

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Current Advances in Computational Strategies for Drug Discovery in Leishmaniasis

Andrés F. Flórez<sup>1,2</sup>, Stanley Watowich<sup>3</sup> and Carlos Muskus<sup>2</sup>

<sup>1</sup>German Cancer Research Center (DKFZ)/ Division Theoretical Systems Biology, Heidelberg

<sup>2</sup>Universidad de Antioquia/Programa de Estudio y Control de Enfermedades Tropicales – PECET, Medellín

<sup>3</sup>University of Texas Medical Branch/Department of Biochemistry and Molecular Biology and the Sealy Center for Structural Biology and Molecular Biophysics, Galveston

<sup>1</sup>Germany

<sup>2</sup>Colombia

<sup>3</sup>USA

## 1. Introduction

Leishmaniasis is a complex disease caused by several species of the *Leishmania* genus ranging in severity from cutaneous and mucocutaneous lesions to the chronic visceral form that if untreated adequately can cause death. It has a worldwide distribution in 98 countries and 85 out of 98 are developing or poor countries. One of the main problems in leishmaniasis is the limited number of drug options along with the adverse effects they can cause including death (Ahasan., et al. 1996; Sundar & Chakravarty 2010; Oliveira., et al. 2011). In addition, there are reports of treatment failures due to increased parasite resistance to the first drug of choice, the antimonials (Faraut-Gambarelli., et al. 1997; Goyeneche-Patino., et al. 2008). Second-choice drugs, such as amphotericin B, pentamidine, paromomycin, and more recently, miltefosine, have also toxic effects that require hospital management (Maltezou 2008; Oliveira., et al. 2011). Miltefosine, the only oral administered drug for leishmaniasis, has not been tested in many *Leishmania* species. Recently, a central nervous system toxicity was reported for liposomal amphotericin B therapy used to treat cutaneous leishmaniasis (Glasser & Murray 2011).

In the search for new drug targets in *Leishmania*, a group of proteins have been proposed based mainly on their known function, the expression level, and localization, or because they are involved in important metabolic processes in the parasite. Topoisomerases (Das., et al. 2008), kinases (de Azevedo & Soares 2009), proteins localized or targeted to lysosomes (Carrero-Lerida., et al. 2009) are some potential *Leishmania* drug targets. However, none of these protein targets have been used to successfully develop new drugs that can substitute the existing therapies.

Currently, the massive genome sequencing of many medically important microorganisms together with protein structure and drug databases and the development of new computational tools, will allow molecular targets and new drugs to be searched in a more rigorous manner. Three *Leishmania* genomes, *L. major*, *L. infantum* and *L. braziliensis*

(Peacock., et al. 2007) have been sequenced and annotated and a fourth species, *L. mexicana* and some *L. major* strains are in the process of being sequenced (GeneDB, <http://www.genedb.org>; University of Washington Genome Sequencing Center, <http://genome.wustl.edu/gsc/gschmpg.html>). The availability of these genomes and the annotated proteins can be used in a rational manner to predict novel drug targets and provide a basis to develop new drugs.

The computational prediction of drugs, in addition to the evaluation of drugs already synthesized and used in other diseases, must be coupled with automated in vitro assessment methodologies of these compounds. In this sense and in the case of *Leishmania*, the use of GFP (Varela., et al. 2009) or luciferase transgenic parasites (Lang., et al. 2005) coupled with techniques such as flow cytometry or fluorometry can be used to rapidly evaluate potential anti-leishmanial drugs. The WHO program for training in tropical diseases research has created a network based on reporter gene technology to foster the process of drug search not only against leishmaniasis but also against other diseases with limited therapeutic options.

## 2. Selection of drug targets

An initial step in the drug discovery process involves the search and selection of the drug target. This target is frequently a protein that is essential for the organism survival or critical for regulating a particular signaling pathway. In the specific case of parasites, the protein target when inhibited should impair or delay parasite viability. The classical approach of finding a new essential protein that can act as a potential target is the experimental characterization by using gene knockout or knock-down strategies in the target organism. Besides essentiality, some targets are selected for being specific for the pathogen; for example, the ergosterol pathway is present in fungi and *Leishmania spp*, but humans only contain the required enzymes for the synthesis of cholesterol. This is the reason why this pathway has been exploited for searching drugs against mycotic pathogens and also *Leishmania*. However, the experimental approach employing interference RNA (RNAi) is not feasible given *Leishmania* species do not carry the machinery for RNAi (Peacock., et al. 2007), with the exception of *Leishmania braziliensis* where some RNAi-associated genes have been found. In addition, depending on the parasite stage the essentiality of a particular protein could change dramatically. With all these constraints, a rational alternative for choosing effective targets is a more systematic study of the biology of the parasite, with the aim of uncovering important mechanisms that are not evident by studying descriptively isolated proteins. A starting point for this “systems view” of the parasite biology in the case of *Leishmania*, was the sequencing of its genome in 2005 (Ivens., et al. 2005). Since then, more high-throughput data have been generated, not at the same rate as other organisms but with important applications for drug discovery in tropical diseases. This leads to an important issue of data analysis, where computational tools can have a role in reducing the ocean of possibilities of finding a drug for this disease, making more efficient and less costly the experimental setup. In the following sections, we will describe the current computational methods that can be applied to find new drug targets, with special application to the *Leishmania* parasite.

### 2.1 Selection of targets by homology searching

The simplest approach for finding a drug target is the homology search of essential proteins. There are several organisms with available essential data at genome-wide level (Forsyth., et al. 2002; Kamath., et al. 2003; Hu., et al. 2007). In model organisms such as yeast, the

phenotypic effects of deletion of particular genes have been shown (Giaever., et al. 2002) and more recently the study of genetic interactions on a large scale (Costanzo., et al. 2010). This has been used to elucidate redundancy and possibly some synergistic effects among genes. Therefore, it is possible to find orthologs in the organism of interest that could be essential by comparing its sequences against the list of essential genes in model organisms. The Database of Essential Genes (<http://tubic.tju.edu.cn/deg/>) (Zhang & Lin 2009) provides information of essential genes in prokaryotes and eukaryotes, and it is also possible to do a BLAST search with the protein of interest. This resource is useful for an exploratory search of essentiality of a particular protein. Another important resource, for drug target data, is the DrugBank database (<http://www.drugbank.ca/>) (Knox., et al. 2011), which can be used to extract drug-target interactions along with additional pharmacological data. The same strategy can be employed in this case; with the advantage that the homology search will also return possible drug candidates that can be tested on the protein found to have homology to the target in DrugBank.

This methodology has been applied in *Pseudomonas aeruginosa* (Sakharkar., et al. 2004) with the aim of detecting new drug targets, given this bacterium is an important problem in nosocomial settings due to the rapid generation of resistance. In *Leishmania*, drug targets can be also identified by this approach. Tools like BLAST or PSI-BLAST can be employed, with PSI-BLAST being more sensitive for detecting distant relationships among proteins (Altschul., et al. 1997). However, some false positives still can occur due to alignments that are optimal according to the algorithm but not biologically meaningful. The E value helps to detect those alignments that are significant. As an example, running a PSI-BLAST search with the *Leishmania major* proteome against the DrugBank database, one can find among the potential *Leishmania* orthologs to known targets, the protein *LmjF36.2430*, which is similar to the sterol 14-  $\alpha$  demethylase in fungi. Drugs such as miconazole are known inhibitors of this enzyme. Interestingly, the protein *LmjF19.0450* belongs to the group of protein kinases conserved in other *Leishmania* species; it is constitutively expressed and has significant similarity to other kinase targets in cancer. These are simple cases of how a homology search can generate a list of potential drug targets using existing genomic data. The main advantage of this methodology is that it offers a quick overview of potential targets and second use of drugs. In addition, the STITCH 2 database (<http://stitch.embl.de/>) (Kuhn., et al. 2010) compiles known and predicted drug-target relationships jointly with biological information about targets in a network-based view.

Despite its simplicity, the homology search strategy has some caveats. Proteins inside the cell perform specific functions depending on their interactions, and these interactions can vary between species. Even if sequences are highly related, pathway conservation is not necessarily present. In addition, temporal regulation is important, as not all the interactions are active at the same time, which can further complicate the analysis. These problems highlight the importance of detecting targets by incorporating more detailed information about the molecular interactions.

## 2.2 Selection of targets by topological analysis of protein networks

In order to better understand complex pathogens such as *Leishmania* and to improve the efficiency of the drug discovery process, it is crucial to gain deeper knowledge about how protein interactions are established and how these interactions are regulated. This is a central issue for a more accurate definition of essentiality and biological robustness. These interactions can be described as a *network*, a representation commonly used to describe

complex systems. The protein interaction network (interactome) describes all possible molecular interactions among proteins. The interactome is composed of *nodes* that represent the molecular components, in this case proteins, and *edges*, that are the interactions between components (Fig. 1). Depending on the biological function of the node, other types of networks can also be constructed; for example, gene networks involving transcription factors as nodes that regulate other genes by binding (edges) and metabolic networks where the nodes are the enzymes connected by the production of some metabolites. The study of networks comes from a mathematical discipline called *graph theory*, and the analysis of the interaction patterns in the network is defined as *network topology*. (Barabasi & Oltvai 2004)

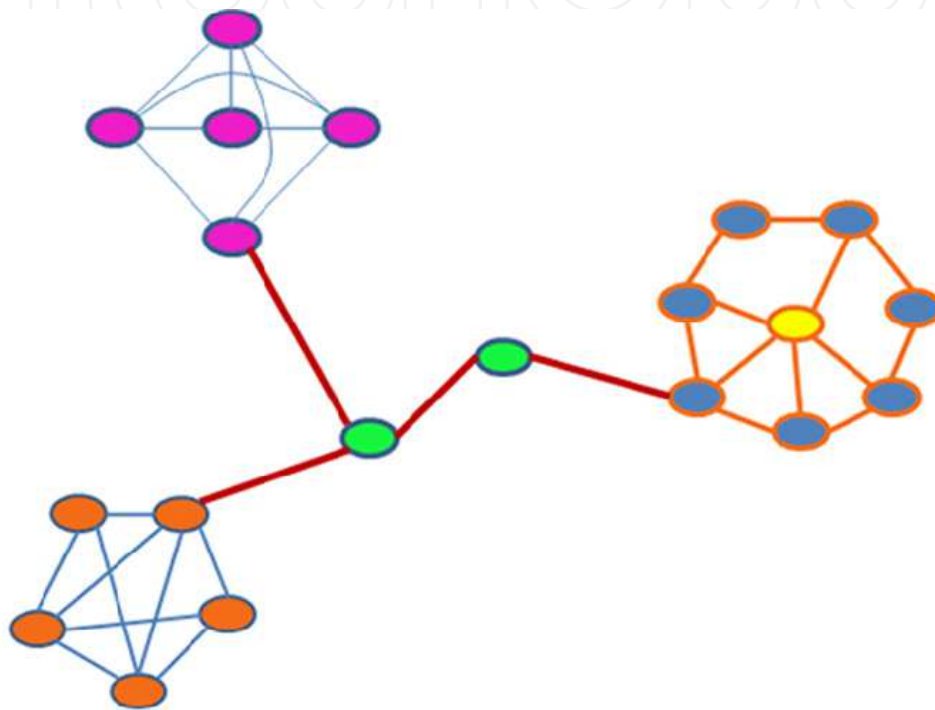


Fig. 1. Schematic representation of a protein network. Yellow circle corresponds to a hub protein, green circles correspond to bottleneck proteins connecting several sub-networks. Lines connecting circles represent the edges of the network.

To detect protein interactions in biological systems, large-scale methods have been developed that can map all possible pairwise interactions. Yeast two-hybrid is a popular technique of this kind, which was used to construct the first interactome (Uetz, et al. 2000). The technique involves the fusion of a protein with a transcription factor DNA-binding domain subunit. This protein is called the *bait*. The second protein is fused onto an activator domain subunit and it is called the *prey*. If the interaction between the bait and the prey is present, the two transcription factor subunits will come closer and the expression of the reporter gene is activated (Osman 2004). The most important limitation of this method is the presence of high number of false positives. However recent evidence has shown that a combination of experimental methods will reduce the number of false interactions (Dreze, et al. 2010).

The initial studies of the yeast interactome revealed that the network structure was not organized randomly, and in fact the organization pattern was similar to other experimentally-observed networks. This particular network structure was called *scale-free*



and it was elucidated by analyzing the number of interactions (or degree distribution) of proteins in the yeast interactome, showing that some nodes were more highly connected than others, and those nodes were in relatively low frequency in the network. This scale-free structure followed a *power law* distribution for the node degree and it described the probability of a node having a certain degree. An interesting consequence of having a scale-free structure is that the network was robust against random deletion of nodes, but susceptible to the deletion of highly connected nodes or *hubs* (Jeong., et al. 2001). The hubs can be detected by measuring the connectivity or degree of the network. In addition, the scale-free network was also susceptible to deletion of other types of nodes that were not highly connected but control the flux of the network; these nodes were called *bottlenecks* (Yu., et al. 2007). A classical example of bottleneck nodes is the scaffold proteins (Good., et al. 2011); these proteins facilitate the communication between signalling pathways very efficiently, although sometimes they are not highly connected. Deleting a bottleneck node will disrupt cellular homeostasis by destroying communication between processes in the cell. This network biology approach becomes an important step in a systems level understanding of the biology of parasites like *Leishmania*, and it becomes very useful for detecting essential nodes that may constitute potential new drug targets.

### 2.2.1 Construction of the *Leishmania* protein interaction network

The analysis of the *Leishmania* protein network could lead to the discovery of new and effective drug targets. However, current protein interaction data in *Leishmania* have only focused on a few specific proteins, and at this time, no yeast two-hybrid data is available for this organism. Despite this limitation, the use of a computationally-predicted protein network from orthology-based methods is a good first step for the exploration of drug targets that may be more informative than a traditional homology search. The results described in the next section will focus on the current status of the predicted *Leishmania major* interactome and will give some directions for future experimental studies for network and target validation.

Even when protein domain sequences are conserved, multiple combinations of these domains enable an organism to rewire the interactome in different ways. This can overcome the problem of the context of the targets that influence essentiality and enable new hubs or protein targets to be detected. A common disadvantage is the bias towards detection of conserved interactions, which could be a caveat in the case of organism-specific interactions that may also be important for survival. These specific interactions will be only detected when more data becomes available, which will also allow existent predictions to be validated.

In our recent study (Florez., et al. 2010), the protein interaction network in *Leishmania major* was predicted using only the parasite protein sequences and several protein interaction databases, in particular iPfam (Finn., et al. 2005), PSIMAP (Park., et al. 2005) and PEIMAP. These databases included protein-protein interactions defined by analysis of structures of protein complexes and experimental data extracted from literature, including high-throughput experiments. From the structures, the analysis of interacting structural domains was mapped to the sequence, using the domain definition by Pfam (Finn., et al. 2006) and SCOP (Hubbard., et al. 1997). These two databases contained information of domains with a systematic classification for protein families. In this particular case the physical distance between adjacent domains within a complex was used as the criteria for the definition of interaction and it was stored in iPfam and PSIMAP databases. This strategy has been used

in other organisms such as fungi and bacteria (He., et al. 2008; Kim., et al. 2008). The domain interaction analysis generated more diversity in the detection of possible interactions because modular exchange of protein domains allowed rewiring the network even if the isolated sequence of the domain was conserved. However, despite the high accuracy of this method, the prediction of protein interactions was limited as there was not an abundance of crystallized protein complexes. The PEIMAP database was also used, and it included sequences of protein interaction pairs detected by several methods, including co-immunoprecipitation (co-IP) and yeast two-hybrid.

To construct the *Leishmania major* network, protein sequences were extracted from the GeneDB database. This database included genomic and proteomic information of pathogens, including protozoan parasites. The protein sequences were aligned to the interacting domain pairs using PSI-BLAST against the SCOP 1.71 database with an E-value cutoff of 0.0001, as described previously (Kim., et al. 2008). The PSI-BLAST tool was used for the alignments because it had the advantage of detecting small conserved sequences, such as small domains that would be otherwise missed by using the standard BLASTP. The same strategy was applied for the alignments concerning the iPfam database. In this case, the domain assignment for the *Leishmania* proteins was carried out using the Pfam database (release 18.0) with the hmmpfam tool employed for the alignments. The final set of predicted interactions was carried out by homology search over the PEIMAP database using BLASTP, with a minimal cutoff of 40% sequence identity and 70% length coverage. The PEIMAP database included protein-protein interaction (PPI) information from six source databases: DIP (Xenarios., et al. 2000), BIND (Bader., et al. 2001), IntAct (Hermjakob., et al. 2004), MINT (Zanzoni., et al. 2002), HPRD(Peri., et al. 2004), and BioGrid (Stark., et al. 2006).

### 2.2.2 Filtering interactions by using a combined confidence score

As discussed earlier, the reliability of this analysis and its bias to certain types of protein interactions was dependent on the experimental method employed. Therefore, it was necessary to combine results from different databases to increase the coverage and the confidence of the predicted interactions. In the *Leishmania major* interactome, we used a simple scoring system to identify high confidence interactions. A previous study classified the experimental methods according to their reliability (Chua., et al. 2006), and we used this data in addition to the significance of the sequence alignments to calculate the confidence of the interactions. This scoring system was called the 'combined score' method, and it was applied for the confidence calculations in the STRING database (von Mering., et al. 2005). This database is useful for searching predicted protein interactions detected by other methods, although the definitions are beyond the scope of this chapter.

The score was calculated according to the formula (1):

$$score = 1 - \prod_{i \in E} (1 - R_i)^n \quad (1)$$

where *score* was the confidence value ranging from 0-1 with 1 equals to 100% accuracy, *E* was the set of methods under analysis (PEIMAP, PSIMAP, iPfam); *R<sub>i</sub>* was the reliability of method *i*, and *n* was the number of interactions predicted by method *i*. The results of these calculations represented pairs of interactions with their respective confidence. With this information, it was possible to select those interactions that fulfilled a particular confidence threshold. In this case, a confidence score of 0.7 was chosen to select the core *Leishmania major* network. The threshold

selection can vary depending on how strongly supported the interactions were required. For us, a 0.70 confidence value gave a smooth fit to the power law distribution and this was an important condition for reliable detection of hubs and bottlenecks.

### 2.2.3 Topological analysis of the network

Topological metrics such as clustering coefficient and mean shortest path help to describe global characteristics of the network. They measure the density of the connections within the network. Highly dense connected networks are characterized by modular components which also maintain the robustness of the network against failures. Biological networks tend to have a modular structure (Jeong., et al. 2001) and one additional way to test for reliability of the predicted network is by comparing the values of the clustering coefficient and mean shortest path to randomly generated networks with the same number of nodes and edges. These metrics should be statistically different between predicted and random networks. In the case of *Leishmania* network, 1,000 random networks were generated and the metrics calculated and compared to the original network.

The power law fitting for the definition of scale-free structure can be calculated using the plug-in Network Analyzer v.2.6.1 (Assenov., et al. 2008) available in the platform Cytoscape (Shannon., et al. 2003). This platform includes a very advanced environment for network visualization and analysis. Network topology metrics, such as betweenness centrality, and connectivity were calculated using the Hubba server (<http://hub.iis.sinica.edu.tw/Hubbawebcite>). (Lin., et al. 2008) A plug-in version of this tool in Cytoscape was recently made available. For the calculation of the metrics, the confidence scores of the interactions were used so the detection could be focused on the nodes most likely to be essential in the group of highly supported interactions. From this analysis, a potential list of targets was selected. However, it was possible that some proteins detected could also be conserved in terms of sequence and function among several organisms including humans. This becomes a problem if drugs targeting some of these proteins interfere with important biological process in humans, generating unwanted toxic effects. To avoid this, an additional filter was used for the list of predicted targets and it consisted of aligning the *Leishmania* proteins to the human proteins and excluding proteins that were conserved between these two species.

### 2.2.4 Prediction of protein function from network clusters

An important feature of network analysis was the prediction of protein function. The normal procedure for inferring function involved a homology search of the unknown protein versus a curated protein database such as UniProt (<http://www.uniprot.org/>). In some occasions, the detection of protein function was not feasible as significant similarity could not be found. When this approach failed, protein interaction network analysis helped to uncover potential functions. The prediction of protein function based on network analysis involved the assumption suggested by experimental data that interacting proteins tended to have related functions. This implied that it was possible to predict the function of neighboring nodes by clustering network modules and knowing the function of some of the nodes inside of the module. This analysis was carried out over the *Leishmania* network using the Markov Clustering (MCL) algorithm (Enright., et al. 2002) which has been demonstrated to be a robust and fast algorithm for detecting clusters or modules in protein networks (Brohee & van Helden 2006). The algorithm was implemented in the NeAT tool (Brohee., et



al. 2008). For proteins of unknown function in the GeneDB database, we predicted their possible biological roles by evaluating the results of Gene Ontology terms for biological processes using the BinGO plug-in available in Cytoscape.

2.2.5 Selection of candidate drug targets from the network analysis

We constructed a protein-protein interaction (PPI) map, combining the results generated by PEIMAP, iPfam and PSIMAP (Fig. 2). The number of interactions detected for each database is described in (Table 1). By merging the data from the different approaches, bias to a specific class of interactions was avoided. The predicted network also contained isolated sub-networks which were difficult to analyze. These sub-networks appeared as a consequence of the inability to assign domains or from the lack of homology of those proteins to the known pairs of protein interactions. These sub-networks could be investigated by further experimental validation of the network. The total number of high confidence predicted interactions were 33,861 for 1,366 nodes

Number of Proteins	PEIMAP		PSIMAP		IPFAM	
8,335	Nodes	Edges	Nodes	Edges	Nodes	Edges
	718	14,839	3,184	158,984	2,336	50,398

Table 1. Number of nodes and predicted interactions for each database.

By using the topological metrics of connectivity and betweenness centrality we identified 384 potential targets. From these targets, those that had homology to human proteins were eliminated. This substantially reduced the number of potential targets, although higher specificity of drug effects was expected. As explained earlier, toxicity becomes a very important issue when designing or searching for a drug, since many clinical trials failed because of undesired and severe side effects. After this filter, the final number of targets was reduced to 142. Further filters can be applied to this list to select those targets that were most attractive for drug design (Table 2).

From the group of targets, 91 kinases were predicted as essential proteins in the network with no homology to the human kinome. Kinases are very important regulators of signaling in the cell, and in the case of *Leishmania*, kinases are crucial to enable the different metabolic changes needed to adapt to a human host. Perhaps by intensive pharmacological investigation, drugs that are very successful in treating cancer (e.g., Gleevec) could be used against *Leishmania* parasites. One particular example from the group of predicted kinases detected on the network is the protein LMPK (*LmjF36.6470*). This protein has been shown to be essential in *Leishmania mexicana* (Wiese 1998) and it has conserved orthologs in other species such as *L. amazonensis*, *L. major*, *L. tropica*, *L. aethiopica*, *L. donovani*, *L. infantum*, and *L. braziliensis* (Wiese & Gorcke 2001). Therefore, this kinase was an interesting candidate for experimental validation and possibly its upstream and down-stream interacting partners could also be inhibited by a combination of drugs. In addition, one of the challenges in this disease is to find a broad-spectrum drug that can have therapeutic effects on several *Leishmania* species that cause different forms of leishmaniasis. Further analysis of this target can help to elucidate drugs or combination of drugs that are active against amastigotes, the stage responsible for the disease in mammals. Three ABC transporters that were *Leishmania* specific -

*LmjF34.0670*, *LmjF27.0470*, *LmjF32.2060* - were also predicted as essential. They confer resistance to antimonials and pentamidine by extruding the drug outside of the cell (Perez-Victoria., et al. 2002). Based upon our analysis, these proteins could be also interesting drug targets due to their role in the homeostasis of the intracellular parasite environment.

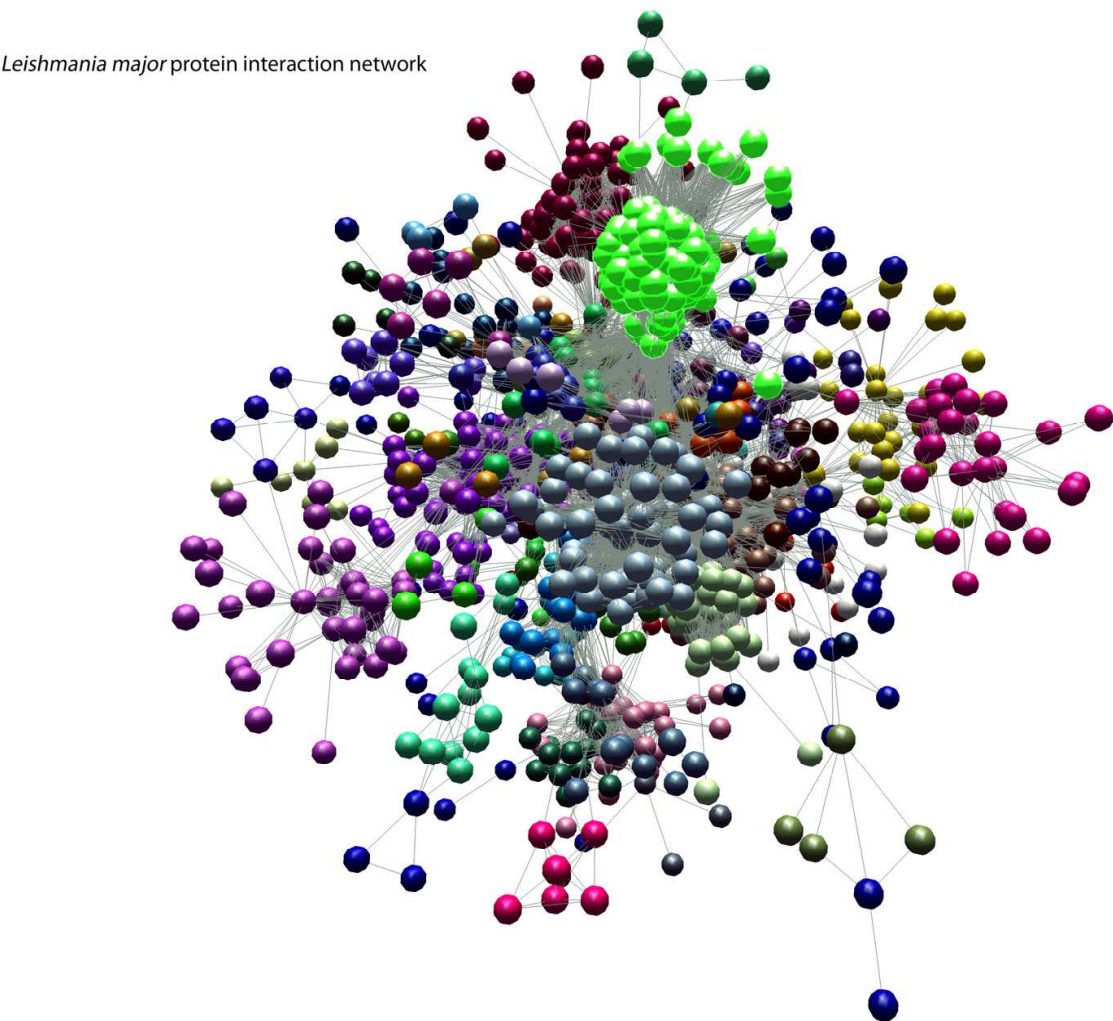


Fig. 2. Visualization of the predicted *Leishmania major* interactome

GeneDB ID	Uniprot ID	Description
LmjF15.0770	Q4QFA8	Protein kinase.
LmjF07.0250	Q4QIR9	Protein kinase
LmjF11.0330	Q4QH47	PIF1 helicase-like protein
LmjF35.2450	Q4FWM4	Hypothetical protein conserved
LmjF25.1990	Q4Q9P0	Protein kinase
LmjF21.0853	Q4QCC1	Hypothetical protein conserved
LmjF27.1800	Q4FYE1	Protein kinase-like
LmjF35.1000	Q4FX16	Casein kinase I
LmjF26.0660	Q4Q9C8	Protein disulfide isomerase
LmjF25.2050	Q4Q9N4	Helicase-like protein

Table 2. Top 10 list of predicted targets from the *L. major* interactome.

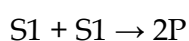
It has been shown that modular organization is a prevalent feature in biology, and this modular organization of pathways can be used to infer protein function (Rives & Galitski 2003). We detected 63 clusters or modules in the network, and assigned potential biological processes to 263 proteins with no prior functional description. By examining the proportion of predicted targets by biological process, 64% of the proteins in the network were predicted to participate in the protein phosphorylation (GO:0006468). In addition, 8% of proteins were predicted to be involved in nucleosome assembly (GO:0006334), 4% in nucleic acid metabolic process (GO:0006139), 4% in electron transport (GO:0006118), 4% in transport processes (GO:0006810), and 2% in protein amino acid alkylation (GO:0006139). The remaining 14% of target proteins were distributed across processes with one protein per process. This result highlighted the importance of protein kinases as the main protein class to characterize and explore as drug targets in *Leishmania* parasites.

### 3. Selection of drug targets by metabolic flux balance analysis and *in silico* deletions

Proteins involved in metabolism constitute another important source of drug targets. The energetic balance in the cell is controlled by enzymes that regulate the transformation of substrates in a coordinated and efficient manner. These enzymes are needed specifically for producing energy or as building blocks for other molecules being essential for the viability of the organism. However, a different approach needs to be used for modeling metabolism because the interactions between enzymes depend upon the rate of turnover of molecules or *fluxes*, not specifically through physical interactions as described for the case of the interactome.

The reconstruction of metabolic networks is more established compared to interactome generation. Since glycolysis was elucidated in 1930, several metabolic pathways have been discovered in many organisms. Metabolic networks reconstructed from this source of data started with *E. coli* (Reed & Palsson 2003) and was followed later on by reconstructions in eukaryotic organisms such as *Saccharomyces cerevisiae* (Duarte, et al. 2004) and *Aspergillus niger* (David, et al. 2003). More recently, metabolic networks of *Plasmodium falciparum* (Plata, et al. 2010) and *Leishmania major* (Chavali, et al. 2008) were reconstructed with the aim of detecting drug targets, and some details about the network generation and analysis will be discussed.

In order to build a metabolic network it is necessary to list all the substances with their concentrations and the reactions between substances. In living systems these reactions are catalyzed by enzymes and the transport processes are carried out by transporters or channels. The reactions are influenced by the *stoichiometric coefficients* which denote the proportion of substrate and product molecules involved in a reaction. The reaction:



describes the generation of product P from S1. Therefore, the stoichiometric coefficients for this particular reaction are -1,-1 and 2 respectively. For a metabolic network consisting of  $m$  substances and  $r$  reactions, the systems dynamics are described by *systems equations* (2) (or balance equations, since the balance of substrate production and degradation is considered):

$$\frac{dS_i}{dt} = \sum_{j=1}^r n_{ij} v_j \quad (2)$$

The term  $n_{ij}$  are the stoichiometric coefficients of metabolite  $i$  in reaction  $j$ . In this case, diffusion is not considered in the system. These equations can be applied to compartments, where the flux between compartments has to be considered as a different reaction. The stoichiometric coefficients  $n_{ij}$  can be combined into the so-called *stoichiometric matrix*, where each column belongs to a reaction and each row to a substance (Klipp et al., 2005).

According to this, a mathematical model for a metabolic network can be described as a system with a vector  $S = (S_1, S_2, \dots, S_n)$  of concentration values for the different species, a vector  $v = (v_1, v_2, \dots, v_r)$  of reaction rates and the stoichiometric matrix  $N$ . With these definitions, the balance equation (3) can be rewritten as follows:

$$\frac{dS}{dt} = Nv \quad (3)$$

### 3.1 Predicted drug targets in the *Leishmania major* metabolic network

The total number of genes included in the network was 560, with 1,112 reactions and 1,101 metabolites. The process for the reconstruction involved different data sources, in particular literature and biological databases. Reaction stoichiometry and subcellular localization were also extracted by examining the existing literature. Some reactions were assigned as non-gene-associated to account for spontaneously generated metabolites. The gene-associated reactions were further adjusted according to the specified constraints.

*Flux balance analysis* is a method that has been used extensively to analyze metabolic networks. The important advantage of this method is that does not require detailed knowledge of the enzyme kinetics. The principle of the method relies on investigation of the fluxes that have the greatest influence in the growth (or production of biomass) by preserving a set of constraints such as physicochemical, thermodynamic, topological and environmental. In the case of the *Leishmania* model, constraints included reaction reversibility rules, promastigote/amastigote protein expression data, various medium conditions and prevalence of transport reactions across cellular compartments were used. The model was simulated under *steady state* conditions, which means that the net change of any metabolite in the network during time should be zero.

Virtual knockouts were carried out over the network with the aim to detect potential drug targets. The knockout genes were classified as being lethal (essential for survival), growth reducing, or with no effect. From this analysis, only 12% of single knockouts were predicted as lethal and 10% as growth reducing. Approximately 83% of all lethal genes belonged to three metabolic processes: lipid, carbohydrate or amino-acid metabolism, highlighting how critical these are to general function.

From the group of lethal genes, the authors selected those that were exclusive or without human orthologs as potential candidates (a strategy that was also employed for the interactome analysis). The gene *LmjF05.0350*, which encodes for trypanothione reductase, was lethal *in silico*. This enzyme participates in the reduction of oxidative agents by using trypanothione and this molecule is only present in kinetoplastids. This enzyme has been studied extensively as drug target, confirming the predictions of the mathematical model (Eberle, et al. 2011). *LmjF31.2940* and *LmjF21.0845*, encoding for squalene synthase and hypoxanthine-guanine phosphoribosyltransferase respectively, were also predicted as lethal in the network. The squalene synthase inhibitors affect the sterol biosynthesis pathway, taking advantage of the trypanosomatid requirement for specific endogenous sterols (e.g.



ergosterol). Interestingly, proteins belonging to this process in *Leishmania* were also detected by homology and interactome analysis showing consistency between methods.

Double knockouts were also simulated to identify lethal combination of genes. Out of 152,520 double deletions, 19,341 were lethal. From this group, 19,285 double deletions were trivial lethal, which means that at least one of the genes involved was lethal in a single deletion. There were 56 non-trivial lethal double deletions that could be interesting to test experimentally. The main participation of these double knockouts was in the lipid and carbohydrate metabolism with 57.6% of the genes in these groups. One explanation for the large number of double deletions that were not essential is the high degree of redundancy in the network. These results show the utility of a different methodology that uses mathematical modeling for the detection of essential genes in metabolism.

#### 4. Discovery of new drugs using virtual screening

Once the protein target is identified, the next step is finding an effective and non-toxic inhibitor that can be administrated in human patients. However, the process for identifying compounds, as in the case for drug targets, requires expensive equipment for testing millions of chemicals with a final result of few *hits* that can advance in further drug development stages. This process is unpractical and time consuming. A computational strategy has been used to improve the drug search by using computational chemistry to select those compounds that are likely to work in the experimental setting, this computational technique is called *virtual screening*. The aim of this technique is to identify novel small molecules that could bind the target of interest. This is carried out by *docking* compound libraries over the protein structure of the target, using optimization algorithms that search for the best conformation of the target and the ligand, which by definition has the lowest conformational energy (McInnes 2007). The libraries are usually compounds available from chemical vendors or predicted in some cases.

The procedure for virtual screening involves the selection of the target as a first step. The target can be chosen by the results from RNAi screenings or from computational approaches as the ones previously described. The target must have a 3D structure with an acceptable resolution, which is critical for the accuracy of the predictions. The structures determined by experimental methods can be easily accessed through the Protein Data Bank (PDB; <http://www.pdb.org/>). This database contains protein structures from different organisms determined by X-ray crystallography or NMR methods. In addition, predicted structures by homology-modeling methods can be found in the ModBase database (<http://modbase.compbio.ucsf.edu/>) (Pieper., et al. 2009). In contrast, the resources for compounds or *ligands* are restricted due to patent protection. However an interesting resource that contains millions of ligands ready to use for virtual screening experiments is the ZINC database (<http://zinc.docking.org/>) (Irwin & Shoichet 2005). This database includes commercial ligands and it is also free for academic use. This combination of resources of protein structures and ligand databases facilitates enormously the development of virtual screening projects which can be very productive for finding new drugs, especially for neglected diseases such as leishmaniasis.

The tools employed for virtual screening are also very diverse. One very popular tool is AutoDock (Goodsell., et al. 1996), which was developed by the Scripps Institute and used in several virtual screening projects, demonstrating good performance and accuracy. However, a recent study (Chang., et al. 2010) compared the accuracy and reproducibility of AutoDock



against a recently developed tool, also freely available, called AutoDock Vina (Trott & Olson 2010). The experiments consisted of screening ligands with known activity against the HIV protease and *decoys* or non-binders. Autodock Vina performed very well in terms of speed, being ~10 times faster, and more accurate in ranking larger molecules compared to AutoDock. We are currently using AutoDock Vina in a virtual screening project called "Drug Search for Leishmaniasis" in association with IBM-World Community Grid (<http://www.worldcommunitygrid.org/research/dsfl/overview.do>) to speed up the process of finding new active compounds.

As an example of the application of this strategy in *Leishmania*, a recent study demonstrated the utility of virtual screening to identify potential MAPK inhibitors. The target, MAPK was first modeled by using *homology modeling* techniques. Essentially, the technique predicts the 3D structure of a particular protein by finding sequence homology to a model protein with experimentally determined structure. This model was refined by molecular dynamics. Structural features, such as ATP binding pocket, phosphorylation lip, and common docking site were identified. Virtual screening was carried out using this target with several compounds from the class of ATP inhibitors. Interestingly, the docking analysis suggested that the indirubin class of molecules could act as putative inhibitors of *Leishmania* MAPK. By testing this result experimentally, the authors found reasonably good correlation between *in vitro* activity and calculated binding energy for indirubin class of inhibitors obtained in the virtual screening study. These molecules make strong hydrogen bonding interactions with Lys43, Arg57, Asp155, Glu94, and Ile96 amino acid residues of the *Leishmania* MAPK model. These residues belong to the catalytic domain and inhibition of the catalytic domain leads to impaired kinase activity (Awale., et al. 2010). This is a clear example of the synergy between computational and experimental methods to accelerate drug discovery.

## 5. Selection of new drugs using machine learning techniques

The *Leishmania* proteome is estimated to contain ~8,150 proteins based on the annotated genome of the sequenced species (Peacock., et al. 2007). However, fewer than 150 proteins have 3D structures in the PDB. This limits the use of docking-based strategies to search for anti-Leishmania compounds. An alternative strategy to associate active compounds with *Leishmania* targets is by using machine learning techniques. This approach is intended to find patterns on protein targets such as domains, post-translational modifications etc, that can be linked to a specific class of compounds. This system will "learn" these patterns and when challenged by proteins from the organism of interest it will predict the potential association for a particular compound. Two studies have applied this strategy to a particular set of protein targets (Bulashevsk., et al. 2009; Thangudu., et al. 2010), employing the different techniques such as support vector machines (SVM) and Bayesian classifiers (BC). As a perspective, these methods could also be applied for drug search in *Leishmania*; however, the definition of protein patterns would be critical for establishing robust drug-target relationships.

## 6. Experimental approaches for drug testing in Leishmania

Several *in vitro* assays to test *Leishmania* susceptibility to new potential inhibitors have been developed for the two stages of *Leishmania*, namely promastigotes and amastigotes. These two stages are morphologically and biochemically different and these differences are likely responsible for their differing susceptibility to proven anti-leishmanial compounds. Assays developed with intracellular amastigotes have the advantage of being more "disease

appropriate" since this is the stage responsible for mammalian disease. In addition, an intracellular *in vitro* model resembles the natural event when the parasite is in the mammalian host. Axenic amastigotes are also employed, although a more efficient alternative is the use of intracellular fluorescent or bioluminescent amastigotes.

The effectiveness of compounds to kill the parasites has been evaluated using different methodological approaches. Several years ago, direct parasite counting was the most used method (Gaspar., et al. 1992; Chan-Bacab., et al. 2003; Khan., et al. 2003). However, this method lacked accuracy and precision, likely due to human errors. This made necessary to develop new automated methods based on colorimetric, radioactive, fluorescent and luminometer detection (Fumarola., et al. 2004). Colorimetric methods like MTT [3-(4,5-dimethylthiazol-2-yl)2,5-diphenyltetrazolium bromide] have been used frequently. Recently, these methods are being replaced by transgenic parasites with reporter genes that do not interfere with cellular mitochondria. Parasites genetically engineered to express green fluorescent protein (GFP) or luciferase, have been developed and are currently used in automatized protocols involving flow cytometer or luminometer.

On the other hand, it is important to measure cytotoxicity to evaluate the possibility that a compound might produce side effects in humans. Mammalian or mouse cell lines are usually employed in these assays. The most used cells are U-937 human histiocytes, TPH-1 human peripheral blood monocytes, and hamster peritoneal macrophages (Robledo., et al. 1999; Weniger., et al. 2001; Varela., et al. 2009; Taylor., et al. 2010). To increase the selectivity of promising drugs, liposomal formulations of the compounds may be evaluated in order to reduce the toxicity as was observed for amphotericin B (Mehta., et al. 1985; Lopez-Berestein 1987).

The leishmanicidal activity of compounds that show high anti-leishmanial activity and low toxicity for mammalian cells *in vitro*, is next evaluated *in vivo*. This is normally performed in mouse or the golden hamster (*Mesocricetus auratus*) models depending on the *Leishmania* subgenus (Travi., et al. 2002). Monkey models can also be used but these studies are limited to only a few laboratories worldwide (Grimaldi., et al. 2010).

## 7. Conclusion

Through the current chapter, the more relevant techniques for finding drugs and targets employing computational approaches were described. Special considerations should be made when using the homology searching approach for finding drug targets as the context of protein interactions is important for the definition of essentiality. Protein interaction network analysis together with metabolic flux balance analysis are becoming useful alternatives to understand protein function and to systematically select drug targets. The selection of new inhibitory compounds can be done by using virtual screening. Predicted structures can also be used in virtual screening when experimentally derived structures are absent. Finally, machine learning techniques are a promising option to search for anti-leishmanial drugs, especially when experimental or predicted structures are not available, as may occur with many *Leishmania* proteins. It is important to state that any computational analysis is considered exploratory, and experimental validation is necessary to guide final decisions about potential compounds that can advance to the next stage of drug discovery. However, by using computational tools the search space for drugs and targets is reduced, allowing more focused experiments that could reduce the cost and time of drug development.

## 8. Acknowledgment

This book chapter was supported by Colciencias, through the projects with the following contracts and codes: Contract 538 with code 111549326124 and contract 197 with code 111551929015.

## 9. References

- Ahasan, H. A., M. A. Chowdhury, M. A. Azhar, A. K. Rafiqueuddin & K. A. Azad (1996). "Deaths in visceral leishmaniasis (Kala-azar) during treatment." *Med J Malaysia* 51(1): 29-32.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller & D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res* 25(17): 3389-3402.
- Assenov, Y., F. Ramirez, S. E. Schelhorn, T. Lengauer & M. Albrecht (2008). "Computing topological parameters of biological networks." *Bioinformatics* 24(2): 282-284.
- Awale, M., V. Kumar, P. Saravanan & C. G. Mohan (2010). "Homology modeling and atomic level binding study of Leishmania MAPK with inhibitors." *J Mol Model* 16(3): 475-488.
- Bader, G. D., I. Donaldson, C. Wolting, B. F. Ouellette, T. Pawson & C. W. Hogue (2001). "BIND--The Biomolecular Interaction Network Database." *Nucleic Acids Res* 29(1): 242-245.
- Barabasi, A. L. & Z. N. Oltvai (2004). "Network biology: understanding the cell's functional organization." *Nat Rev Genet* 5(2): 101-113.
- Brohee, S., K. Faust, G. Lima-Mendez, G. Vanderstocken & J. van Helden (2008). "Network Analysis Tools: from biological networks to clusters and pathways." *Nat Protoc* 3(10): 1616-1629.
- Brohee, S. & J. van Helden (2006). "Evaluation of clustering algorithms for protein-protein interaction networks." *BMC Bioinformatics* 7: 488.
- Bulashevskaya, A., M. Stein, D. Jackson & R. Eils (2009). "Prediction of small molecule binding property of protein domains with Bayesian classifiers based on Markov chains." *Comput Biol Chem* 33(6): 457-460.
- Carrero-Lerida, J., G. Perez-Moreno, V. M. Castillo-Acosta, L. M. Ruiz-Perez & D. Gonzalez-Pacanowska (2009). "Intracellular location of the early steps of the isoprenoid biosynthetic pathway in the trypanosomatids *Leishmania major* and *Trypanosoma brucei*." *Int J Parasitol* 39(3): 307-314.
- Costanzo, M., A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Koh, K. Toufighi, S. Mostafavi, J. Prinz, R. P. St Onge, B. VanderSluis, T. Makhnevych, F. J. Vizeacoumar, S. Alizadeh, S. Bahr, R. L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z. Y. Lin, W. Liang, M. Marback, J. Paw, B. J. San Luis, E. Shuteriqi, A. H. Tong, N. van Dyk, I. M. Wallace, J. A. Whitney, M. T. Weirauch, G. Zhong, H. Zhu, W. A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pal, F. P. Roth, G. Giaever, C. Nislow, O. G. Troyanskaya, H. Bussey, G. D. Bader, A. C. Gingras, Q. D. Morris, P. M. Kim, C. A. Kaiser, C. L. Myers, B. J. Andrews & C. Boone (2010). "The genetic landscape of a cell." *Science* 327(5964): 425-431.
- Chan-Bacab, M. J., E. Balanza, E. Deharo, V. Munoz, R. D. Garcia & L. M. Pena-Rodriguez (2003). "Variation of leishmanicidal activity in four populations of *Urechites andrieuxii*." *J Ethnopharmacol* 86(2-3): 243-247.

- Chang, M. W., C. Ayeni, S. Breuer & B. E. Torbett (2010). "Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina." *PLoS One* 5(8): e11955.
- Chavali, A. K., J. D. Whittemore, J. A. Eddy, K. T. Williams & J. A. Papin (2008). "Systems analysis of metabolism in the pathogenic trypanosomatid *Leishmania major*." *Mol Syst Biol* 4: 177.
- Chua, H. N., W. K. Sung & L. Wong (2006). "Exploiting indirect neighbours and topological weight to predict protein function from protein-protein interactions." *Bioinformatics* 22(13): 1623-1630.
- Das, B. B., A. Ganguly & H. K. Majumder (2008). "DNA topoisomerases of *Leishmania*: the potential targets for anti-leishmanial therapy." *Adv Exp Med Biol* 625: 103-115.
- David, H., M. Akesson & J. Nielsen (2003). "Reconstruction of the central carbon metabolism of *Aspergillus niger*." *Eur J Biochem* 270(21): 4243-4253.
- de Azevedo, W. F., Jr. & M. B. Soares (2009). "Selection of targets for drug development against protozoan parasites." *Curr Drug Targets* 10(3): 193-201.
- Dreze, M., D. Monachello, C. Lurin, M. E. Cusick, D. E. Hill, M. Vidal & P. Braun (2010). "High-quality binary interactome mapping." *Methods Enzymol* 470: 281-315.
- Duarte, N. C., M. J. Herrgard & B. O. Palsson (2004). "Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model." *Genome Res* 14(7): 1298-1309.
- Eberle, C., B. S. Lauber, D. Fankhauser, M. Kaiser, R. Brun, R. L. Krauth-Siegel & F. Diederich (2011). "Improved inhibitors of trypanothione reductase by combination of motifs: synthesis, inhibitory potency, binding mode, and antiprotozoal activities." *ChemMedChem* 6(2): 292-301.
- Enright, A. J., S. Van Dongen & C. A. Ouzounis (2002). "An efficient algorithm for large-scale detection of protein families." *Nucleic Acids Res* 30(7): 1575-1584.
- Faraut-Gambarelli, F., R. Piarroux, M. Deniau, B. Giusiano, P. Marty, G. Michel, B. Faugere & H. Dumon (1997). "In vitro and in vivo resistance of *Leishmania infantum* to meglumine antimoniate: a study of 37 strains collected from patients with visceral leishmaniasis." *Antimicrob Agents Chemother* 41(4): 827-830.
- Finn, R. D., M. Marshall & A. Bateman (2005). "iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions." *Bioinformatics* 21(3): 410-412.
- Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. Sonnhammer & A. Bateman (2006). "Pfam: clans, web tools and services." *Nucleic Acids Res* 34(Database issue): D247-251.
- Florez, A. F., D. Park, J. Bhak, B. C. Kim, A. Kuchinsky, J. H. Morris, J. Espinosa & C. Muskus (2010). "Protein network prediction and topological analysis in *Leishmania major* as a tool for drug target selection." *BMC Bioinformatics* 11: 484.
- Forsyth, R. A., R. J. Haselbeck, K. L. Ohlsen, R. T. Yamamoto, H. Xu, J. D. Trawick, D. Wall, L. Wang, V. Brown-Driver, J. M. Froelich, K. G. C, P. King, M. McCarthy, C. Malone, B. Misiner, D. Robbins, Z. Tan, Z. Y. Zhu Zy, G. Carr, D. A. Mosca, C. Zamudio, J. G. Foulkes & J. W. Zyskind (2002). "A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*." *Mol Microbiol* 43(6): 1387-1400.
- Fumarola, L., R. Spinelli & O. Brandonisio (2004). "In vitro assays for evaluation of drug activity against *Leishmania* spp." *Res Microbiol* 155(4): 224-230.
- Gaspar, R., F. R. Oppendoes, V. Preat & M. Roland (1992). "Drug targeting with polyalkylcyanoacrylate nanoparticles: in vitro activity of primaquine-loaded



- nanoparticles against intracellular *Leishmania donovani*." *Ann Trop Med Parasitol* 86(1): 41-49.
- Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A. P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kotter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis & M. Johnston (2002). "Functional profiling of the *Saccharomyces cerevisiae* genome." *Nature* 418(6896): 387-391.
- Glasser, J. S. & C. K. Murray (2011). "Central nervous system toxicity associated with liposomal amphotericin B therapy for cutaneous leishmaniasis." *Am J Trop Med Hyg* 84(4): 566-568.
- Good, M. C., J. G. Zalatan & W. A. Lim (2011). "Scaffold proteins: hubs for controlling the flow of cellular information." *Science* 332(6030): 680-686.
- Goodsell, D. S., G. M. Morris & A. J. Olson (1996). "Automated docking of flexible ligands: applications of AutoDock." *J Mol Recognit* 9(1): 1-5.
- Goyeneche-Patino, D. A., L. Valderrama, J. Walker & N. G. Saravia (2008). "Antimony resistance and trypanothione in experimentally selected and clinical strains of *Leishmania panamensis*." *Antimicrob Agents Chemother* 52(12): 4503-4506.
- Grimaldi, G., Jr., R. Porrozzio, K. Friedrich, A. Teva, R. S. Marchevsky, F. Vieira, N. Miekeley & F. J. Paumgartten (2010). "Comparative efficacies of two antimony regimens to treat *Leishmania braziliensis*-induced cutaneous Leishmaniasis in rhesus macaques (*Macaca mulatta*)." *Antimicrob Agents Chemother* 54(1): 502-505.
- He, F., Y. Zhang, H. Chen, Z. Zhang & Y. L. Peng (2008). "The prediction of protein-protein interaction networks in rice blast fungus." *BMC Genomics* 9: 519.
- Hermjakob, H., L. Montecchi-Palazzi, C. Lewington, S. Mudali, S. Kerrien, S. Orchard, M. Vingron, B. Roechert, P. Roepstorff, A. Valencia, H. Margalit, J. Armstrong, A. Bairoch, G. Cesareni, D. Sherman & R. Apweiler (2004). "IntAct: an open source molecular interaction database." *Nucleic Acids Res* 32(Database issue): D452-455.
- Hu, W., S. Sillaots, S. Lemieux, J. Davison, S. Kauffman, A. Breton, A. Linteau, C. Xin, J. Bowman, J. Becker, B. Jiang & T. Roemer (2007). "Essential gene identification and drug target prioritization in *Aspergillus fumigatus*." *PLoS Pathog* 3(3): e24.
- Hubbard, T. J., A. G. Murzin, S. E. Brenner & C. Chothia (1997). "SCOP: a structural classification of proteins database." *Nucleic Acids Res* 25(1): 236-239.
- Irwin, J. J. & B. K. Shoichet (2005). "ZINC--a free database of commercially available compounds for virtual screening." *J Chem Inf Model* 45(1): 177-182.
- Ivens, A. C., C. S. Peacock, E. A. Worthey, L. Murphy, G. Aggarwal, M. Berriman, E. Sisk, M. A. Rajandream, E. Adlem, R. Aert, A. Anupama, Z. Apostolou, P. Attipoe, N. Bason, C. Bauser, A. Beck, S. M. Beverley, G. Bianchetti, K. Borzym, G. Bothe, C. V. Bruschi, M. Collins, E. Cadag, L. Ciarloni, C. Clayton, R. M. Coulson, A. Cronin, A. K. Cruz, R. M. Davies, J. De Gaudenzi, D. E. Dobson, A. Duesterhoeft, G. Fazelina, N. Fosker, A. C. Frasch, A. Fraser, M. Fuchs, C. Gabel, A. Goble, A. Goffeau, D. Harris, C. Hertz-Fowler, H. Hilbert, D. Horn, Y. Huang, S. Klages, A. Knights, M. Kube, N. Larke, L. Litvin, A.



- Lord, T. Louie, M. Marra, D. Masuy, K. Matthews, S. Michaeli, J. C. Mottram, S. Muller-Auer, H. Munden, S. Nelson, H. Norbertczak, K. Oliver, S. O'Neil, M. Pentony, T. M. Pohl, C. Price, B. Purnelle, M. A. Quail, E. Rabbinowitsch, R. Reinhardt, M. Rieger, J. Rinta, J. Robben, L. Robertson, J. C. Ruiz, S. Rutter, D. Saunders, M. Schafer, J. Schein, D. C. Schwartz, K. Seeger, A. Seyler, S. Sharp, H. Shin, D. Sivam, R. Squares, S. Squares, V. Tosato, C. Vogt, G. Volckaert, R. Wambutt, T. Warren, H. Wedler, J. Woodward, S. Zhou, W. Zimmermann, D. F. Smith, J. M. Blackwell, K. D. Stuart, B. Barrell & P. J. Myler (2005). "The genome of the kinetoplastid parasite, *Leishmania major*." *Science* 309(5733): 436-442.
- Jeong, H., S. P. Mason, A. L. Barabasi & Z. N. Oltvai (2001). "Lethality and centrality in protein networks." *Nature* 411(6833): 41-42.
- Kamath, R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, M. Gotta, A. Kanapin, N. Le Bot, S. Moreno, M. Sohrmann, D. P. Welchman, P. Zipperlen & J. Ahringer (2003). "Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi." *Nature* 421(6920): 231-237.
- Khan, K. M., M. Rasheed, U. Zia, S. Hayat, F. Kaukab, M. I. Choudhary, R. Atta ur & S. Perveen (2003). "Synthesis and in vitro leishmanicidal activity of some hydrazides and their analogues." *Bioorg Med Chem* 11(7): 1381-1387.
- Kim, J. G., D. Park, B. C. Kim, S. W. Cho, Y. T. Kim, Y. J. Park, H. J. Cho, H. Park, K. B. Kim, K. O. Yoon, S. J. Park, B. M. Lee & J. Bhak (2008). "Predicting the interactome of *Xanthomonas oryzae* pathovar *oryzae* for target selection and DB service." *BMC Bioinformatics* 9: 41.
- Klipp, E., Herwig, R., Kowald, A., Wierling, C. & Lehrach, H (2005). *Systems Biology in Practice: Concepts, Implementation and Application*. Wiley-VCH Verlag GmbH & Co. KGaA, ISBN: 9783527310784, Weinheim, Germany.
- Knox, C., V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo & D. S. Wishart (2011). "DrugBank 3.0: a comprehensive resource for 'omics' research on drugs." *Nucleic Acids Res* 39(Database issue): D1035-1041.
- Kuhn, M., D. Szklarczyk, A. Franceschini, M. Campillos, C. von Mering, L. J. Jensen, A. Beyer & P. Bork (2010). "STITCH 2: an interaction network database for small molecules and proteins." *Nucleic Acids Res* 38(Database issue): D552-556.
- Lang, T., S. Goyard, M. Lebastard & G. Milon (2005). "Bioluminescent *Leishmania* expressing luciferase for rapid and high throughput screening of drugs acting on amastigote-harboring macrophages and for quantitative real-time monitoring of parasitism features in living mice." *Cell Microbiol* 7(3): 383-392.
- Lin, C. Y., C. H. Chin, H. H. Wu, S. H. Chen, C. W. Ho & M. T. Ko (2008). "Hubba: hub objects analyzer--a framework of interactome hubs identification for network biology." *Nucleic Acids Res* 36(Web Server issue): W438-443.
- Lopez-Berestein, G. (1987). "Liposomes as carriers of antimicrobial agents." *Antimicrob Agents Chemother* 31(5): 675-678.
- Maltezou, H. C. (2008). "Visceral leishmaniasis: advances in treatment." *Recent Pat Antiinfect Drug Discov* 3(3): 192-198.
- McInnes, C. (2007). "Virtual screening strategies in drug discovery." *Curr Opin Chem Biol* 11(5): 494-502.
- Mehta, R. T., G. Lopez-Berestein, R. L. Hopfer, K. Mehta, R. A. White & R. L. Juliano (1985). "Prophylaxis of murine candidiasis via application of liposome-encapsulated

- amphotericin B and a muramyl dipeptide analog, alone and in combination." *Antimicrob Agents Chemother* 28(4): 511-513.
- Oliveira, L. F., A. O. Schubach, M. M. Martins, S. L. Passos, R. V. Oliveira, M. C. Marzochi & C. A. Andrade (2011). "Systematic review of the adverse effects of cutaneous leishmaniasis treatment in the New World." *Acta Trop* 118(2): 87-96.
- Osman, A. (2004). "Yeast two-hybrid assay for studying protein-protein interactions." *Methods Mol Biol* 270: 403-422.
- Park, D., S. Lee, D. Bolser, M. Schroeder, M. Lappe, D. Oh & J. Bhak (2005). "Comparative interactomics analysis of protein family interaction networks using PSIMAP (protein structural interactome map)." *Bioinformatics* 21(15): 3234-3240.
- Peacock, C. S., K. Seeger, D. Harris, L. Murphy, J. C. Ruiz, M. A. Quail, N. Peters, E. Adlem, A. Tivey, M. Aslett, A. Kerhornou, A. Ivens, A. Fraser, M. A. Rajandream, T. Carver, H. Norbertczak, T. Chillingworth, Z. Hance, K. Jagels, S. Moule, D. Ormond, S. Rutter, R. Squares, S. Whitehead, E. Rabbinowitsch, C. Arrowsmith, B. White, S. Thurston, F. Bringaud, S. L. Baldauf, A. Faulconbridge, D. Jeffares, D. P. Depledge, S. O. Oyola, J. D. Hilley, L. O. Brito, L. R. Tosi, B. Barrell, A. K. Cruz, J. C. Mottram, D. F. Smith & M. Berriman (2007). "Comparative genomic analysis of three *Leishmania* species that cause diverse human disease." *Nat Genet* 39(7): 839-847.
- Perez-Victoria, J. M., A. Di Pietro, D. Barron, A. G. Ravelo, S. Castanys & F. Gamarro (2002). "Multidrug resistance phenotype mediated by the P-glycoprotein-like transporter in *Leishmania*: a search for reversal agents." *Curr Drug Targets* 3(4): 311-333.
- Peri, S., J. D. Navarro, T. Z. Kristiansen, R. Amanchy, V. Surendranath, B. Muthusamy, T. K. Gandhi, K. N. Chandrika, N. Deshpande, S. Suresh, B. P. Rashmi, K. Shanker, N. Padma, V. Niranjana, H. C. Harsha, N. Talreja, B. M. Vrushabendra, M. A. Ramya, A. J. Yatish, M. Joy, H. N. Shivashankar, M. P. Kavitha, M. Menezes, D. R. Choudhury, N. Ghosh, R. Saravana, S. Chandran, S. Mohan, C. K. Jonnalagadda, C. K. Prasad, C. Kumar-Sinha, K. S. Deshpande & A. Pandey (2004). "Human protein reference database as a discovery resource for proteomics." *Nucleic Acids Res* 32(Database issue): D497-501.
- Pieper, U., N. Eswar, B. M. Webb, D. Eramian, L. Kelly, D. T. Barkan, H. Carter, P. Mankoo, R. Karchin, M. A. Marti-Renom, F. P. Davis & A. Sali (2009). "MODBASE, a database of annotated comparative protein structure models and associated resources." *Nucleic Acids Res* 37(Database issue): D347-354.
- Plata, G., T. L. Hsiao, K. L. Olszewski, M. Llinas & D. Vitkup (2010). "Reconstruction and flux-balance analysis of the *Plasmodium falciparum* metabolic network." *Mol Syst Biol* 6: 408.
- Reed, J. L. & B. O. Palsson (2003). "Thirteen years of building constraint-based in silico models of *Escherichia coli*." *J Bacteriol* 185(9): 2692-2699.
- Rives, A. W. & T. Galitski (2003). "Modular organization of cellular networks." *Proc Natl Acad Sci U S A* 100(3): 1128-1133.
- Robledo, S. M., A. Z. Valencia & N. G. Saravia (1999). "Sensitivity to Glucantime of *Leishmania viannia* isolated from patients prior to treatment." *J Parasitol* 85(2): 360-366.
- Sakharkar, K. R., M. K. Sakharkar & V. T. Chow (2004). "A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*." *In Silico Biol* 4(3): 355-360.
- Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski & T. Ideker (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome Res* 13(11): 2498-2504.

- Stark, C., B. J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz & M. Tyers (2006). "BioGRID: a general repository for interaction datasets." *Nucleic Acids Res* 34(Database issue): D535-539.
- Sundar, S. & J. Chakravarty (2010). "Antimony toxicity." *Int J Environ Res Public Health* 7(12): 4267-4277.
- Taylor, V. M., D. L. Munoz, D. L. Cedeno, I. D. Velez, M. A. Jones & S. M. Robledo (2010). "Leishmania tarentolae: utility as an in vitro model for screening of antileishmanial agents." *Exp Parasitol* 126(4): 471-475.
- Thangudu, R. R., M. Tyagi, B. A. Shoemaker, S. H. Bryant, A. R. Panchenko & T. Madej (2010). "Knowledge-based annotation of small molecule binding sites in proteins." *BMC Bioinformatics* 11: 365.
- Travi, B. L., Y. Osorio, P. C. Melby, B. Chandrasekar, L. Arteaga & N. G. Saravia (2002). "Gender is a major determinant of the clinical evolution and immune response in hamsters infected with Leishmania spp." *Infect Immun* 70(5): 2288-2296.
- Trott, O. & A. J. Olson (2010). "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading." *J Comput Chem* 31(2): 455-461.
- Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields & J. M. Rothberg (2000). "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*." *Nature* 403(6770): 623-627.
- Varela, M. R., D. L. Munoz, S. M. Robledo, B. K. Kolli, S. Dutta, K. P. Chang & C. Muskus (2009). "Leishmania (Viannia) panamensis: an in vitro assay using the expression of GFP for screening of antileishmanial drug." *Exp Parasitol* 122(2): 134-139.
- von Mering, C., L. J. Jensen, B. Snel, S. D. Hooper, M. Krupp, M. Foglierini, N. Jouffre, M. A. Huynen & P. Bork (2005). "STRING: known and predicted protein-protein associations, integrated and transferred across organisms." *Nucleic Acids Res* 33(Database issue): D433-437.
- Weniger, B., S. Robledo, G. J. Arango, E. Deharo, R. Aragon, V. Munoz, J. Callapa, A. Lobstein & R. Anton (2001). "Antiprotozoal activities of Colombian plants." *J Ethnopharmacol* 78(2-3): 193-200.
- Wiese, M. (1998). "A mitogen-activated protein (MAP) kinase homologue of *Leishmania mexicana* is essential for parasite survival in the infected host." *EMBO J* 17(9): 2619-2628.
- Wiese, M. & I. Gorcke (2001). "Homologues of LMPK, a mitogen-activated protein kinase from *Leishmania mexicana*, in different *Leishmania* species." *Med Microbiol Immunol* 190(1-2): 19-22.
- Xenarios, I., D. W. Rice, L. Salwinski, M. K. Baron, E. M. Marcotte & D. Eisenberg (2000). "DIP: the database of interacting proteins." *Nucleic Acids Res* 28(1): 289-291.
- Yu, H., P. M. Kim, E. Sprecher, V. Trifonov & M. Gerstein (2007). "The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics." *PLoS Comput Biol* 3(4): e59.
- Zanzoni, A., L. Montecchi-Palazzi, M. Quondam, G. Ausiello, M. Helmer-Citterich & G. Cesareni (2002). "MINT: a Molecular INteraction database." *FEBS Lett* 513(1): 135-140.
- Zhang, R. & Y. Lin (2009). "DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes." *Nucleic Acids Res* 37(Database issue): D455-458.



### **Current Topics in Tropical Medicine**

Edited by Dr. Alfonso Rodriguez-Morales

ISBN 978-953-51-0274-8

Hard cover, 564 pages

**Publisher** InTech

**Published online** 16, March, 2012

**Published in print edition** March, 2012

Tropical Medicine has emerged and remained as an important discipline for the study of diseases endemic in the tropic, particularly those of infectious etiology. Emergence and reemergence of many tropical pathologies have recently aroused the interest of many fields of the study of tropical medicine, even including new infectious agents. Then evidence-based information in the field and regular updates are necessary. Current Topics in Tropical Medicine presents an updated information on multiple diseases and conditions of interest in the field. It includes pathologies caused by bacteria, viruses and parasites, protozoans and helminths, as well as tropical non-infectious conditions. Many of them are considering not only epidemiological aspects, but also diagnostic, therapeutic, preventive, social, genetic, bioinformatic and molecular ones. With participation of authors from various countries, many from proper endemic areas, this book has a wide geographical perspective. Finally, all of these characteristics, make an excellent update on many aspects of tropical medicine in the world.

#### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Andrés F. Flórez, Stanley Watowich and Carlos Muskus (2012). Current Advances in Computational Strategies for Drug Discovery in Leishmaniasis, Current Topics in Tropical Medicine, Dr. Alfonso Rodriguez-Morales (Ed.), ISBN: 978-953-51-0274-8, InTech, Available from: <http://www.intechopen.com/books/current-topics-in-tropical-medicine/current-advances-in-computational-strategies-for-drug-discovery-in-leishmaniasis->

**INTech**  
open science | open minds

#### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

#### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen