Chapter 13

Identifying Antimalarial Drug Targets by Cellular Network Analysis

Kitiporn Plaimas and Rainer König

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/65432

Abstract

Malaria is one of the most deadly parasitic infectious diseases and identifying novel drug targets is mandatory for the development of new drugs. To find drug targets, metabolic and signaling networks have been constructed. These networks have been investigated by graph theoretical methods. Furthermore, mechanistic models have been set up based on stoichiometric equations. At equilibrium, production and consumption of internal metabolites need to be balanced leading to a large set of flux equations, and this can be used for metabolic flux simulations to identify drug targets. Analysis of flux variability and knockout simulations were applied to detect potential drug targets whose absence reduces the predicted biomass production and hence viability of the parasite in the host cell. Furthermore, not only the parasite was studied, but also the interaction between the host and the parasite, and, based on experimental expression data, stage-specific metabolic models of the parasite were developed, particularly during the red-blood cell stage. In this chapter, these various network-based approaches for drug target prediction will be explained and summarized.

Keywords: network-based analysis, drug targets, flux balance analysis, malaria

1. Introduction

Network-based analysis has become an important tool in biomedical research. It facilitates the investigation and understanding of a system as a whole, not only its single components. For this, first the networks need to be constructed and then investigated employing different analysis or modeling techniques. According to the applied methodological approaches to analyze these networks, one may distinguish cellular network models for signal transduction, gene regulation and metabolism. The network constructions based on information are compiled from databases and are assembled in an automated way often followed by manual refinement. Network-based models have been applied to study the cellular mechanisms of a large variety of diseases elucidating, for example, tumor growth, malfunctioning

of the differentiation of immune cells, or identifying drug targets of invasive pathogens [1, 2]. To find drug targets for the treatment of malaria, metabolic and signaling networks have been constructed and intensively investigated. This chapter will introduce the reader into the basic principles of constructing and applying such cellular networks. It then leads through the application of these systems biology approaches to predict drug targets followed by a small section exemplarily showing an experimental validation for these predictions.

2. Construction of cellular networks

Proteins are involved in all cellular functions. These cellular processes can be put up as cellular networks, which describe associations among these proteins and other cellular compounds such as metabolites and nucleic acids. These cellular networks can conceptually be divided into three distinct parts: the cell signaling, the transcriptional regulatory network, and the metabolic network. The best observed and modeled network is the metabolic network while the complex system of signal transduction is rather captured statistically investigating the experimental information about proteins and their expressed genes of network models basing on protein-protein interactions [3]. The transcriptional regulatory network links transcriptional regulators to their target genes [4]. The simplest form of a network is a network represented by an undirected graph $G = (V, E)$ consisting of nodes V and edges (connections, links) E between these nodes. Each node $i \in V$ represents a unique cellular entity such as enzymes, genes, and proteins, while each edge $(i, j) \in E$ represents an observed interaction between two nodes i and j. A metabolic network model can be constructed as a bipartite graph consisting of two disjoint sets of nodes (reaction and metabolite nodes, see **Figure 1**) [5]. The direction of edges in the metabolic networks is given by the flux from the substrate to the product of a biochemical reaction. An edge indicates that a metabolite is either a substrate or a product of a reaction. The distinction between substrates and products of a reaction is only possible if the graph is directed, that is, if the set of edges E consists of ordered pairs of vertices. This distinction is often useful when modeling metabolic fluxes but may be neglected in simpler models [6]. As a bipartite graph, the metabolic network can be represented as an adjacency matrix of $m \times n$ dimensions, where m is the number of metabolites and n is the number of reactions. More specific models of metabolic networks concerning the stoichiometry can also be represented as an adjacency matrix using stoichiometric coefficients of chemical reactions as weights for the edges between metabolites and reactions. As shown in **Figure 1**, our small example network consists of three reactions (R1, R2, and R3) and six metabolites (A, B, C, D, E, and F):

R1:	A	⇔	B
R2:	2 B + C	→	E + F
R3:	2 E	→	B + D

http://dx.doi.org/10.5772/65432

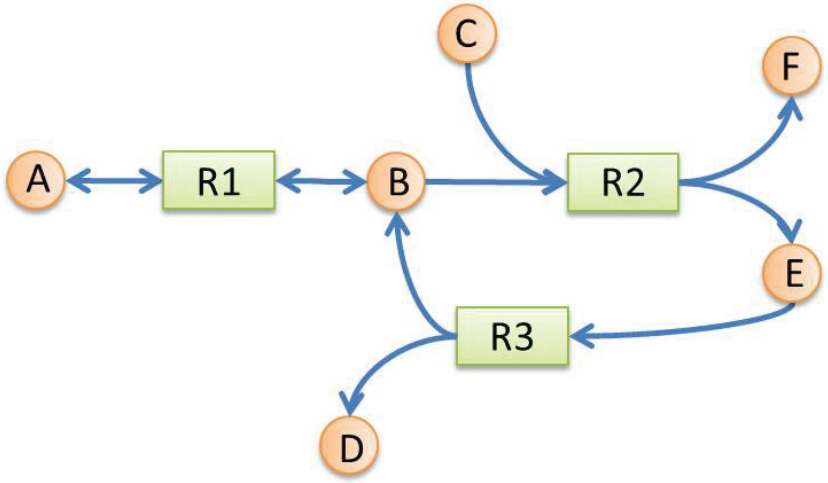

Figure 1. Graphical view of a metabolic network model as a bipartite graph consisting of two disjoint sets of nodes (reactions and metabolites). This network consists of three reactions (R1, R2 and R3) and six metabolites (A, B, C, D, E, F). R1 is a reversible reaction, the other reactions are irreversible.

The stoichiometric matrix or the adjacency matrix containing stoichiometric coefficients of each reaction equation is

$$s = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

where the rows correspond to metabolites A, B, C, D, E, and F, and the columns correspond to reactions R1, R2, and R3, respectively. R1 is a reversible reaction. Metabolic networks for *Plasmodium spp* can be constructed using the databases PlasmoCyc [7], Malaria Parasite Metabolic Pathways (MPMP) [8], The Kyoto Encyclopedia of Genes and Genomes (KEGG), http://www.genome.jp/kegg/, and from models in the literature [9]. Unspecific compounds such as water, ATP, ADP, etc., may be discarded for these rather general models but need to be considered for more detailed models when, for example, employing flux balance analysis (see below). Cellular networks can be analyzed mechanistically or statistically by their topological features. In the following, we explain briefly some of these topological features.

3. Topological features for statistical analyses of cellular networks

Several computational techniques have been developed to identify essential genes and drug targets *in silico* for a therapy against malaria. To construct an undirected graph for metabolism,

the network representation of a reaction-pair network can be used instead of a bipartite graph. In this representation, enzymes are linked if there is at least one metabolite, which is produced by one of the enzymes and which serves as a substrate for the other. For these simple networks, the network topology can be described by characteristic properties. Similarly, protein interaction networks can be analyzed to get specific characteristics for signal transduction [3, 10, 11]. These characteristics either hint directly to essential genes (serving as drug targets) or can be used when comparing the full network with a network in which one of the nodes (enzymes or signaling proteins) is targeted by a drug.

3.1. Diameter and density of a network

The diameter of a network is the largest distance of all shortest paths between two nodes (reactions, signaling molecules) in the network. The density of a network is the ratio of the edges (links, connections) between two reactions divided by all possible edges of all reactions. These two properties can be used to determine the robustness of a network. In recent studies, a reaction was said to be essential if the mutated or targeted network showed a larger diameter after removing the reaction [12, 13].

3.2. Scale-freeness of networks

Networks can be distinguished by their degree distributions where the degree of a node $v \in V$ is defined as the number of edges between v and its adjacent nodes. Many degree distributions of naturally occurring networks follow power laws [6] $P(k) \sim k^{-\gamma}$ where $\gamma > 0$ is a constant depending on the network and is usually in the range of 2–3. P is the probability to draw a node with degree k. Networks with a power law distribution are also called *scale-free* networks [6, 14]. Basically, these scale-free networks consist of few highly connected vertices, so-called *hubs*, and many less connected vertices [15]. Most real-world networks including metabolic networks are approximately scale-free networks [6]. **Figure 2** shows the degree distribution of the metabolic network of *Plasmodium falciparum,* which is fitted by a power-law distribution. Scale-free networks generally have a small diameter [16], as in particular the highly connected nodes connect nodes within only a few links. Additionally, these networks are highly connected [17]. The benefit of such a highly connected and scale-free architecture is its robustness against single "attacks," that is, a failure of a single node in the system, as it is statistically more probable that vertices with lower degrees are hit from which the general structure of the network is not affected. The scale-free topology provides robustness to the network with increases flexibility to random perturbations where the loss of individual nodes usually has no effect on the overall network topology. Nevertheless, such a network is susceptible to targeted attacks at highly connected critical hubs [18], and mutations affecting hubs are more likely to cause a defect [17].

3.3. Clustering coefficient

The clustering coefficient is used to estimate the local density of links (edges) in the network. It describes the connectedness among neighbors and helps to estimate the probability of local alternative paths of signaling or metabolic fluxes (e.g., after targeting). The clustering coefficient of a node v is defined as the ratio of the number of connecting edges among all neighbors

http://dx.doi.org/10.5772/65432

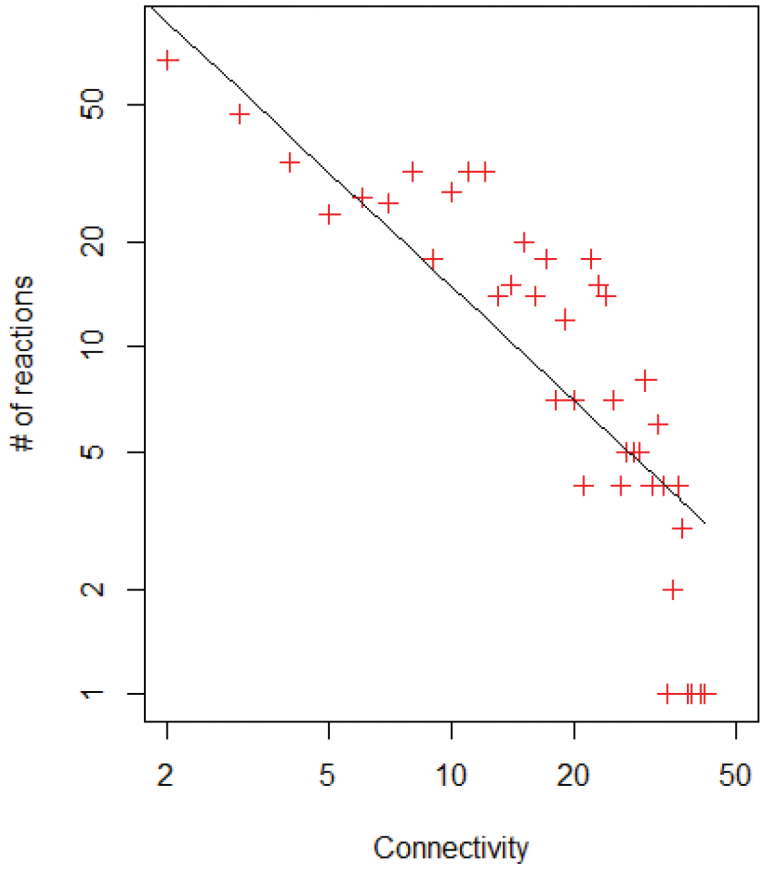

Figure 2. Degree distribution of the metabolic network of *P. falciparum* using the (most suitable) network of [13].

of v and the total number of edges among them that could be possible. This means, if all neighbors are connected among themselves, the clustering coefficient becomes one, if none of the neighbors is connected with any other neighbor, it is zero [6, 19, 20]. In **Figure 3**, the observed reaction in dark has three neighbors and two edges among its neighbors. Having three neighbors, there are six possible connections among neighbors. Thus, the clustering coefficient of the observed reaction in this example can be computed as $\frac{2}{6} = \frac{1}{3}$.

3.4. Centrality

Descriptors for *node centrality* are quite powerful for describing the potential of essentiality of a node. They describe not only the impact of the node to its direct vicinity but also the contribution of a node to the global structure of the network. The simplest of all centrality measures is the connectivity, or degree k, which is just the number of links connecting the node with other nodes. In a cellular network, the degree is commonly used to describe an important node as it

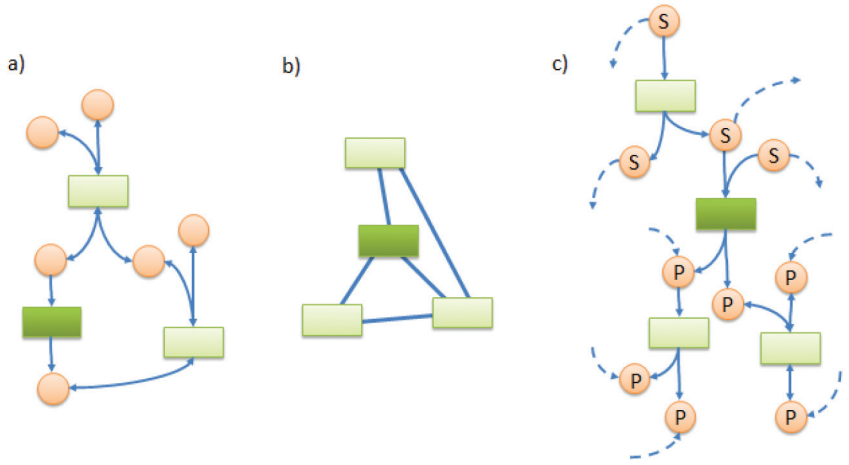

Figure 3. Illustration of the concepts of the topology features. Circles represent metabolites, rectangles reactions, arrows directions of the metabolic flux, lines represent links between two neighboring reactions and dark rectangles represent the investigated reactions. (a) The observed reaction is a chokepoint because it is the only reaction consuming the upstream metabolite. (b) The metabolic network in a reaction-pair representation for computing the clustering coefficient. The observed reaction has three neighbors (degree of 3) and there are two links among these neighbors. Therefore, the clustering coefficient for this observed reaction is 1/3. (c) Graphical illustration of the way to compute producibility of the observed reaction from its substrates (S) to its products (P). Possible alternative pathways to consume substrates S for producing products P are represented by dashed arrows. The percentage of the products that can be produced from the substrates is the producibility of the observed reaction.

is known that often essential genes are nodes in the network with a high degree (so-called hubs). Another commonly used centrality measure is betweenness centrality. Betweenness centrality is the frequency of a node to be part of the shortest paths connecting all pairs of nodes in the network [21].

3.5. Choke points and load points

In metabolic networks, Samal *et al*. found out that most reactions identified as essential are involved in the consumption or production of metabolites with low connectivity [22]. This is because these nodes are more likely to be the limiting factor for consuming or producing these metabolites. In the extreme case, an enzyme is the only enzyme, which consumes or produces a certain compound. Blocking such a reaction may cause severe effects to the cell as, for example, it may cause an assembly of toxic compounds, which cannot be degraded anymore or a lack of substrates for important processes further downstream of the enzyme. Hence, a choke point reaction was defined as a reaction that uniquely consumes or produces a certain metabolite in the metabolic network [23, 24]. This concept has been successfully applied to identify drug targets for *Plasmodium spp.* [24, 25]. Load scores are defined as hot spots in the metabolic network (enzymes or metabolites) based on the ratio of the number of shortest paths (connecting any two enzymes or metabolites in the whole network) passing through a metabolite or enzyme and the number of nearest neighbor links [23].

3.6. Producibility (by deviations)

A reaction is determined to be potentially essential when basically the mutated network cannot yield the products of the reaction from upstream substrates of the reaction using other pathways linking the substrates to the products (see **Figure 3**). The percentage of the products that can be produced from the substrates, the so-called "producibility," can be used to examine the essentiality of the observed reaction [13].

3.7. Applying these topology-based methods to predict drug targets for *Plasmodium* spp.

The concept of choke points and load points was successfully applied to estimate the essentiality of an enzyme in *Plasmodium* [23, 24, 26]. Yeh *et al*. initially applied a chokepoint analysis for *P. falciparum*. Strikingly, they found that 87% of known drug targets with biological evidence are chokepoints according to their analysis [24]. In line, they identified three targets of clinically proven malaria drugs, dihydrofolate reductase, dihydropteroate synthase, and 1-deoxy-D-xylulose 5-phosphate reductoisomerase as chokepoints. Rahman and Schomburg performed a chokepoint and load score analysis for several other organisms [23, 26]. In Fatumo *et al*., we performed a chokepoint analysis together with our developed producibility concept to obtain a more reliable list of potential drug targets in *P. falciparum*. For example, we identified deoxyhypusine synthase involved in spermidine metabolism, which is a known drug target in *P. falciparum*, *Anopheles stephensi,* and *Trypanosoma evansi* [27]. This enzyme was detected by intersecting the predicted targets from a chokepoint and a producibility analysis [26].

Protein-protein interactions were inferred by a high-throughput method (yeast-2-hybrid) and assembled for a signaling network of *P. falciparum.* This has been performed for the first time by Suthram in 2005 identifying conserved proteins, pathways, and interactions [28]. The network was then analyzed by using a network alignment approach comparing the networks across organisms, by using various graph theoretical measures and an *in silico* knock-out strategy to identify potential drug targets [11, 12, 28, 29]. With this, conserved pathways and proteins between organisms were identified hinting for essentiality. The study showed that a few interactions were conserved among the analyzed organisms, demonstrating that the protein interaction network of *Plasmodium* is distinctively different from the others. Interestingly, a conserved protein complex was found in calmodulin-mediated endocytosis. Indeed, inhibition of calmodulin resulted in attenuated growth [30] and reduced chloroquine extrusion in malarial parasites diminishing drug resistance to chloroquine [31]. Additionally, endocytosis was found to be related to these mechanisms [32]. Thus, the proximity of calmodulin to the formation of endocytic vacuoles in *Plasmodium* provides an interesting link to discover strategies coping drug resistance mechanisms of *Plasmodium* [28].

Recently, Bhattacharyya and Chakrabarti analyzed a large-scale protein-protein interaction network of *Plasmodium* and identified potential drug targets using various graph theoretical measures such as centrality measures. They also used an *in silico* knock-out strategy to study the perturbation due to a loss of a protein in the network [12]. With this, approximately 270 proteins of *P. falciparum* were identified as potential drug targets including proteins, which play crucial roles in intra-pathogen network integrity, stage specificity but also interact with

various human proteins involved in multiple metabolic pathways within the host cell. Most of the housekeeping proteins were found to be potential targets [12].

Interactions between the human host and the parasite have been intensively studied [11–13, 33]. The comparison of several reconstructed network models has been performed to find the best suitable reconstruction for detecting drug targets *in silico*. This was performed on a metabolic network reconstruction based on automatically inferred enzymes and compared with a reconstructed model that based only on enzymes whose coding genes were known. These networks were analyzed with criteria for defining essential enzymes including chokepoints, betweenness centrality, connectivity, and the diameter of the networks. Comparing the modeling results with a comprehensive list of known drug targets for *P. falciparum* showed that the most suitable network model was constructed using only enzymes from the parasite alone, which coding genes were known [13].

Chen *et al*. developed a network-based approach to predict malaria-associated genes by a random walk algorithm [33]. They first constructed separate gene networks of the human genome and of the parasite genome and then connected them with known host-pathogen protein interactions. Known malaria target genes were used as the seeds (a set of nodes at which the search started) in a random walk algorithm to prioritize genes. The random walk algorithm then iteratively explored the global structure of the network starting at a set of nodes (seeds) to estimate the probability of a node being reached from the seeds. These probability scores can be viewed as the influential impact over the network imposed by the set of seed nodes. Finally, all the genes were ranked according to their probability scores. Manually examining the top 50 predicted human genes, interesting proteins such as TLR4 and P53 were found to be associated with malaria [33].

4. *In silico* modeling using flux balance analysis to identify drug targets

Flux balance analysis (FBA) is a computational approach to estimate the quantitative flux of metabolites through a mechanistic model of metabolism. Thereby, it is possible to predict the growth rate of an organism or the rate of production of an important metabolite [9, 34–36]. Biochemical stoichiometric equations are used to assemble a set of constraints to limit the feasible search space. The idea is that, at equilibrium, production and consumption of internal metabolites are balanced. This leads to a large set of equations in which the net production flux equals the net consumption flux for each internal metabolite. Additionally, allowable fluxes of any reaction are bounded at plausible maximum and minimum fluxes. Bounds may also be taken from the literature. These balances and bounds define the space of allowable flux distributions of a system, that is, the allowed combinations of fluxes for each reaction. To get a phenotype or modeling prediction from these constraints, an optimization criterion is put up. For example, in the case of predicting growth, the objective is to optimize biomass production which is the rate at which metabolic compounds are converted into the physiological portions of biomass constituents most importantly of nucleic acids, amino acids and lipids. Together with the constraints, this is mathematically formulated as a system of linear equations which is solved using linear programming based programs. Flux variability and knockout simulations

http://dx.doi.org/10.5772/65432

are analyzed to detect potential drug targets whose absence reduces the biomass production and hence viability of the parasite in the host cell. By simulating a reconstructed metabolic network of an organism of interest, first a "wildtype" model is investigated and the growth rate of the wildtype under specific bounds (or conditions) obtained. Performing a single gene (or reaction) knockout/deletion under the same condition by limiting its corresponding fluxes to zero (knockout simulation), the fluxes are calculated simulating an organism effected to a drug (targeting the deleted enzyme) and the growth rate is compared to the wildtype. A knocked out gene (or reaction) is predicted to be essential under the given condition if the mutant model yields a much lower growth rate compared to the wildtype. Flux balance analysis is a widely used and well-established technique to assess the essentiality of genes and hence potential drug targets [9, 34–36]. The beauty of this approach is that it does not depend on specific enzymatic parameters for each enzyme like their Michaelis Menten constants, etc., but are rather basing on simple stoichiometric equations. To some extent, the only experimental parameters are the boundary conditions. The drawback is that often several solutions can come out which are mathematically equally good, but physiologically very different leading to follow-up analyses of each of these solutions. Nevertheless, the approach was used for several genome-scale metabolic network constructions, followed by flux simulations of the inner metabolites of *Plasmodium* spp. to identify drug targets. It also enables to embed the metabolism of *Plasmodium* spp. into the metabolism of its environment, for example, human red blood cells [9]. Furthermore, experimental data on a systems view can be embedded using microarray or sequencing based gene expression data and with this, stage-specific metabolic models of the parasite were developed, particularly during the red-blood cell stage [9]. To better understand flux balance analysis and its potential, we will give a brief introduction into the mathematical secrets of it in the next section (which can be skipped without losing the track to understand the subsequent sections).

4.1. Flux balance analysis formulation

Let s_{ij} be the stoichiometric coefficient of metabolite i in reaction j, which specifies the number of metabolites produced or consumed by reaction j. $s_{ij} > 0$ indicates that reaction j produces metabolite i, while $s_{ij} < 0$ indicates that reaction j consumes metabolite i. $s_{ij} = 0$ means that metabolite i does not participate in reaction j. For example, considering a reaction $A + 2B \rightarrow C$, the stoichiometric coefficients of A, B, and C are −1, −2, and 1, respectively. The stoichiometric coefficients s_{ij} can be combined into the so-called *stoichiometric matrix* $S = (s_{ij})$. A rate of concentration change of a metabolite can be formulated by the set of system equations:

$$\frac{dx_i}{dt} = \sum_j s_{ij} v_j \tag{1}$$

where x_i is the concentration of metabolite i, s_{ij} is the stoichiometric coefficient, and v_j is the consumption/production rate of reaction j. Based on the assumption of mass conservation at steady state in the cell, internal metabolite concentrations are constant over time. Therefore, the concentration change of each internal metabolite i is zero, which means $\frac{dx_i}{dt} = 0$. With this assumption, equation (1) can be formulated as

$$\sum_j S_{ij} v_j = Sv = 0 \tag{2}$$

where S is the $m \times n$ stoichiometric matrix of m metabolites and n reactions in the network. The vector v represents all reaction rates (also called metabolic fluxes) in the metabolic network. The ranges of individual metabolic fluxes are constrained by $\alpha_j \leq v_j \leq \beta_j$ where α_j and β_j are the minimal and maximal fluxes of reaction j, respectively. These inequality constraints allow reversibility. If a reaction is reversible, the flux of the reaction v_j can either be negative or positive. A positive flux indicates the forward direction while a negative flux indicates a backward direction. If we want to block a reaction (knockout simulation), we can constrain the flux of this reaction to be equal to zero ($v_j = 0$). In addition, the benefit of these inequality constraints is to simulate metabolic capabilities under certain conditions such as a glucose minimal medium condition, which we can model by constraining the flux of the glucose uptake rate in a specific range of values and set the uptake rates of all other carbon source to zero. Finally, the set or subspace of vector v that satisfies all constrains and the ranges of individual metabolic fluxes is a set of feasible fluxes covering all feasible capabilities of the metabolic network under the given specific condition. Using an optimization criterion, such as to optimize the biomass of the cell yields then only one or a few out of these solutions. The biomass production rate can be defined by a reaction or several reactions that produce the metabolic building blocks of a cell (e.g., amino acids and nucleotides) or macromolecules that form the biomass in a physiological composition. The physiological biomass composition of a given organism comprises the relative amounts of the important molecules and can be found in the literature [9, 36]. The flux of the biomass production is associated with the specific growth rate of an organism. Finally, the obtained growth rate of the mutant (with a reaction knocked out) is compared to the growth rate of the wildtype to predict a gene or an enzyme to be essential. This section was taken from Ref. [37].

4.1.1. Applying FBA to predict drug targets

FBA has been widely used to predict essential genes of the human malaria parasite *P. falciparum* [9, 36, 38]. A metabolic network reconstruction of *P. falciparum* was developed with 1001 reactions and 616 metabolites [36]. The model allowed predicting the phenotype (growth) of experimental gene knockouts. Validating the predictions with drug inhibition assays yielded approximately 90% accuracy. Several modifications on the linear programming implementation were studied to make the static FBA model more realistic. For example, gene expression profiles of the malaria parasite were integrated into metabolic models [9, 36, 38]. In the study of Plata *et al*. [36], the maximum flux of the associated reactions was constrained by their expression level while Huthmacher *et al*. [9] used a method proposed by Shlomi *et al*. [39]. This method is a modification to flux balance analysis (FBA) by adding binary variables for each reaction. These binary variables act like an on/off switch according to the expression level. The mathematical objective is to maximize the number of non-zero fluxes for the reactions with switched-on-state. Dholakia *et al*. analyzed many available *omics* resources of stage-specific expression and used pathway tools from the BioCyc database to analyze flux distributions with respect to gene expression for identifying drug targets, and in particular in the erythrocytic stage-specific metabolism of the parasite. Based on the FBA approach, Plata *et al*. identified 40 enzymatic drug

http://dx.doi.org/10.5772/65432

targets. All of these enzymes had no or very low sequence similarity to human proteins which made them more attractive as this facilitated designing drugs targeting these enzymes and not human host factors. This set of genes consisted of six genes associated with isoprenoid metabolism, three genes involved in nucleotide metabolism, and the rest of genes related to CoA, shikimate, and folate biosynthesis. In addition, one of predicted essential genes, nicotinate nucleotide adenylyltransferase, was selected to be tested further in an experimental assay. This enzyme has been known for anti-microbial development [40] but not in *Plasmodium* spp. yet. Thus, in Plata et al., the experimental validation was done in *P. falciparum* by inhibiting this enzyme by a small-molecule inhibitor from [41] resulting in blocking host cell escape and reinvasion by arresting the parasites in the trophozoite growth stage [36]. Hence, FBA allowed the construction of stage-specific metabolic networks for different stages of the parasites and gave the opportunity to find drug targets for these stages. Additionally, also host-parasite interactions can be studied using FBA [9]. In the study by Huthmacher et al. [9], a host-parasite network was constructed and the metabolic fluxes for each blood life cycle stage were predicted employing gene expression data of the different stages. Knock-out simulations identified 307 indispensable metabolic reactions for the parasite. Of 57, 35 experimentally validated essential enzymes were recovered. Another set of 16 enzymes were predicted, if additionally assuming that nutrient uptake from the host cell is limited and all reactions catalyzed by the inhibited enzyme are blocked. An interesting modification to flux balance analysis was developed by a two-stage flux balance analysis to identify drug targets by comparing the differences of fluxes between a drug treated and untreated condition [42]. This approach was applied to find drug targets in *Plasmodium*, which is described in more detail in the next section.

4.2. Finding multiple drug targets to treat a drug resistant *Plasmodium* strain

Recently, Phaiphinit et al. reconstructed the metabolic network of *P. falciparum* in the human host red blood cell using flux balance analysis [35]. This model was used to analyze two specific metabolic models: a model for the parasite when having invaded the red blood cell without any treatment and, in turn, the treated situation, when a drug like chloroquine acts by inhibiting the hemozoin formation causing a high production rate of harmful heme. The process of identifying target combinations consisted of two main steps (**Figure 4**):

Step 1—Developing two multi-cellular metabolic models: The model was constructed for the situation of the parasite being inside the red blood cell of the human host. All metabolites of the parasite in exchange with the external environment were taken from the red blood cell. To find potential reactions which could harm *P. falciparum* by getting exposed to severe toxicity, the flux distribution of the multi-cellular metabolic model was calculated for two conditions. The first condition was the untreated situation where the parasite was able to get rid of toxins from hemoglobin degradation after consuming hemoglobin from the red blood cell. The second condition mimicked the treated situation in which the toxins could not be degraded. The difference in flux distributions between the two conditions was assumed to be the effect from the drug which disturbed the parasite.

Step 2—Finding the optimal drug target: The reactions in the parasite which were disturbed from the drug in the treated situation may suit as drug targets for a combined treatment, or if

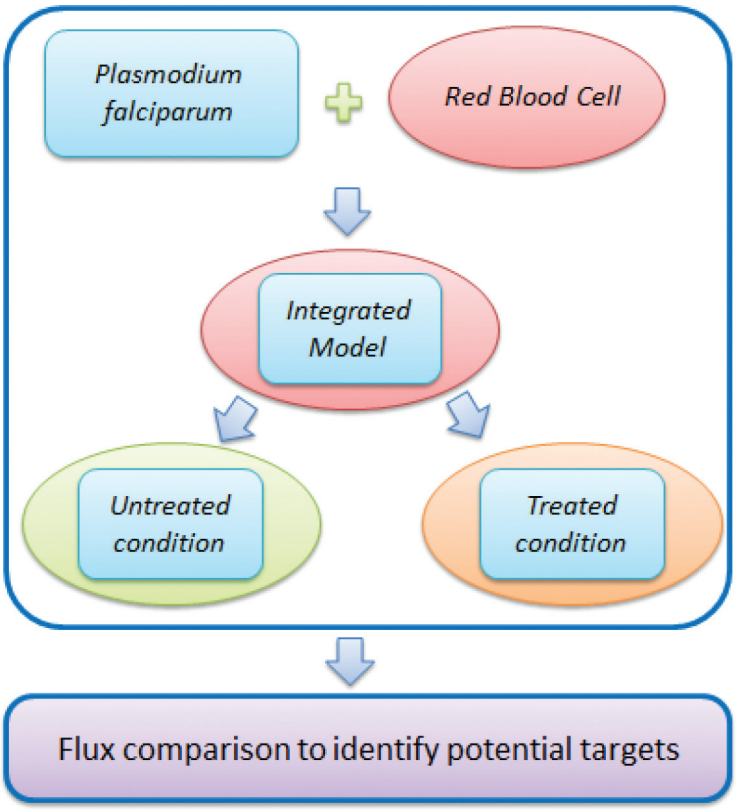

Figure 4. The workflow to identify drug targets by the comparison of treated and untreated conditions. First, *in silico* models of *Plasmodium falciparum* and the human red blood cell are combined as an integrated model. This integrated model is characterized by two specific conditions (treated and untreated). The flux rates of all reactions in both situations are compared to identify a set of potential drug targets.

the parasite gets resistant to the first drug (chloroquine). In particular, reactions for which no flux was predicted in the treated scenario were promising targets because they may have a similar treatment effect when targeted compared to the original drug and may suit as drug targets against strains which are resistant to the first drug.

FBA was used to get the flux distributions for the untreated and the treated conditions. For the untreated condition, the objective was to maximize the production rate of biomass according to Ref. [36], including the Na^+/K^+ ratio based potential at the ATPase, which plays an important role for the homeostasis of red blood cells [43, 44]. In the treated condition, the drug usually inhibits the detoxification process of the parasite harming the parasite due to the toxicity of free heme. Thus, during the treated condition, the (toxic) flux of heme production should be an additional objective to ensure that the toxic flux is not zero when identifying reactions or enzymes to be blocked during the treatment. The flux distributions of both models were then compared to obtain a list of candidate targets by the criteria that the reactions with zero fluxes

http://dx.doi.org/10.5772/65432

in the treated condition but non-zero fluxes in the untreated condition could be potential targets for inhibiting heme detoxification.

With this method, 23 enzymes were identified as candidate targets, which mostly were in pyruvate metabolism and the citrate cycle. The optimal set of multiple targets for blocking the detoxification was a set of a heme ligase, adenosine transporter, myo-inositol 1-phosphate synthase, ferrodoxim reductase-like protein, and the guanine transporter. Purine transporters have been known as the major route of purine into the parasitized red blood cell. In the development of anti-malarial drugs, inhibitors targeting purine transport are of pharmaceutical interest and are investigated. Likewise, adenosine transport and its inhibitor have been studied in infected and uninfected human erythrocytes recently [45]. In summary, this shows an efficient way to identify useful target combinations in the development of novel antimalarial drugs [35].

5. Experimental validation, a case study

Typically, after the computational network analysis, a list of potential drug targets is assembled and needs to be validated experimentally. Exemplarily, in one study of a topological network analysis, 22 potential targets were proposed [26]. Using a refined network comprising also the host enzymes led to a refined set of the five potential drug targets (glutamyl–tRNA (gln) amidotransferase, hydroxyethylthiazole kinase, deoxyribose–phosphate aldolase, pseudouridylate synthase, and deoxyhypusine synthase) [46]. The next step was to find effective inhibitors to block these enzymes. Many reported inhibitors can be collected from databases like the Brenda Enzyme database [47], Drugbank [48], and from companies like Sigma (http://www.sigma.com), or by scanning the literature. In this example, a study was found, in which Jahn and coworkers used 6-diazo-5-oxonorleucine (DON) to be an effective inhibitor of glutamyl-tRNA(Gln) amidotransferase in *Chlamydomonas reinhardtii* [49]. Accordingly, an experimental viability assay (IC_{50} analysis) was performed and showed that DON suits as a valid agent against *P. falciparum* (laboratory strain Dd2) in *Plasmodium* infected blood cultures. Strikingly, this was confirmed by an *in vivo* study using *Plasmodium berghei* infected Swiss albino mice. All treated mice survived whereas all untreated died [45].

6. Conclusions

Even though the number of deaths caused by malaria has diminished considerably, it is still a challenge to treat the effected patients and clear off the pathogen after infection. In particular, there are increasingly more strains getting resistant against common treatments, and hence there is a striking demand to find new targets for therapy.

The computational approaches introduced here show some convincing results. However, it needs to be shown that these predictions are experimentally confirmed and finally make their way from the bench to the bedside.

Various techniques of network-based analyses to identify potential drug targets of *Plasmodium* have been described in more detail including the construction of cellular networks, the

analysis of topological features, as well as *in silico* models based on flux balance analysis. To construct a network, one needs to consider the network types which are suitable to find the targets of interest. Moreover, the consideration of the interactions between host and pathogen makes the network more realistic, but, however, also more complex to obtain drug targets. Analyzing topological features seems to be a comfortable way to retrieve interesting targets; however, the *in silico* models using flux balance may reflect much more detailed relations of the biochemical reactions in a cell. All of the methods described in this chapter provided promising results, some with experimental evidence. It is to be noted that they have been widely used for a large variety of other organisms as well.

Even though all these presented concepts have the very same aim to find a target, their results are quite heterogeneous lists of different predicted drug targets, some of them validated by experimental assays. As a future aspect, a data and method integration needs to be performed leading to a *consistent* set of targets independent from the data it bases on, and, at its best, being consistent with a larger set of experimental data sets and validations.

Author details

Kitiporn Plaimas[1,2] and Rainer König[3,4]*

*Address all correspondence to: rainer.koenig@uni-jena.de

1 AVIC Research Center, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

2 Omics Sciences and Bioinformatics Center, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

3 Integrated Research and Treatment Center, Center for Sepsis Control and Care (CSCC), Jena University Hospital, Jena, Germany

4 Network Modeling, Leibniz Institute for Natural Product Research and Infection Biology—Hans Knöll Institute, Jena, Germany

References

[1] Ahn YY, et al., Metabolic network analysis-based identification of antimicrobial drug targets in category A bioterrorism agents. PLoS One, 2014. **9**(1): p. e8519.

[2] Peng Q, Schork NJ, Utility of network integrity methods in therapeutic target identification. Front Genet, 2014. **5**: p. 12.

[3] Acencio ML, Lemke N, Towards the prediction of essential genes by integration of network topology, cellular localization and biological process information. BMC Bioinformatics, 2009. **10**: p. 290.

[4] Luscombe NM, et al., Genomic analysis of regulatory network dynamics reveals large topological changes. Nature, 2004. **431**(7006): pp. 308–312.

[5] König R, Eils R, Gene expression analysis on biochemical networks using the Potts spin model. Bioinformatics, 2004. **20**(10): pp. 1500–1505.

[6] Barabasi AL, Oltvai ZN, Network biology: understanding the cell's functional organization. Nat Rev Genet, 2004. **5**(2): pp. 101–113.

[7] Karp PD, et al., Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Res, 2005. **33**(19): pp. 6083–6089.

[8] Ginsburg H, Progress in in silico functional genomics: the malaria metabolic pathways database. Trends Parasitol, 2006. **22**(6): pp. 238–240.

[9] Huthmacher C, et al., Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis. BMC Syst Biol, 2010. **4**: p. 120.

[10] Gursoy A, Keskin O, Nussinov R, Topological properties of protein interaction networks from a structural perspective. Biochem Soc Trans, 2008. **36**(Pt 6): pp. 1398–1403.

[11] Swann J, et al., Systems analysis of host-parasite interactions. Wiley Interdiscip Rev Syst Biol Med, 2015. **7**(6): pp. 381–400.

[12] Bhattacharyya M, Chakrabarti S, Identification of important interacting proteins (IIPs) in *Plasmodium falciparum* using large-scale interaction network analysis and in-silico knockout studies. Malar J, 2015. **14**: p. 70.

[13] Fatumo S, et al., Comparing metabolic network models based on genomic and automatically inferred enzyme information from Plasmodium and its human host to define drug targets in silico. Infect Genet Evol, 2011. **11**(4): pp. 708–715.

[14] Barabasi AL, Albert R, Emergence of scaling in random networks. Science, 1999. **286** (5439): pp. 509–512.

[15] Almaas E, Biological impacts and context of network theory. J Exp Biol, 2007. **210**(9): pp. 1548–1558.

[16] Milgram S, Small-world problem. Psychol Today, 1967. **1**(1): pp. 61–67.

[17] Zhu X, Gerstein M, and Snyder M, Getting connected: analysis and principles of biological networks. Genes Dev, 2007. **21**(9): pp. 1010–1024.

[18] Albert R, Jeong H, and Barabasi AL, Error and attack tolerance of complex networks. Nature, 2000. **406**(6794): pp. 378–382.

[19] Wagner A, Fell DA, The small world inside large metabolic networks. Proc Biol Sci, 2001. **268**(1478): pp. 1803–1810.

[20] Zur H, Ruppin E, and Shlomi T, iMAT: an integrative metabolic analysis tool. Bioinformatics, 2010. **26**(24): pp. 3140–3142.

[21] Estrada E, Protein bipartivity and essentiality in the yeast protein-protein interaction network. J Proteome Res, 2006. **5**(9): pp. 2177–2184.

[22] Samal A, et al., Low degree metabolites explain essential reactions and enhance modularity in biological networks. BMC Bioinformatics, 2006. **7**: p. 118.

[23] Rahman SA, Schomburg D, Observing local and global properties of metabolic pathways: 'load points' and 'choke points' in the metabolic networks. Bioinformatics, 2006. **22**(14): pp. 1767–1774.

[24] Yeh I, et al., Computational analysis of *Plasmodium falciparum* metabolism: organizing genomic information to facilitate drug discovery. Genome Res, 2004. **14**(5): pp. 917–924.

[25] Bonday ZQ, et al., Import of host delta-aminolevulinate dehydratase into the malarial parasite: identification of a new drug target. Nat Med, 2000. **6**(8): pp. 898–903.

[26] Fatumo S, et al., Estimating novel potential drug targets of *Plasmodium falciparum* by analysing the metabolic network of knock-out strains in silico. Infect Genet Evol, 2009. **9** (3): pp. 351–358.

[27] Moritz E, et al., The efficacy of inhibitors involved in spermidine metabolism in *Plasmodium falciparum*, *Anopheles stephensi* and *Trypanosoma evansi*. Parasitol Res, 2004. **94**(1): pp. 37–48.

[28] Suthram S, Sittler T, Ideker T, The Plasmodium protein network diverges from those of other eukaryotes. Nature, 2005. **438**(7064): pp. 108–112.

[29] Kelley BP, et al., PathBLAST: a tool for alignment of protein interaction networks. Nucleic Acids Res, 2004. **32**(Web Server issue): pp. W83–W88.

[30] Scheibel LW, et al., Calcium and calmodulin antagonists inhibit human malaria parasites (*Plasmodium falciparum*): implications for drug design. Proc Natl Acad Sci U S A, 1987. **84** (20): pp. 7310–7314.

[31] Sanchez CP, et al., Evidence for a substrate specific and inhibitable drug efflux system in chloroquine resistant *Plasmodium falciparum* strains. Biochemistry, 2004. **43**(51): pp. 16365–16373.

[32] Hoppe HC, et al., Antimalarial quinolines and artemisinin inhibit endocytosis in Plasmodium falciparum. Antimicrob Agents Chemother, 2004. **48**(7): pp. 2370–2378.

[33] Chen Y, Xu R, Network-based gene prediction for *Plasmodium falciparum* malaria towards genetics-based drug discovery. BMC Genomics, 2015. **16**(Suppl 7): p. S9.

[34] Orth JD, Thiele I, Palsson BO, What is flux balance analysis? Nat Biotechnol, 2010. **28**(3): pp. 245–248.

[35] Phaiphinit S, et al., In silico multiple-targets identification for heme detoxification in the human malaria parasite *Plasmodium falciparum*. Infect Genet Evol, 2016. **37**: pp. 237–244.

[36] Plata G, et al., Reconstruction and flux-balance analysis of the *Plasmodium falciparum* metabolic network. Mol Syst Biol, 2010. **6**: p 408.

[37] Plaimas K, Computational Analysis of the Metabolic Network of Microorganisms to Detect Potential Drug Targets, Doctoral thesis, Heidelberg University, 2011.

[38] Dholakia N, Dhandhukia P, Roy N, Screening of potential targets in *Plasmodium falciparum* using stage-specific metabolic network analysis. Mol Divers, 2015. **19**(4): pp. 991–1002.

[39] Shlomi T, et al., Network-based prediction of human tissue-specific metabolism. Nat Biotechnol, 2008. **26**(9): pp. 1003–1110.

[40] Magni G, et al., NAD(P) biosynthesis enzymes as potential targets for selective drug design. Curr Med Chem, 2009. **16**(11): pp. 1372–1390.

[41] Sorci L, et al., Targeting NAD biosynthesis in bacterial pathogens: structure-based development of inhibitors of nicotinate mononucleotide adenylyltransferase NadD. Chem Biol, 2009. **16**(8): pp. 849–861.

[42] Li Z, Wang RS, Zhang XS, Two-stage flux balance analysis of metabolic networks for drug target identification. BMC Syst Biol, 2011. **5**(Suppl 1): p. S11.

[43] Mauritz JM, et al., The homeostasis of *Plasmodium falciparum*-infected red blood cells. PLoS Comput Biol, 2009. **5**(4): p. e1000339.

[44] Wiback SJ, Palsson BO, Extreme pathway analysis of human red blood cell metabolism. Biophys J, 2002. **83**(2): pp. 808–818.

[45] Quashie NB, Ranford-Cartwright LC, de Koning HP, Uptake of purines in *Plasmodium falciparum*-infected human erythrocytes is mostly mediated by the human equilibrative nucleoside transporter and the human facilitative nucleobase transporter. Malar J, 2010. **9**: p. 36.

[46] Plaimas K, et al., Computational and experimental analysis identified 6-diazo-5-oxonorleucine as a potential agent for treating infection by *Plasmodium falciparum*. Infect Genet Evol, 2013. **20**: pp. 389–395.

[47] Schomburg I, Chang A, Placzek S, Sohngen C, Rother M, Lang M, Munaretto C, Ulas S, Stelzer M, Grote A, et al., BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res, 2013. **41**(Database issue): pp. D764–D772.

[48] Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B, Hassanali M, DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res. 2008. **36**(Database issue): pp. D901–D906.

[49] Jahn D, Kim YC, Ishino Y, Chen MW, Soll D. Purification and functional characterization of the Glu–tRNA(Gln) amidotransferase from Chlamydomonas reinhardtii. J Biol Chem, 1990. **265**: pp. 8059–8064.