alternative substrate for their synthesis by the cell, e.g., heme. Or they serve as a single source of multiple components that are available in the rich medium or can be derived from it, e.g., hemoglobin, a source of various amino acids that exist in the rich medium.

To identify the essential functions served by the non-constitutive components of the IMM (Table 3) in iPfa, we analyzed the essentiality of the backbone moieties present in these components. Here, we removed the whole set of molecules that contained a backbone moiety (one moiety at a time) from the rich medium and found that five of the eight backbone moieties identified in the IMM were essential. This result suggests two fulfilled conditions for the five backbone moieties. There is a lack of other molecules (besides the 36 non-constitutive substrates of the IMM, Table 3) in the rich medium that can directly provide the five moieties, and *P. falciparum* is unable to synthesize the moieties from any other substrate available in the rich medium.

The essentiality analysis of the backbone moieties also indicates that other molecules (besides the 36 non-constitutive substrates of the IMM, Table 3) in the rich medium could provide three of the eight backbone moieties. The three backbone moieties are the sources of carbon, phosphate, and purines. We next performed iterative IMM analyses until we found new sets of substrates that provided the three backbone moieties (Fig 3). Interestingly, all of the new sets of substrates contained the three backbone moieties (Fig 3). The essentiality analysis (Materials and Methods) of all molecules that contained the three backbone moieties confirmed that these three backbone moieties are also essential. This observation suggests that iPfa does not possess the biosynthetic capabilities to produce the three backbone moieties and indicates that *P. falciparum* must scavenge the eight backbone moieties from the host cell or its surrounding.

We validated our results with available experimental work on the nutritional requirements of *P. falciparum* [43, 46, 48–51]. Although glucose, hypoxanthine, and pantothenate are

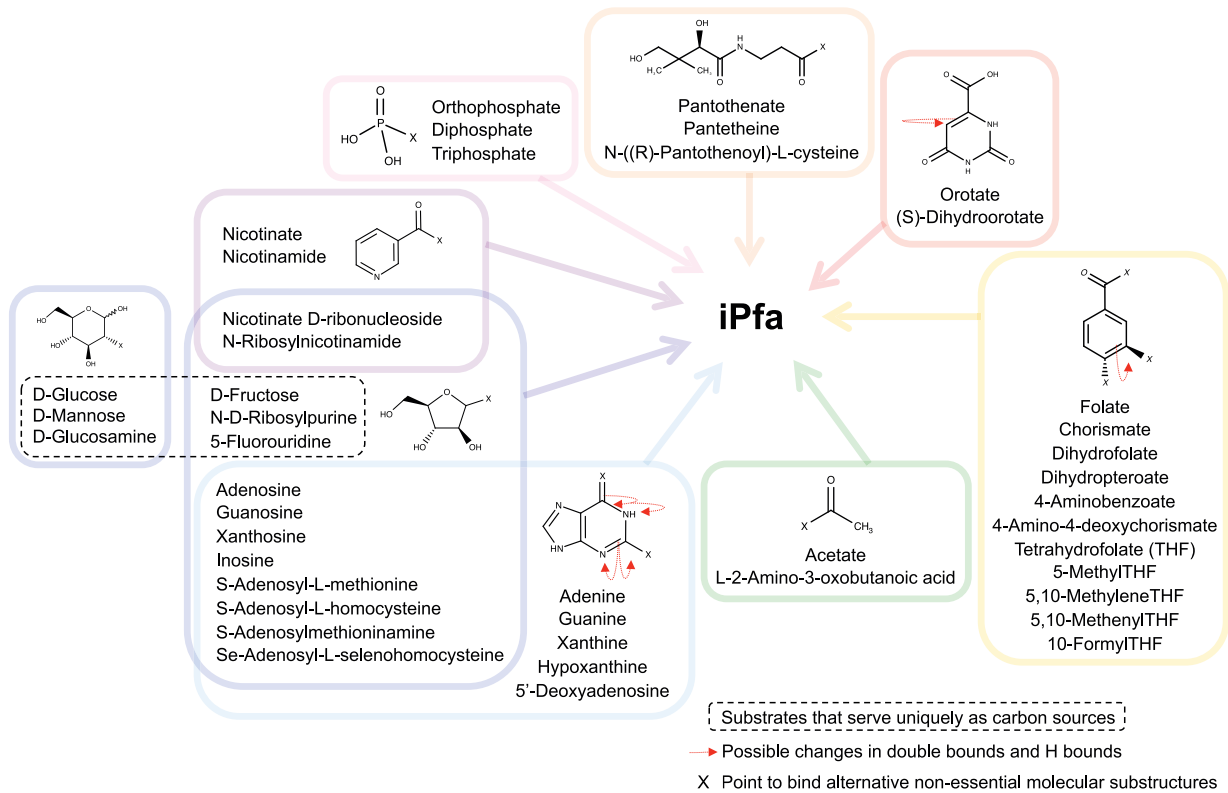


Fig 3. Nutritional requirements of *P. falciparum* concerning essential backbone moieties. Note: the presence of orthophosphate and S-adenosylmethioninamine in the same medium is not allowed. When *P. falciparum* grows on orthophosphate and S-adenosylmethioninamine as the only sources of phosphate and purines, it cannot synthesize enough ATP. The ATP limitation impedes the production of other phosphorylated nucleotides, sugar nucleotides, sphingomyelin and phospholipids in the stoichiometrically required amounts. As suggested throughout the manuscript, these metabolic processes are thermodynamically dependent (more details in the [S1 Methods](#)).

doi:10.1371/journal.pcbi.1005397.g003

known to be the preferred carbon, purine, and CoA sources respectively [43, 51], we have identified for the first time all alternative metabolites that are thermodynamically allowed to serve these functions (Fig 3). We further suggest that *P. falciparum* might use acetate or 2-amino-3-oxobutanoate to produce acetyl-CoA in the cytosol, which is then converted into the biomass building block UDP-N-acetyl-D-glucosamine.

Orotate is an intermediate metabolite in the canonical and highly conserved biosynthetic pathway of UMP, a common precursor of pyrimidine nucleotides. TFA of iPfa suggested that the biosynthetic pathway of pyrimidines is thermodynamically infeasible and hence the orotate backbone moiety, provided by orotate and (S)-dihydroorotate, needs to be taken up. However, the biosynthetic pathway of UMP, which involves six enzymes, is known to be functional [40] and essential in *Plasmodium* spp. given that the malaria parasites cannot salvage pyrimidines [50, 52]. We performed a comparative FBA against TFA to identify the thermodynamically constrained reaction(s) and found that the dihydroorotase reaction (DHOase, E.C. 3.5.2.3), which produces (S)-dihydroorotate from carbamoyl-L-aspartate, was not thermodynamically feasible. The malarial enzyme DHOase shares some characteristics with both types I (e.g., mammals) and II (e.g., *Escherichia coli*) enzymes [53]. Previous experimental [54] and computational [55] studies in DHOases of type II have hypothesized that channeling exists from carbamoyl phosphate to (S)-dihydroorotate through the enzymes E.C. 2.1.3.2 and E.C. 3.5.2.3. When TFA was used to simulate the presence of such channel in iPfa (Materials and Methods

and [S1 Methods](#)), the pyrimidine biosynthetic pathway became thermodynamically feasible, and this backbone moiety could be synthesized by the parasite. Interestingly, previous attempts to develop the malaria parasites axenically have suggested that orotate incorporation can be used as a direct measurement of parasite development [50]. The presence of orotate in the medium has been shown to inhibit the function of the DHOase enzyme [53].

Gene essentiality studies in the IMM. The essentiality of the medium components suggests a complex interplay between the chemical structure of the substrates and the catalytic capabilities of the cell. To identify these links in *P. falciparum*, we evaluated single gene essentiality in iPfa for each of the alternative IMMs ([S1 Methods](#)). We identified 96 essential genes; the 63 genes predicted to be essential in the *in silico* rich medium (less restrictive condition) ([S2 Table](#)) remained indispensable in the alternative IMMs (more restrictive condition). We identified nine essential genes that were common in all IMMs ([S5 Table](#)), and the remaining 24 essential genes were distributed among the alternative IMMs with the number of genes varying among the alternative IMMs ([S5 Table](#)).

Discussion

In this work, we reconstructed a GEM for *P. falciparum* and performed thermodynamics-based flux analysis to study the essential metabolic capabilities and the nutritional requirements of the parasite. Besides, we identified thermodynamic bottlenecks and hypothesized substrate channeling in the metabolic pathways. The results from this work, which have been compared with available experimental evidence, demonstrate that iPfa and its TFA can be used to fill in the knowledge gaps regarding the critical aspects of *P. falciparum* metabolism.

The highest accuracy score (59%) in the prediction of essential genes for growth was obtained using TFA with metabolomics data. Such approach revealed up to 63 lethal genes and 26 synthetic lethal pairs in iPfa ([S2 Table](#) and [S3 Table](#)). Existing empirical evidence supports the essential function of 35 genes, and 26 genes have been targeted with antimalarial drugs. Our analysis identifies 28 essential genes for which no experimental validation was found in the literature. Previous studies have hypothesized the essential function of 15 of the 28 genes. Overall, 13 essential metabolic capabilities had not been previously identified and one of them, which is critical for the production of cardiolipin, is not similar to any human protein ([S2 Table](#)). The 28 non-supported predictions require further testing for their validation as potential antimalarial drug targets, as suggested in the iterative cycle of systems biology [11].

The analysis of metabolic tasks identifies the effect of the gene knockouts on the parasite's growth and suggests that the disruption of the 63 essential genes impair the production of up to 30 biomass precursors ([S2 Table](#)). When thermodynamics is taken into account, the number of metabolic tasks impacted by the knockout of one gene increases and it provides additional evidence for the complex interactions between genes, reactions, and metabolites in the metabolic networks. Such is the case of genes required to produce pyrimidines, such as the orotidine 5'-monophosphate decarboxylase, which represents an attractive antimalarial drug target ([S2 Table](#)). The knockout of orotidine 5'-monophosphate decarboxylase in iPfa impedes the assembly of pyrimidines, sugar nucleotides, and some DNA nucleotides. This result suggests the thermodynamically dependent function of the metabolic pathways producing phosphorylated nucleotides and sugar nucleotides in *P. falciparum*.

We identify eighteen bottleneck metabolites in iPfa, whose intracellular concentrations determine the directionality of a set of reactions and render eight genes essential ([Table 2](#)). The bottleneck metabolites are phosphorylated nucleotides, sugar nucleotides, and intermediates in the production of phospholipids. The concentrations of the bottleneck metabolites play

a critical role on the directionality of reactions in the cytosol and ER that are mainly involved in the production of phospholipids. Specifically, the genes identified as essential were required to maintain the redox and energy balance in the ER. Additional analysis using the TFA framework helped us quantify the effect of the metabolite concentrations on the reaction directionalities and can guide future metabolomics studies on the malaria parasites ([S1 Methods](#), [S1 Dataset](#)). The results of these studies further suggest that there exists in *P. falciparum* a thermodynamic dependency between the production of phospholipids, phosphorylated nucleotides, and sugar nucleotides through the concentrations of the bottleneck metabolites. The knowledge of metabolic subsystems that are thermodynamically dependent can be used to design drugs that synergistically target these parts of the metabolism and prevent their metabolic regulation and rise of drug-resistant parasites.

Through the analysis of thermodynamic bottlenecks, we identify two thermodynamically infeasible pathways, i.e. the Kennedy and the pyrimidine biosynthesis pathway, that are functional in *P. falciparum* [40]. The presence of substrate channeling between ethanolamine phosphate and PE in the Kennedy pathway and between carbamoyl phosphate and (S)-dihydroorotate in the pyrimidine biosynthetic pathway turns them thermodynamically feasible. Hypotheses on substrate channeling through these pathways have been suggested in other cell types based on the metabolite concentration levels and thermodynamic studies [42, 54, 56]. We demonstrate that network thermodynamics can provide a framework for integrating and testing hypotheses on enzyme mechanisms and their function in metabolic networks [14, 16]. The final validation of substrate channeling in *P. falciparum* requires the experimental characterization of the enzymatic mechanisms using techniques such as isotope dilution [18].

The definition and identification of the IMM provide important insight into the nutritional requirements of *P. falciparum*. The IMM analysis suggests that *P. falciparum* requires at least 23 substrates for growth and 13 of them are indispensable ([Table 3](#)). Certain substrates, such as pantothenate [43], hypoxanthine [44] and nicotinate [45], which are provided by default *in vitro*, might be substituted by other substrates to support growth. The ability of substrates to substitute for each other is better understood by considering backbone moieties, which are defined here as molecular substructures shared between substrates that can replace each other to support growth. We find that eight backbone moieties exist in the metabolic network of *P. falciparum* and that seven substrates can provide these eight backbone moieties. The analysis of the requirement of the IMM components in a rich medium indicated that *P. falciparum* is auxotrophic for the eight backbone moieties ([S1 Methods](#)). The malaria parasites might grow on a medium with fewer metabolites that contain the essential backbone moieties. However, not all combinations of substrates support the growth of *P. falciparum*. An infeasible combination involves orthophosphate and S-adenosylmethioninamine as the only sources of phosphate and purines. These substrates do not allow the distribution of ATP for the production of phosphorylated nucleotides, sugar nucleotides, sphingomyelin and phospholipids ([S1 Methods](#)). The knowledge of nutritional requirements based on molecular substructures will be used to simplify to its maximum extent the medium formulation, design synthetic media [47], and study the uptake mechanism and the metabolic fate of the backbone moieties.

Based on available high-throughput gene knockout data for *P. berghei* [57, 58], we expect that most of the genes that TFA predicted to be essential for *P. falciparum* development ([S2 Table](#)) will also be essential *in vivo*. The study of *P. falciparum* metabolism on particular life stages metabolism will allow the identification of additional essential metabolic capabilities and nutritional requirements, such as the need to uptake oleic acid [48]. New methods, such as high-throughput gene knockout [57, 58] and metabolomics techniques, can be applied to *Plasmodium* species and can provide large omics data sets on specific life stages of the malaria infection. iPfa is an ideal scaffold to integrate such context-specific information and to perform

integrative studies of *P. falciparum* metabolism. A future comparison of *in vivo* knockout high-throughput data and the essentiality analysis on the IMM will advance our understanding of the substrates that are accessible to the malaria parasite in the host. The GEM iPfa should be further developed to keep it up-to-date, and the detailed description of the reconstruction process provided in the [S1 Methods](#) serves as a reference for this purpose. Modifications of iPfa will allow the study of context-specific cases, i.e., modeling of the blood stage or liver stage metabolism [59] ([S1 Methods](#)).

In conclusion, the results from the TFA of iPfa fill a significant knowledge gap regarding the essential metabolic capabilities and needs of *P. falciparum*. The studies on bottleneck metabolites and substrate channeling provide hypotheses for future analysis of the endometabolome and the enzyme mechanisms in malaria parasites. The model, the analysis framework and the results presented here are a valuable resource that can facilitate the ongoing experimental efforts to obtain a better understanding of *P. falciparum*'s physiology and identify novel drug targets for antimalarial intervention.

Materials and methods

Reconstruction of iPfa

The *P. falciparum* 3D7 genome sequence was retrieved from PlasmoDB [60]. The enzymatic functions of the proteins were annotated using the RAVEN Toolbox [24] ([S6 Table](#)). A detailed description of the reconstruction process, the software and the data used is provided in the [S1 Methods](#).

Flux balance analysis (FBA)

FBA is a well-established approach for the computational analysis of large-scale metabolic networks [9, 10]. FBA was performed on iPfa, with the maximization of biomass growth as the objective function.

Thermodynamics-based flux analysis (TFA)

We integrated the thermodynamic properties of the metabolites and reactions in iPfa in the form of thermodynamic constraints following the systematic approach defined within the framework of Thermodynamics-based Flux Analysis (TFA) [13, 61, 62], which has been also referred to as Thermodynamics-based metabolic Flux Analysis (TMFA) [13] and Thermodynamics-based Flux Balance Analysis (TFBA) [15, 63]. Thermodynamic constraints determine the feasible range of $\Delta_r G'$ and hence reduce the uncertainty in the reaction directionalities and with it the feasible solution space that is characteristic of highly underdetermined problems like the analysis of metabolic networks with Flux Balance Analysis (FBA). These constraints in iPfa accounted for the intracellular conditions, like the pH, the membrane potential and the intracellular concentration ranges of the metabolites ([S1 Methods](#)). TFA further allowed the integration of experimentally measured concentration ranges of metabolites [19–22].

Integration of metabolomics data sets

We created a thermodynamically curated version of iPfa where we allowed the concentration of every intracellular metabolite to vary between 1 μM and 50 mM, which is the physiological range used in similar TFA studies [12, 13]. The concentrations of metabolites outside the parasite cell were allowed to vary between 0.01 μM and 100 mM. We also generated various thermodynamically curated versions of iPfa integrating one-at-a-time each of the ten metabolomics data set considered in this study [19–21]. Nine metabolomics data sets were measured

with NMR: eight were obtained from different isolates of *P. falciparum* trophozoite-infected red blood cells [20] and one from isolated trophozoites [19]. The remaining data set was measured with LC-MS in isolated trophozoites [21]. The physiological or generic range of concentration (defined above) was considered for a metabolite if no data was available in the metabolomics sets. For the metabolites present in more than one intracellular compartment of iPfa, the same concentration range was defined in all of these compartments. We also generated a combined metabolomics data set [19–22], in which one unique concentration range was calculated for each metabolite appearing in multiple data sets. This unique concentration range comprised all the measured concentration values for that metabolite, i.e. the minimum value and the maximum value measured. Overall, there are absolute concentrations ranges for a total of 61 metabolites [19–22].

We translated the reported values and measurement errors into ranges of concentration expressed in mol/L_{cell}. The concentration ranges are integrated within TFA and constrain the allowable $\Delta_r G'$ range of the reactions in which the metabolites participate and with it the flux ranges of the neighboring reactions. We integrated the metabolomics data sets one-by-one and simultaneously to study the number of bidirectional reactions and the predictions of essential genes. The essential genes varied among data sets, and the data set measured with LC-MS [21] and the combined metabolomics data set allowed the identification of the maximum number of essential genes.

Gene essentiality per metabolic task

This analysis allowed the identification of the metabolic tasks or biomass building blocks that could not be performed or produced, respectively, upon knockout of the essential genes. The analysis involved two steps: first, the essential genes for biomass growth were identified following standard procedures (S1 Methods). Second, the disruption of each essential gene was applied in iPfa, and the production of each biomass building block was tested. The building blocks that could not be produced individually were identified. An MILP formulation was defined to identify the groups of building blocks that could not be produced at the same time due to stoichiometric requirements. A gene was defined as essential when its knockout led to a specific growth rate slower than 10% of the optimal growth value predicted (S1 Methods). The optimal growth value was the same in all scenarios, i.e. using FBA and TFA. The threshold used does not have any impact on the identification of essential genes and reactions in this study (see Fig SII A and Fig SII B in S1 Methods) since no knockouts led to *in silico* growth reducing phenotypes. No additional filtering was applied to identify essential genes and pairs of genes.

Comparison of predictions with experimental data

The genes predicted as essential in iPfa with FBA and TFA were compared with experimental information available in the literature. We used primarily experimental data for *P. falciparum* (*in vitro*) available in the literature. Information for other *Plasmodium* spp. was also used for comparison. An accuracy score was calculated as $(TP+TN)/(TP+TN+FP+FN)$, where TP (true positive) and TN (true negative) define predictions that correctly simulate growth and non-growth, respectively, based on the available experimental data. While FP (false positive) and FN (false negative) describe predictions that incorrectly simulate growth and non-growth, respectively. The accuracy score served to evaluate the prediction of essential genes. For the predictions with TFA and metabolomics data integrated, an accuracy score of 59% is obtained based on the data available for *P. falciparum* in the literature (references reported in S2 Table and previous reviews [64]). Overall in the literature, we found information about the essential

function of 71 genes out of the total 325 genes in iPfa. We also evaluated the predictions with the available data from the high-throughput genetic screening of the mouse malaria parasite *P. berghei* in the blood stages (PlasmoGEM data) [57, 58]. The accuracy score based on PlasmoGEM data is 51%, and it provides information for 233 orthologous genes out of the total 325 genes in iPfa.

We further tested whether the enzymes predicted as essential in iPfa are similar to any human protein. We define *similarity* between the essential enzymes in iPfa and the human proteome when there exists at least one matching protein in non-redundant human proteome (txid9606) with the standard protein BLAST settings in the bioinformatics software blastp version 2.6.0+ [65]: cutoff value 1E-10, matrix BLOSUM62 [66].

Sensitivity analysis on the concentration bounds in TFA

We reduced the lower bound ten orders of magnitudes (minimum intracellular concentration of 0.1 fM) and increased the upper bound ten fold (maximum intracellular concentration of 500 mM). We then recalculated the feasible range of $\Delta_r G'$ and tested the gene essentiality of the three genes. These are the three genes that were further identified as essential with TFA using the generic concentration ranges (1 μ M—50 mM) with reference to the FBA prediction (S2 Table). The minimum concentration bound required to predict each of the three genes as essential with TFA was identified and reported in Table 2.

Identification of bottleneck metabolites

This study identified the metabolites responsible for the thermodynamic bottlenecks, which determined the directionality of a set of reactions and allowed the identification of eight genes as essential (Table 2). iPfa was used with generic concentration ranges (1 μ M—50 mM) or simultaneous integration of the experimental concentration ranges, and each of the eight genes was knocked out separately (S2 Table). These models were feasible with FBA, but not with TFA. An MILP formulation was defined to search for the minimal number of metabolites whose concentration ranges should be relaxed to make the model feasible in TFA. All the alternative solutions were obtained, and the minimal sets were formed by picking one metabolite from each alternative. The minimal set should involve at least one metabolite from each alternative. The metabolites that were shared among more alternatives appeared in the minimal sets.

Thermodynamic Variability Analysis (TVA) and inference of reaction directionality

The Gibbs free energy of a reaction ($\Delta_r G'$) is a measure of its thermodynamic potential and based on the second law of thermodynamics, defines the thermodynamically feasible direction under which the reaction can operate, defined as reaction directionality [14–16]. The $\Delta_r G'$ accounts for the standard Gibbs free energy of a reaction ($\Delta_r G^{\circ}$), which was calculated using the group contribution method (GCM) [61, 62], and for the concentration of the metabolites involved in the reactions. The uncertainty associated with the values of $\Delta_r G^{\circ}$ and the metabolite concentrations is expressed within TFA in the form of ranges that ultimately define the feasible range of $\Delta_r G'$ associated with each reaction. The range of $\Delta_r G'$ was known with Thermodynamic Variability Analysis (TVA) [13], which follows the same principles as the Flux Variability Analysis (FVA) [67]. We applied TVA with a requirement of 90% of the optimal growth. When the whole range of $\Delta_r G'$ is negative, then we define the reaction to be unidirectional in the forward direction. When the range of $\Delta_r G'$ is positive and negative, then we define the reaction as bidirectional.

Substrate channeling integration within TFA

The presence of substrate channeling between two enzymes (E_1 and E_2) was simulated within the TFA framework. The reactions R_1 and R_2 catalyzed by E_1 and E_2 were lumped to define an overall reaction ($L = R_1 + R_2$). The overall reaction eliminates the common intermediates and recycled metabolites since they are produced in the first reaction R_1 and consumed in the second reaction R_2 and vice versa, respectively. For example, if R_1 involves the transformation $A + B \rightarrow I + C$ and R_2 is defined as $I + D \rightarrow P + B$, the final lumped or overall reaction L is defined as $A + D \rightarrow P + C$. The overall reaction L does not involve the intermediate metabolite I and the recycled metabolite B . It is important to note that reactions R_1 and R_2 should be first defined in the direction that satisfies mass balances and allows flux through the pathway of study. Then the lumped reaction can be formed. For a pathway of study that contains reactions R_1 and R_2 the starting and product metabolites are usually A and P , respectively. Then, the reactions R_1 and R_2 should be defined as $A + B \rightarrow I + C$ and $I + D \rightarrow P + B$. Other definitions, such as $A + B \rightarrow I + C$ and $P + B \rightarrow I + D$, do not satisfy mass balances. In this study, the reactions of the pyrimidine biosynthesis pathway and the Kennedy pathway were defined to allow production of UMP and PE, respectively. The overall reaction was then allowed to be bidirectional. Thermodynamic properties were calculated for the overall reactions and TFA and Thermodynamic Variability Analysis (TVA) [13] were performed to determine the thermodynamically feasible directionality of the overall reactions. Substrate channeling was suggested when the directionality of the overall reaction allowed flux through the biosynthetic pathway that would be otherwise thermodynamically infeasible. The directionality of the reactions was known from the sign of the range in the $\Delta_r G'$ obtained from TVA.

Studies on *in silico* minimal medium (IMM)

The IMM analysis was performed following the strategy defined before [3]. Here, TFA was applied with the combined metabolomics data set integrated into iPfa (S1 Methods).

Identification of groups of substrates that can substitute for each other for growth

This study was performed on the substrates of iPfa to validate their ability to substitute for each other and support growth. Groups of substrates were created based on the IMM analysis. Substrates were grouped if they never appeared in the same IMM and presented a common molecular substructure or backbone moiety. We validated these groups in two steps. First, we tested that iPfa could not grow when all substrates in a group were removed from the *in silico* medium of 52 metabolites identified in the IMM. Second, we tested that under such conditions, the inclusion of each substrate of the group individually allowed simulated growth in iPfa. This analysis was also performed in the rich medium of 236 substrates to validate the groups of substrates that contained the three backbone moieties (the sources of carbon, phosphate, and purine).

The tools and methods used are described in more detail in S1 Methods.

Supporting information

S1 Methods. Describes in more detail the tools and methods used in this study. (PDF)

S1 Table. Gap-filling reactions and alternatives that determine the thermodynamic feasibility of the metabolic tasks in iPfa and the *in silico* rich medium. ¹Gap-filling reactions are defined with their R-5 digit identifier as obtained from the KEGG database. Transport

reactions are marked with T_. Cellular compartments are defined with the reaction name: _r, endoplasmic reticulum; _c, cytosol; _m, mitochondrion; _a, apicoplast. Note that the same gap-filling reaction can be suggested for more than one metabolic task.

(XLSX)

S2 Table. Single essentiality predictions in iPfa. ¹Genes predicted as essential for iPfa growth in the *in silico* rich medium (composed of 236 substrates, [S1 Methods](#)). ²Life stage of the malaria infection at which the experiment reported in the literature was performed: B, blood stages; L, liver stage; M, mosquito stage. The experimental description of the gene is defined in parenthesis: (e) essential; (gr) growth reducing; (ne) non-essential. ³Impact of the gene knock-out on the metabolic tasks of iPfa is explainable based on the annotated function of the gene. ⁴This gene is involved in an essential pathway and is thus presumed to be essential. ⁵None of the reactions associated with this gene is single essential; the gene is essential when two or more associated reactions are knocked out. ⁶We define *similarity* between the essential enzymes in iPfa and the human proteome when there exists at least one matching protein in non-redundant human proteome (txid9606) with the standard protein BLAST settings in the bioinformatics software blastp version 2.6.0+ [65]: cutoff value 1E-10, matrix BLOSUM62 [66].

(XLSX)

S3 Table. Synthetic lethal pairs in iPfa. ¹Pairs of genes predicted as essential for iPfa growth with TFA and the MS data (2012, Vo Duy *et al.*) in the *in silico* rich medium (composed of 236 substrates, [S1 Methods](#)).

(XLSX)

S4 Table. List of reactions that become unidirectional with TFA. ¹Reaction I.D. in iPfa. Cellular compartments are defined after “_” in the reaction I.D.: (_c) cytosol, (_m) mitochondrion, (_r) endoplasmic reticulum, (_e) extracellular. Transport reactions are marked with T_. ²Values of the standard Gibbs free energy of reaction ($\Delta G_o'$) calculated with the group contribution method (GCM), as explained in the [S1 Methods](#). Note that $\Delta G_o'$ does not determine the reaction directionality. ³Reaction directionalities determined with Thermodynamic Variability Analysis (TVA). ⁴The reactions R02057_r, R01468_r, R02038_r, and T_c_to_r_C00189 are unidirectional (forwards) using TVA and generic concentration ranges, or blocked using TVA and metabolomics data (see [Table 2](#)).

(XLSX)

S5 Table. Single essentiality predictions in iPfa among the alternative *in silico* minimal media (IMMs). ¹Genes predicted as essential for iPfa growth with TFA and the MS data (2012, Vo Duy *et al.* [21]) in the alternative IMMs (10032 alternative IMMs, each composed of 23 substrates, see [Table 2](#)). ²Genes commonly essential in all alternative IMMs. ³If more than one reaction is associated with the gene, one enzyme name was chosen.

(XLSX)

S6 Table. Genes of *P. falciparum* and associations with K-Orthology groups obtained with the RAVEN Toolbox. ¹Gene IDs were obtained from PlasmoDB. ²K-ID or Orthology groups were obtained from the KEGG database. ³The E-value denotes the goodness of the annotation to the assigned metabolic function. The annotation process is described in detail in [S1 Methods](#).

(XLSX)

S7 Table. Localization evidence and scores for the genes in the draft metabolic network of iPfa. ¹Localizations from the ApiLoc database. Intracellular compartments are defined by

letters, such as c, cytosol; m, mitochondrion; a, apicoplast; er, endoplasmic reticulum; fv, food vacuole; e, extracellular space/parasitophorous vacuole; n, nucleus; ag, Golgi apparatus. (XLSX)

S1 Dataset. Ranking of metabolites based on the *Reduction of Uncertainty (RoU) in the Gibbs free energy of the reactions ($\Delta_r G'$)*. The concentration range of each metabolite was integrated separately in iPfa following two approaches. In the first approach, the concentration of the metabolite was integrated at-a-time in all compartments of iPfa where the metabolite appears (sheet S1 A and sheet S1 B in S1 Dataset). In the second approach, the concentration of the metabolite was integrated separately in each compartment of iPfa where the metabolite appears (sheet S1 C and sheet S1 D in S1 Dataset). The RoU in the $\Delta_r G'$ was calculated in both approaches. The reactions in iPfa were ranked based on two ranking criteria: the number of reactions impacted by each metabolite (sheet S1 A and sheet S1 C in S1 Dataset), and the global RoU (sheet S1 B and sheet S1 D in S1 Dataset) as defined in the [S1 Methods](#). Note: we consider that a metabolite impacts a reaction if the RoU in $\Delta_r G'$ is higher than 0.1%. (XLSX)

Acknowledgments

We thank Dr. Rebecca Oppenheim for her valuable contribution in the compartmentalization process of iPfa and the careful review of the manuscript. We acknowledge the MalarX collaborators for their feedback in the reconstruction process of iPfa and its analysis.

Author Contributions

Conceptualization: ACP ST MA VH.

Formal analysis: ACP ST MA VH.

Investigation: ACP ST MA DSF VH.

Methodology: ACP ST MA VH.

Project administration: VH.

Software: ACP MA.

Supervision: VH.

Validation: ACP ST DSF VH.

Writing – original draft: ACP VH.

Writing – review & editing: ACP ST MA DSF VH.

References

1. World Health Organization. World Malaria Report, 2015. Geneva, Switzerland: WHO, 2016.
2. Bosi E, Monk JM, Aziz RK, Fondi M, Nizet V, Palsson BO. Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. *Proc Natl Acad Sci USA*. 2016 Jun 28; 113(26):E3801–9. doi: [10.1073/pnas.1523199113](https://doi.org/10.1073/pnas.1523199113) PMID: [27286824](https://pubmed.ncbi.nlm.nih.gov/27286824/)
3. Tymoshenko S, Oppenheim RD, Agren R, Nielsen J, Soldati-Favre D, Hatzimanikatis V. Metabolic Needs and Capabilities of *Toxoplasma gondii* through Combined Computational and Experimental Analysis. *PLoS Comput Biol*. 2015; 11(5):1–28.

4. Famili I, Forster J, Nielson J, Palsson BO. *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc Natl Acad Sci USA*. 2003 Nov 11; 100(23):13134–9. doi: [10.1073/pnas.2235812100](https://doi.org/10.1073/pnas.2235812100) PMID: [14578455](https://pubmed.ncbi.nlm.nih.gov/14578455/)
5. Plata G, Hsiao T-L, Olszewski KL, Llinás M, Vitkup D. Reconstruction and flux-balance analysis of the *Plasmodium falciparum* metabolic network. *Mol Syst Biol*. 2010; 6(408).
6. Huthmacher C, Hoppe A, Bulik S, Holzhütter H-G. Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst Biol*. 2010; 4(120).
7. Bazzani S, Hoppe A, Holzhütter H-G. Network-based assessment of the selectivity of metabolic drug targets in *Plasmodium falciparum* with respect to human liver metabolism. *BMC Syst Biol*. 2012; 6(118).
8. Fang X, Reifman J, Wallqvist A. Modeling metabolism and stage-specific growth of *Plasmodium falciparum* HB3 during the intraerythrocytic developmental cycle. *Mol Biosyst*. 2014; 10(10):2526–37. doi: [10.1039/c4mb00115j](https://doi.org/10.1039/c4mb00115j) PMID: [25001103](https://pubmed.ncbi.nlm.nih.gov/25001103/)
9. Varma A, Palsson BO. Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. *Nat Biotechnol*. 1994 Oct; 12(10):994–8.
10. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? *Nat Biotechnol*. 2010; 28(3):245–8. doi: [10.1038/nbt.1614](https://doi.org/10.1038/nbt.1614) PMID: [20212490](https://pubmed.ncbi.nlm.nih.gov/20212490/)
11. Kitano H. Systems biology: a brief overview. *Science*. 2002; 295(5560):1662–4. doi: [10.1126/science.1069492](https://doi.org/10.1126/science.1069492) PMID: [11872829](https://pubmed.ncbi.nlm.nih.gov/11872829/)
12. Henry CS, Jankowski MD, Broadbelt LJ, Hatzimanikatis V. Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biophys J*. 2006; 90(4):1453–61. doi: [10.1529/biophysj.105.071720](https://doi.org/10.1529/biophysj.105.071720) PMID: [16299075](https://pubmed.ncbi.nlm.nih.gov/16299075/)
13. Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-based metabolic flux analysis. *Biophys J*. 2007; 92(5):1792–805. doi: [10.1529/biophysj.106.093138](https://doi.org/10.1529/biophysj.106.093138) PMID: [17172310](https://pubmed.ncbi.nlm.nih.gov/17172310/)
14. Soh KC, Hatzimanikatis V. Network thermodynamics in the post-genomic era. *Curr Opin Microbiol*. 2010; 13(3):350–7. doi: [10.1016/j.mib.2010.03.001](https://doi.org/10.1016/j.mib.2010.03.001) PMID: [20378394](https://pubmed.ncbi.nlm.nih.gov/20378394/)
15. Soh KC, Hatzimanikatis V. Constraining the flux space using thermodynamics and integration of metabolomics data. In: Krömer JO, Nielsen LK, Blank LM, editors. *Metabolic Flux Analysis*. 1191. New York: Springer Protocols; 2014. p. 49–63.
16. Ataman M, Hatzimanikatis V. Heading in the right direction: thermodynamics-based network analysis and pathway engineering. *Curr Opin Biotechnol*. 2015; 36:176–82. doi: [10.1016/j.copbio.2015.08.021](https://doi.org/10.1016/j.copbio.2015.08.021) PMID: [26360871](https://pubmed.ncbi.nlm.nih.gov/26360871/)
17. Chakrabarti A, Miskovic L, Soh KC, Hatzimanikatis V. Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints. *Biotechnol J*. 2013; 8(9):1043–57. doi: [10.1002/biot.201300091](https://doi.org/10.1002/biot.201300091) PMID: [23868566](https://pubmed.ncbi.nlm.nih.gov/23868566/)
18. Spivey HO, Ovadi J. Substrate channeling. *Methods*. 1999 Oct; 19(2):306–21. doi: [10.1006/meth.1999.0858](https://doi.org/10.1006/meth.1999.0858) PMID: [10527733](https://pubmed.ncbi.nlm.nih.gov/10527733/)
19. Teng R, Junankar PR, Bubb Wa, Rae C, Mercier P, Kirk K. Metabolite profiling of the intraerythrocytic malaria parasite *Plasmodium falciparum* by (1)H NMR spectroscopy. *NMR Biomed*. 2009; 22(3):292–302. doi: [10.1002/nbm.1323](https://doi.org/10.1002/nbm.1323) PMID: [19021153](https://pubmed.ncbi.nlm.nih.gov/19021153/)
20. Teng R, Lehane Adele M, Winterberg M, Shafik Sarah H, Summers Robert L, Martin Rowena E, et al. 1H-NMR metabolite profiles of different strains of *Plasmodium falciparum*. *Biosci Rep*. 2014; 34(6):685–99.
21. Vo Duy S, Besteiro S, Berry L, Perigaud C, Bressolle F, Vial HJ, et al. A quantitative liquid chromatography tandem mass spectrometry method for metabolomic analysis of *Plasmodium falciparum* lipid related metabolites. *Anal Chim Acta*. 2012; 739:47–55. doi: [10.1016/j.aca.2012.06.016](https://doi.org/10.1016/j.aca.2012.06.016) PMID: [22819049](https://pubmed.ncbi.nlm.nih.gov/22819049/)
22. Sanz S, Bandini G, Ospina D, Bernabeu M, Mariño K, Fernández-Becerra C, et al. Biosynthesis of GDP-fucose and other sugar nucleotides in the blood stages of *Plasmodium falciparum*. *J Biol Chem*. 2013; 288(23):16506–17. doi: [10.1074/jbc.M112.439828](https://doi.org/10.1074/jbc.M112.439828) PMID: [23615908](https://pubmed.ncbi.nlm.nih.gov/23615908/)
23. Thiele I, Palsson BO. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc*. 2010; 5(1):93–121. doi: [10.1038/nprot.2009.203](https://doi.org/10.1038/nprot.2009.203) PMID: [20057383](https://pubmed.ncbi.nlm.nih.gov/20057383/)
24. Agren R, Liu L, Shoae S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. *PLoS Comput Biol*. 2013; 9(3):e1002980. doi: [10.1371/journal.pcbi.1002980](https://doi.org/10.1371/journal.pcbi.1002980) PMID: [23555215](https://pubmed.ncbi.nlm.nih.gov/23555215/)
25. Oppenheim RD, Creek DJ, Macrae JI, Modrzyńska KK, Pino P, Limenitakis J, et al. BCKDH: the missing link in apicomplexan mitochondrial metabolism is required for full virulence of *Toxoplasma gondii* and *Plasmodium berghei*. *PLoS Pathog*. 2014; 10(7):e1004263. doi: [10.1371/journal.ppat.1004263](https://doi.org/10.1371/journal.ppat.1004263) PMID: [25032958](https://pubmed.ncbi.nlm.nih.gov/25032958/)

26. Naik RS, Branch OH, Woods AS, Vijaykumar M, Perkins DJ, Nahlen BL, et al. Glycosylphosphatidylinositol anchors of *Plasmodium falciparum*: molecular characterization and naturally elicited antibody response that may provide immunity to malaria pathogenesis. *J Exp Med*. 2000 Dec 04; 192(11):1563–76. PMID: [11104799](#)
27. Jensen MD, Conley M, Helstowski LD. Culture of *Plasmodium falciparum*: the role of pH, glucose, and lactate. *J Parasitol*. 1983; 69(6):1060–7. PMID: [6371212](#)
28. Martin RE, Kirk K. Transport of the essential nutrient isoleucine in human erythrocytes infected with the malaria parasite *Plasmodium falciparum*. *Blood*. 2007; 109(5):2217–24. doi: [10.1182/blood-2005-11-026963](#) PMID: [17047158](#)
29. Yeh E, DeRisi JL. Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage *Plasmodium falciparum*. *PLoS Biol*. 2011; 9(8):e1001138. doi: [10.1371/journal.pbio.1001138](#) PMID: [21912516](#)
30. Jomaa H, Wiesner J, Sanderbrand S, Altinccek B, Weidemeyer C, Hintz M, et al. Inhibitors of the non-mevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science*. 1999 Sep 03; 285(5433):1573–6. PMID: [10477522](#)
31. Cassera MB, Merino EF, Peres VJ, Kimura EA, Wunderlich G, Katzin AM. Effect of fosmidomycin on metabolic and transcript profiles of the methylerythritol phosphate pathway in *Plasmodium falciparum*. *Mem Inst Oswaldo Cruz*. 2007 Jun; 102(3):377–83. PMID: [17568945](#)
32. Odom AR, Van Voorhis WC. Functional genetic analysis of the *Plasmodium falciparum* deoxyxylulose 5-phosphate reductoisomerase gene. *Mol Biochem Parasitol*. 2010 Apr; 170(2):108–11. Pubmed Central PMCID: 2814890. doi: [10.1016/j.molbiopara.2009.12.001](#) PMID: [20018214](#)
33. Imlay LS, Armstrong CM, Masters MC, Li T, Price KE, Edwards RL, et al. *Plasmodium* IspD (2-C-Methyl-D-erythritol 4-Phosphate Cytidyltransferase), an Essential and Druggable Antimalarial Target. *ACS Infect Dis*. 2015 Apr 10; 1(4):157–67. Pubmed Central PMCID: 4714788. doi: [10.1021/id500047s](#) PMID: [26783558](#)
34. Gerold P, Schwarz RT. Biosynthesis of glycosphingolipids de-novo by the human malaria parasite *Plasmodium falciparum*. *Mol Biochem Parasitol*. 2001 Jan 15; 112(1):29–37. PMID: [11166384](#)
35. Bello AM, Poduch E, Fujihashi M, Amani M, Li Y, Crandall I, et al. A potent, covalent inhibitor of orotidine 5'-monophosphate decarboxylase with antimalarial activity. *J Med Chem*. 2007 Mar 8; 50(5):915–21. doi: [10.1021/jm060827p](#) PMID: [17290979](#)
36. Roggero R, Zufferey R, Minca M, Richier E, Calas M, Vial H, et al. Unraveling the mode of action of the antimalarial choline analog G25 in *Plasmodium falciparum* and *Saccharomyces cerevisiae*. *Antimicrob Agents Chemother*. 2004 Aug; 48(8):2816–24. doi: [10.1128/AAC.48.8.2816-2824.2004](#) PMID: [15273086](#)
37. Krnajski Z, Gilberger TW, Walter RD, Cowman AF, Muller S. Thioredoxin reductase is essential for the survival of *Plasmodium falciparum* erythrocytic stages. *J Biol Chem*. 2002 Jul 19; 277(29):25970–5. doi: [10.1074/jbc.M203539200](#) PMID: [12004069](#)
38. Tyagi C, Bathke J, Goyal S, Fischer M, Dahse HM, Chacko S, et al. Targeting the intersubunit cavity of *Plasmodium falciparum* glutathione reductase by a novel natural inhibitor: computational and experimental evidence. *Int J Biochem Cell Biol*. 2015 Apr; 61:72–80. doi: [10.1016/j.biocel.2015.01.014](#) PMID: [25660424](#)
39. Naidoo K, Coetzer TL. Reduced glycerol incorporation into phospholipids contributes to impaired intra-erythrocytic growth of glycerol kinase knockout *Plasmodium falciparum* parasites. *Biochim Biophys Acta*. 2013 Nov; 1830(11):5326–34. doi: [10.1016/j.bbagen.2013.08.006](#) PMID: [23954205](#)
40. Otto TD, Wilinski D, Assefa S, Keane TM, Sarry LR, Bohme U, et al. New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Molecular Microbiology*. 2010 Apr; 76(1):12–24. doi: [10.1111/j.1365-2958.2009.07026.x](#) PMID: [20141604](#)
41. Déchamps S, Wengelnik K, Berry-Sterkers L, Cerdan R, Vial HJ, Gannoun-Zaki L. The Kennedy phospholipid biosynthesis pathways are refractory to genetic disruption in *Plasmodium berghei* and therefore appear essential in blood stages. *Mol Biochem Parasitol*. 2010; 173(2):69–80. doi: [10.1016/j.molbiopara.2010.05.006](#) PMID: [20478340](#)
42. van Hellemond JJ, Slot JW, Geelen MJ, van Golde LM, Vermeulen PS. Ultrastructural localization of CTP:phosphoethanolamine cytidyltransferase in rat liver. *J Biol Chem*. 1994 Jun 3; 269(22):15415–8. PMID: [8195180](#)
43. Divo AA, Geary TG, Davis NL, Jensen JB. Nutritional requirements of *Plasmodium falciparum* in culture. I. Exogenously supplied dialyzable components necessary for continuous growth. *J Protozool*. 1985; 32(1):59–64. PMID: [3886898](#)
44. Asahi H, Kanazawa T, Kajihara Y, Takahashi K, Takahashi T. Hypoxanthine: a low molecular weight factor essential for growth of erythrocytic *Plasmodium falciparum* in a serum-free medium. *Parasitology*. 1996 Jul; 113 (Pt 1):19–23.

45. O'Hara JK, Kerwin LJ, Cobbold Sa, Tai J, Bedell Ta, Reider PJ, et al. Targeting NAD⁺ metabolism in the human malaria parasite *Plasmodium falciparum*. *PLoS One*. 2014; 9(4):e94061. doi: [10.1371/journal.pone.0094061](https://doi.org/10.1371/journal.pone.0094061) PMID: [24747974](https://pubmed.ncbi.nlm.nih.gov/24747974/)
46. Desai SA. Insights gained from *P. falciparum* cultivation in modified media. *ScientificWorldJournal*. 2013; 2013:363505. doi: [10.1155/2013/363505](https://doi.org/10.1155/2013/363505) PMID: [23956690](https://pubmed.ncbi.nlm.nih.gov/23956690/)
47. Pillai AD, Nguitragool W, Lyko B, Dolinta K, Butler MM, Nguyen ST, et al. Solute restriction reveals an essential role for clag3-associated channels in malaria parasite nutrient acquisition. *Mol Pharmacol*. 2012 Dec; 82(6):1104–14. Pubmed Central PMCID: 3502622. doi: [10.1124/mol.112.081224](https://doi.org/10.1124/mol.112.081224) PMID: [22949525](https://pubmed.ncbi.nlm.nih.gov/22949525/)
48. Mi-Ichi F, Kano S, Mitamura T. Oleic acid is indispensable for intraerythrocytic proliferation of *Plasmodium falciparum*. *Parasitology*. 2007 Nov; 134:1671–7. doi: [10.1017/S0031182007003137](https://doi.org/10.1017/S0031182007003137) PMID: [17610764](https://pubmed.ncbi.nlm.nih.gov/17610764/)
49. Mi-ichi F, Kita K, Mitamura T. Intraerythrocytic *Plasmodium falciparum* utilize a broad range of serum-derived fatty acids with limited modification for their growth. *Parasitology*. 2006 Oct; 133:399–410. doi: [10.1017/S0031182006000540](https://doi.org/10.1017/S0031182006000540) PMID: [16780611](https://pubmed.ncbi.nlm.nih.gov/16780611/)
50. Trager W. Malaria parasites (*Plasmodium lophurae*) developing extracellularly in vitro: incorporation of labeled precursors. *J Protozool*. 1971 Aug; 18(3):392–9. PMID: [5132315](https://pubmed.ncbi.nlm.nih.gov/5132315/)
51. Geary TG, Divo AA, Bonanni LC, Jensen JB. Nutritional requirements of *Plasmodium falciparum* in culture. III. Further observations on essential nutrients and antimetabolites. *J Protozool*. 1985; 32(4):608–13. PMID: [2866244](https://pubmed.ncbi.nlm.nih.gov/2866244/)
52. Cassera MB, Zhang Y, Hazleton KZ, Schramm VL. Purine and pyrimidine pathways as targets in *Plasmodium falciparum*. *Curr Top Med Chem*. 2011 Aug; 11(16):2103–15. PMID: [21619511](https://pubmed.ncbi.nlm.nih.gov/21619511/)
53. Krungkrai SR, Wutipraditkul N, Krungkrai J. Dihydroorotase of human malarial parasite *Plasmodium falciparum* differs from host enzyme. *Biochem Biophys Res Commun*. 2008 Feb 15; 366(3):821–6. doi: [10.1016/j.bbrc.2007.12.025](https://doi.org/10.1016/j.bbrc.2007.12.025) PMID: [18082626](https://pubmed.ncbi.nlm.nih.gov/18082626/)
54. Christopherson RI, Jones ME. The overall synthesis of L-5,6-dihydroorotate by multienzymatic protein pyr1-3 from hamster cells. Kinetic studies, substrate channeling, and the effects of inhibitors. *J Biol Chem*. 1980; 255(23):11381–95. PMID: [6108323](https://pubmed.ncbi.nlm.nih.gov/6108323/)
55. Zhang PF, Martin PD, Purcarena C, Vaishnav A, Brunzelle JS, Fernando R, et al. Dihydroorotase from the Hyperthermophile *Aquifex aeolicus* Is Activated by Stoichiometric Association with Aspartate Transcarbamoylase and Forms a One-Pot Reactor for Pyrimidine Biosynthesis. *Biochemistry*. 2009 Feb 3; 48(4):766–78. doi: [10.1021/bi801831r](https://doi.org/10.1021/bi801831r) PMID: [19128030](https://pubmed.ncbi.nlm.nih.gov/19128030/)
56. Zhang PF, Martin PD, Purcarena C, Vaishnav A, Brunzelle JS, Fernando R, et al. Dihydroorotase from the hyperthermophile *Aquifex aeolicus* is activated by stoichiometric association with aspartate transcarbamoylase and forms a one-pot reactor for pyrimidine biosynthesis. *Biochemistry*. 2009 Feb 3; 48(4):766–78. doi: [10.1021/bi801831r](https://doi.org/10.1021/bi801831r) PMID: [19128030](https://pubmed.ncbi.nlm.nih.gov/19128030/)
57. Gomes Ana R, Bushell E, Schwach F, Girling G, Anar B, Quail Michael A, et al. A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. *Cell Host Microbe*. 2015; 17(3):404–13. doi: [10.1016/j.chom.2015.01.014](https://doi.org/10.1016/j.chom.2015.01.014) PMID: [25732065](https://pubmed.ncbi.nlm.nih.gov/25732065/)
58. Schwach F, Bushell E, Gomes AR, Anar B, Girling G, Herd C, et al. PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites. *Nucleic Acids Res*. 2015 Jan; 43(Database issue):D1176–82. doi: [10.1093/nar/gku1143](https://doi.org/10.1093/nar/gku1143) PMID: [25593348](https://pubmed.ncbi.nlm.nih.gov/25593348/)
59. Graewe S, Stanway RR, Rennenberg A, Heussler VT. Chronicle of a death foretold: *Plasmodium* liver stage parasites decide on the fate of the host cell. *FEMS Microbiol Rev*. 2012 Jan; 36(1):111–30. doi: [10.1111/j.1574-6976.2011.00297.x](https://doi.org/10.1111/j.1574-6976.2011.00297.x) PMID: [22092244](https://pubmed.ncbi.nlm.nih.gov/22092244/)
60. PlasmoDB. 2016 [cited 2014–2015]. Available from: <http://plasmodb.org/plasmo/>.
61. Jankowski MD, Henry CS, Broadbelt LJ, Hatzimanikatis V. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys J*. 2008; 95(3):1487–99. Pubmed Central PMCID: PMC2479599. doi: [10.1529/biophysj.107.124784](https://doi.org/10.1529/biophysj.107.124784) PMID: [18645197](https://pubmed.ncbi.nlm.nih.gov/18645197/)
62. Jol SJ, Kümmel A, Hatzimanikatis V, Beard Da, Heinemann M. Thermodynamic calculations for biochemical transport and reaction processes in metabolic networks. *Biophys J*. 2010; 99(10):3139–44. Pubmed Central PMCID: PMC2980713. doi: [10.1016/j.bpj.2010.09.043](https://doi.org/10.1016/j.bpj.2010.09.043) PMID: [21081060](https://pubmed.ncbi.nlm.nih.gov/21081060/)
63. Soh KC, Miskovic L, Hatzimanikatis V. From network models to network responses: integration of thermodynamic and kinetic properties of yeast genome-scale metabolic networks. *FEMS Yeast Res*. 2012 Mar; 12(2):129–43. doi: [10.1111/j.1567-1364.2011.00771.x](https://doi.org/10.1111/j.1567-1364.2011.00771.x) PMID: [22129227](https://pubmed.ncbi.nlm.nih.gov/22129227/)
64. Tymoshenko S, Oppenheim RD, Soldati-Favre D, Hatzimanikatis V. Functional genomics of *Plasmodium falciparum* using metabolic modelling and analysis. *Brief Funct Genomics*. 2013 Jul; 12(4):316–27. Pubmed Central PMCID: 3743259. doi: [10.1093/bfgp/elt017](https://doi.org/10.1093/bfgp/elt017) PMID: [23793264](https://pubmed.ncbi.nlm.nih.gov/23793264/)

65. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997 Sep 01; 25(17):3389–402. Pubmed Central PMCID: 146917. PMID: [9254694](#)
66. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, et al. Protein database searches using compositionally adjusted substitution matrices. *Febs Journal.* 2005 Oct; 272(20):5101–9. doi: [10.1111/j.1742-4658.2005.04945.x](#) PMID: [16218944](#)
67. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng.* 2003 Oct; 5(4):264–76. PMID: [14642354](#)