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

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

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

16

Stochastic Simulation
for Biochemical Reaction
Networks in Infectious Disease

Shailza Singh and Sonali Shinde
National Centre for Cell Science NCCS Complex, Ganeskhind,
*Pune University Campus, Pune*
India

1. Introduction

Modern computational science is converging on complex problems in the general field of systems biology. There is now a credible possibility of modeling drug delivery vesicles (liposomes) and their properties with qualitative and quantitative insight coming from atomistic calculations. Understanding the formation of vesicles from phospholipids bilayers and their fluidity and permeability properties is the basis of a large number of applications in the domain of drug delivery, in particular release of the active species according to the pH or ionic concentration changes. Prediction of structural changes (phase transition in particular) of membranes by modification of one or several constituents or addition of external molecular species may have potential therapeutic applications. Understanding the basic principles of biomembranes (lipid bilayers) governing and mediating various biologically relevant processes on the microscopic level is one of the great challenges in biology (Singh 2011). Molecular dynamic (MD) simulations are well suited for detailed analysis of the interactions between lipid bilayers and various small molecules, including water, chemicals, co-enzymes, peptides, oligonucleotides and proteins, as evidenced by the extensive body of published literature. Density Functional Theory (DFT) methods and recent theoretical studies of the properties of phospholipid molecules, consistent QM/MM and MM methodologies as well as their counterparts as function of time, Born-Oppenheimer dynamics (BOMD) and molecular dynamics (MD) are applied to understand and predict the properties of liposomes in water, in particular, fluidity, and their evolution in the presence of ions and biological agents. This has subsequently been achieved through the Dipalmitoylphosphatidylcholine (DPPC) bilayer systems and its partitioning with an ether lipid drug. Hence, membrane-lipid therapy plays a dominant role in molecular medicine.

1.1 Cellular model of lipid membranes and drug partitioning

The model of the DPPC bilayer was used for simulation of the phospholipid environment around the drug. The lipid parameters were taken from the literature (Lipka et al., 1984). The partial atomic charges were calculated at DFT/B3LYP/6- 31G** level using Gaussian 09 package. Also, the partitioning of large molecular systems was done by following ONIOM (Our-own-N-layered integrated molecular orbital and molecular mechanics) method,

whereby the whole system is partitioned into onion-skin-like layers, using QM/MM. The total energy can be obtained from following:

$$E = E[Low, real] + E[Med, int] + E[High, inner] - E[Low, int] - E[Med, inner] = E6 + E5 + E4 - E3 - E2$$
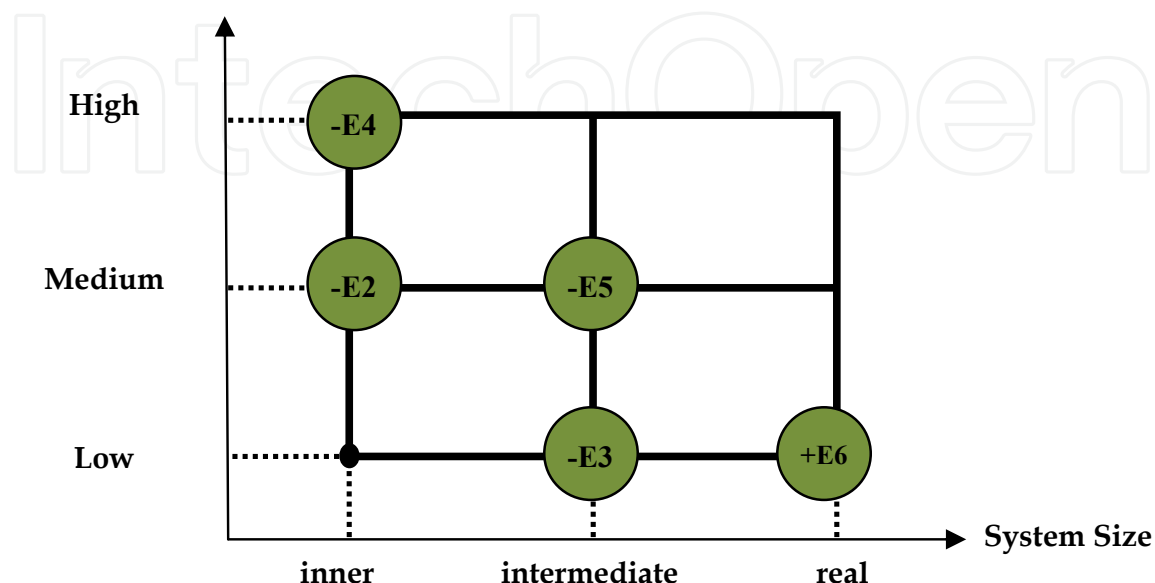
where real denotes the entire system, which is treated at low level, while int denotes the part of the system partitioned to be the intermediate layers for which the energy is calculated at both medium and high levels. Here inner denotes the inner layer of the system partitioning, whose energy is calculated at both high and medium levels.

Chemical potential profile, μ(z), of solute across a lipid bilayer and adjacent water phase is obtained by inserting the solute numerous times into randomly selected positions in the system obtained by MD simulation and calculating the interaction energy, E(z), between the inserted solute and all the molecules in the system. From <exp (−E(z)/RT)>t, where <…>t denotes the thermal average over insertions of solute with randomly chosen orientations into configurations of the system at depth z unperturbed by the solute, the excess chemical potential, μ(z), and thereby the free energy of transfer, ΔG(z), from bulk water to the bilayer interior at the depth z are obtained: ΔG(z) = μ (z) – μ (water)

$$= -RT\ln <e^{-E(z)/RT}> / <e^{-E(water)/RT}>$$

This simulation may enable us to understand the thermodynamic interactions between the lipid and drug molecules, the role of electrostatics in such adducts, and the permeability of drug (Fig. 1A & 1B). Degree of penetration into the membrane interior obtained from computer simulations may also be useful in estimating the extent to which a drug or prodrug that is unstable in a given solution or biological fluid might be protected when bound to a lipid bilayer membrane (Fig.1B&C). Multi-scale modeling and simulation revealed that the interactions between the drug and membrane are both electrostatic and hydrophobic. Increasing the chain length and lipophilicity may strengthen the binding interactions and overcome the problem of drug resistance.
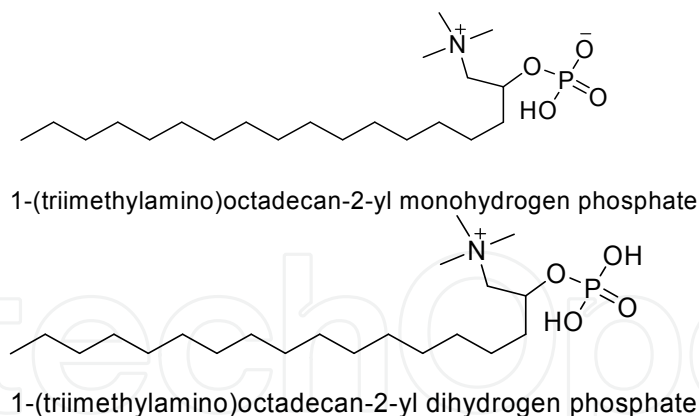
1-(triimethylamino)octadecan-2-yl monohydrogen phosphate

1-(triimethylamino)octadecan-2-yl dihydrogen phosphate

Fig. 1. A) Neutral and charged states of miltefosine
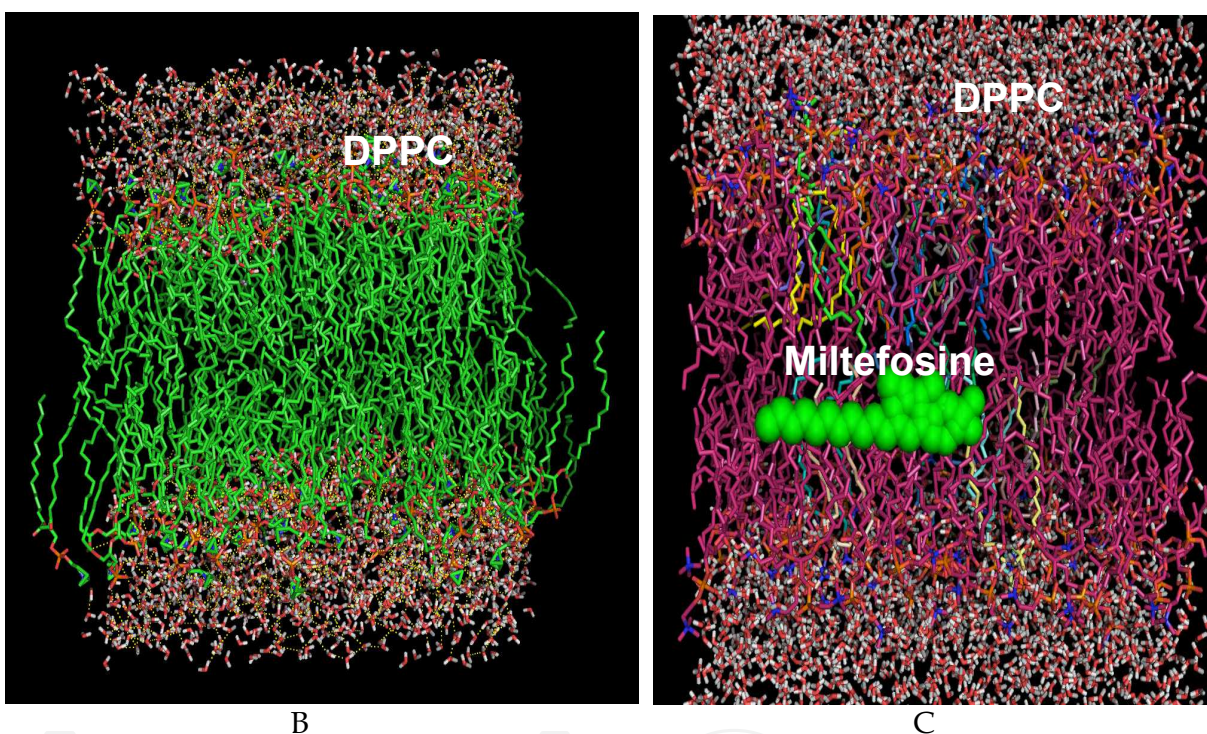

B   C

Fig. 1. B) Model representing DPPC bilayer system, C) Miltefosine with the lipid bilayer system.

2. Membrane permeability, lipid metabolism and multi-drug resistance in *Leishmania major*

Membrane impermeability is the major contributing factor to multidrug resistance in clinical isolates of *Leishmania major* (Perez Victoria et al., 2006). Reductionism which dominated biological research for over a century has provided a wealth of information about individual cellular components and their functions. Understanding the structure and the dynamics of the complex intercellular web of interactions that contribute to the structure and function of a living cell is a key challenge for biology in the 21st century.

The recent developments in genome sequencing (complete for *L. major, L. braziliensis*, and *L. infantum*) and lipidomic tools conceive a powerful platform for the study of lipid

metabolism in Leishmania. Lipid concentration changes in biological systems reflect regulation at multiple spatial and dynamic scales, e.g., biochemical reactions in the cells, intercellular lipid trafficking, changes in cell membrane composition, systemic lipid metabolism or lipid oxidation [Niemela et al., 2009]. Excluding the transport reactions, the percentage of intracellular reactions participating in lipid metabolism in *L.major* was the greatest, when compared to reactions across other metabolic subsystems [Chavali et al., 2008]. In order to address the intricacies of lipid regulation, it is quite obligatory to analyze the lipid metabolic mechanisms (Fig.2).
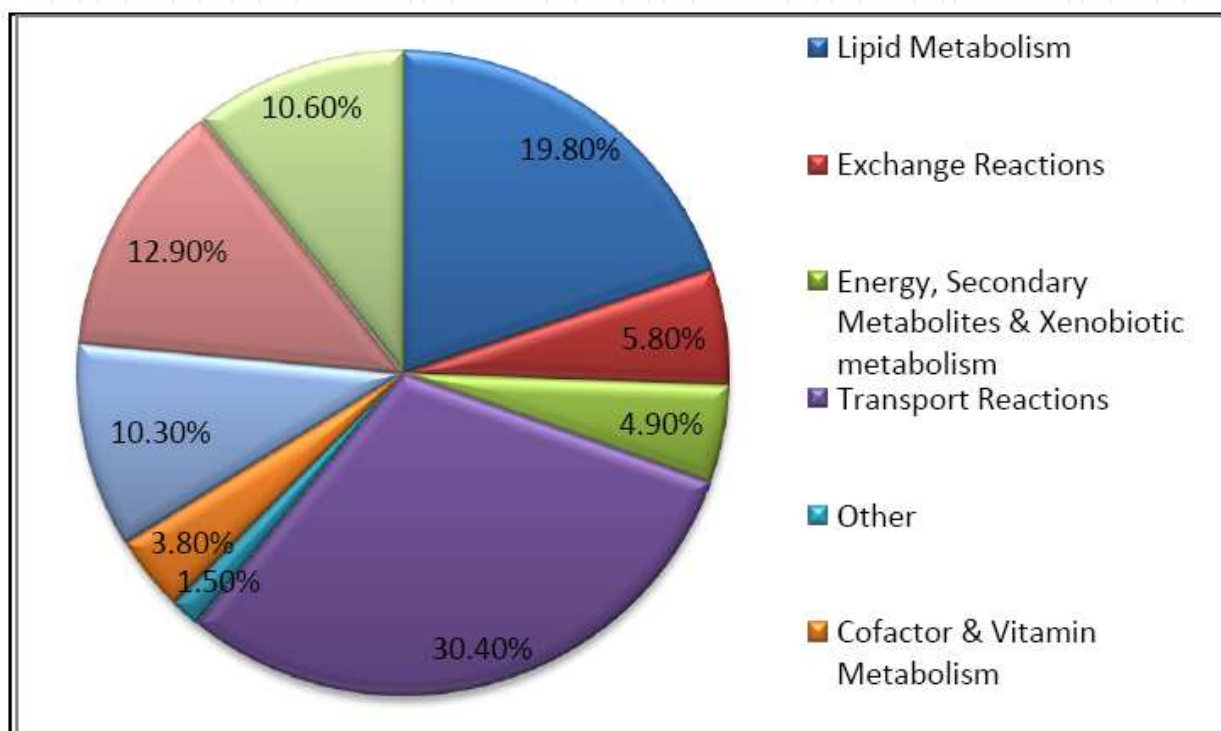

Fig. 2. Distribution of Metabolic Pathways according to the Cellular Processes. Neglecting the Transport Reactions, the Lipid Metabolism shows the highest no. of intracellular reactions signifying the abundance of the lipid biochemical networks.

The *L.major* pathogen is decoded at the interactome level, and the study is converged into the Lipid Metabolic Pathway, to serve the purpose of targeting the probable ligand via Liposomal Drug Delivery. Lipids play an active role in a variety of dynamic processes involving the membranes that compartmentalize the cell. To achieve this, cell must regulate the mechanical properties of the membrane and it can do so partly by controlling its lipid composition. The underlying biophysical question is to understand the variability and chemical diversity of membrane lipid composition, the mechanical properties of the membrane and the associated protein functions. Lipid metabolism is essential to the development, multiplication, virulence and differentiation of Leishmania species in the host, thus making the pathways for synthesis of parasite lipids good targets for development of new anti-leishmanial drugs [Zufferey and Mamoun, 2005; Zhang K. et al., 2003; Zufferey et al., 2003; Haughan et al.,1995]. Thereby, the diversity between Leishmania and mammals in lipid composition could be exploited towards selective chemotherapy. Lipids being a crucial entity of the pathogen cell surface are given

extensive importance and moreover due to the liposomal drug delivery system, targeting the lipids organized in the membrane of the pathogen would be an added advantage for the smooth release of the encapsulated drug.

The cell surface of Leishmania spp. has been shown to play a role in host-pathogen interactions and in the ability of the pathogen to evade the host-immune system. LPGs are essential for adhesion of the parasite to the midgut of the insect and therefore are important for transmission of the parasite to the human host. The basic LPG structure in all Leishmania species consists of several domains like the GPI anchor, with an 1-O-alkyl-2-lyso-phosphatidyl inositol lipid joined to a heptasaccharide glycan core, a long phosphoglycan (PG) domain composed of (Gal β1-4manα1-PO4)n repeat units (n=15-30), and a small neutral oligosaccharide cap [Bergter and Vermelho, 2010]. In different Leishmania species the phosphoglycan repeat units contain additional substitutions that mediate key roles in stage-specific adhesion. Deletion of LPG in *L.major* indicates that the glycosylated structures are involved in resistance to oxidative stress and the human immune system [Oppenheimer et al., 2011]. Henceforth, it is evident that lipids are essential cell constituents and therefore must be constantly synthesized to allow multiplication of the parasite. This suggests that the pathways leading to their synthesis are essential for parasite proliferation and pathogenesis and thus offers a reasonable target for rational design of novel antileishmanial drugs. Gal*f* plays an important role in host specific cell recognition, parasitic growth and pathogenesis. Since Gal*f* is not present in humans, the Galf biosynthetic pathway is an attractive target for the development of novel anti-parasitic drugs. In this pathway, UDP-galactopyranose mutase (UGM) catalyzes the conversion of UDP-galactopyranose to form UPD-Galf, which serves as a precursor for all the Galf found on the cell surface. Deletion of UGM gene may lead to the identification of specific inhibitors of this enzyme. [Oppenheimer et al., 2011]

Interactome Network Building is a crucial step for advancing towards Mathematical modeling, using the Systems Biology Markup Language. Metabolic network reconstruction has become an indispensable tool for studying the systems biology of metabolism. The focus is laid onto reviewing the mathematical foundations for analyzing chemical reactions, and describing how these systems of coupled chemical reactions can provide insight into the behavior of complex regulatory mechanisms. This is accomplished using the simulation mechanisms based on the certain defined kinetic laws such as Convenience Kinetics, Generalized Mass Action Kinetics and the Hill-Hinze equation. A saturable rate law "Convenience Kinetics" - a generalised form of Michaelis-Menten kinetics, covers all possible stoichiometries, describes enzyme regulation by activators and inhibitors, and can be derived from a rapid-equilibrium random-order enzyme mechanism. Convenience kinetics is used to translate a biochemical network – manually into a dynamical model with rational biological properties. It implements enzyme saturation and regulation by activators and inhibitors, covers all possible reaction stoichiometries, and can be specified by a small number of parameters. Their mathematical forms make it especially suitable for parameter assessment and optimisation. To ensure thermodynamic correctness, it is written in terms of thermodynamically independent parameters [Liebermeister and Klipp, 2006]. Moreover, the classes of generalized kinetic expressions of GMA also ensure consistency with the extended thermodynamic conditions [Horn and Jackson, 2004].

Simulation can provide the ultimate detail concerning individual particle motions as a function of time [Karplus and McCammon, 2002]. Thus, they can be used to address specific questions about the properties of a model system, often more easily than experiments on the actual system. These techniques are carried out upon treating a biological process as a system of equations, represented by their rate constants and other parameters and simulating their interactions through numerical techniques. The numerical simulations capture the effects of genes and their expression level through the time-course of evolution at molecular concentration. Thus, it is a biochemical characterization of the enzymatic network using the systems biology approach by interconnecting mathematical, biological, and physical sciences with the aim to capture the behavioural sciences of the enzymes, in the temporal and spatial scales, eventually leading to the discrete stochastic kinetic simulation. The integration of genomic and physiological information is now increasingly important with the emergence of "systems biology", which attempts to simultaneously study the expression patterns and activity of all genes, together with proteomic and metabolomic data [Cavalieri and Filippo, 2005]. When combined with a specific method of analysis like flux balance analysis (FBA), constraint-based models can be used to generate quantitative predictions (e.g., growth rate of an organism) and yield testable hypotheses for future experimental investigations [Roberts et al., 2009; Lee et al., 2006]. This permits an iterative process of model development, hypothesis generation and testing, and further model development and refinement [Reed et al., 2006]. Hence, looking at the disease as "perturbation of networks" can provide such a framework which may offer insights from systems biology into the practicalities of personalized, preventive, predictive and participative (P4) medicine [Antonio del Sol et al., 2010]. Using this approach, a base is formulated for the appropriate selection of an inhibitor, taking into account the dynamics of the enzyme kinetics.

2.1 Lipophosphoglycan & GIPL structure

The parasite surface is the primary interface of Leishmania with the host, and is comprised largely of three abundant classes of glycosylphosphatidylinositol (GPI)-anchored molecules: lipophosphoglycan (LPG), a smaller heterogeneous group of glycoinositolphospholipids (GIPLs), and proteins such as gp63, gp46, PSA-2, and proteophosphoglycans (PPGs) [G.F. Spath, 2000]. This LPG is held together with a phosphoinositide membrane anchor, and has a tripartite structure consisting of a lipid domain, a neutral hexasaccharide, and a phosphorylated galactose-mannose, with a termination in a neutral cap [G.F. Spath, 2000] (Fig.3). LPG is in abundance in the cell membrane and is extended from the cell surface. In the mammalian host, LPG confers resistance to complement-mediated lysis, oxidative stress, and inhibits phagolysosomal fusion. In Leishmania major, the core of the abundant surface lipophosphoglycan (LPG) is structurally related to that of the smaller glycosylinositolphospholipids (GIPLs) in containing galactosylfuranose (Galf) residues in a Gal0–2Galf (β3)Man motif [Zhang et al., 2004]. However, deletion of the putative Galf - transferase (Galf T) LPG1 affected Galf incorporation in LPG but not GIPLs [G.F. Spath, 2000]. Thus, LPG and LPG1 satisfy the requirements for virulence factors genes as set forth by the modern ''Molecular Koch's postulates'', establishing LPG itself as a Leishmania virulence factor [Falkow S, 1988].
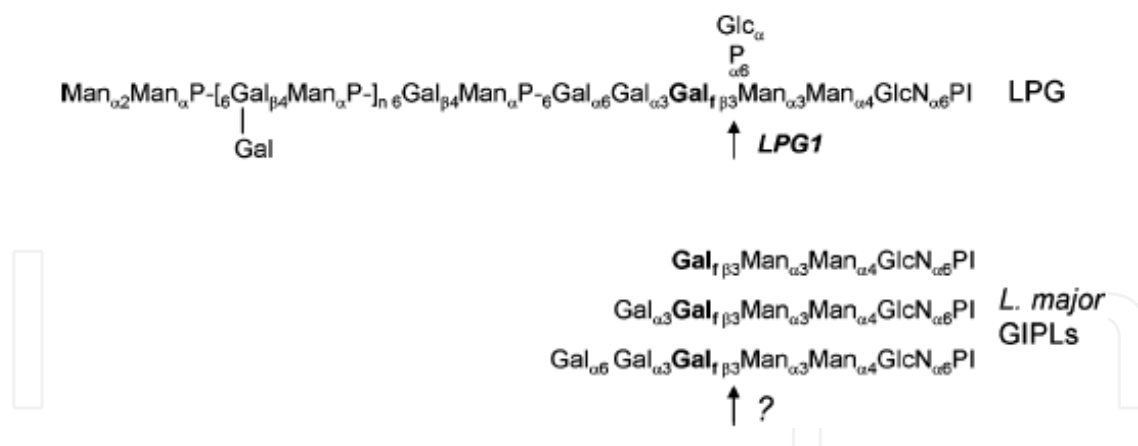
$$\text{Man}_{\alpha 2}\text{Man}_{\alpha}\text{P-}[_6\text{Gal}_{\beta 4}\text{Man}_{\alpha}\text{P-}]_{n\ 6}\text{Gal}_{\beta 4}\text{Man}_{\alpha}\text{P-}_6\text{Gal}_{\alpha 6}\text{Gal}_{\alpha 3}\textbf{Gal}_{f\,\beta 3}\text{Man}_{\alpha 3}\text{Man}_{\alpha 4}\text{GlcN}_{\alpha 6}\text{PI} \quad \text{LPG}$$

with Glc$_\alpha$ – P$_{\alpha 6}$ above, Gal below the Man$_\alpha$, and ↑ LPG1

$$\textbf{Gal}_{f\,\beta 3}\text{Man}_{\alpha 3}\text{Man}_{\alpha 4}\text{GlcN}_{\alpha 6}\text{PI}$$

$$\text{Gal}_{\alpha 3}\textbf{Gal}_{f\,\beta 3}\text{Man}_{\alpha 3}\text{Man}_{\alpha 4}\text{GlcN}_{\alpha 6}\text{PI}$$

$$\text{Gal}_{\alpha 6}\text{Gal}_{\alpha 3}\textbf{Gal}_{f\,\beta 3}\text{Man}_{\alpha 3}\text{Man}_{\alpha 4}\text{GlcN}_{\alpha 6}\text{PI}$$

L. major GIPLs

↑ ?

Fig. 3. Structure of GIPLs and LPGs in *L. major*.

2.2 Stochastic simulation

The study of metabolic networks is of high relevance, because of their implications for the basic understanding of living cells and organisms and for medical applications, especially with respect to discovering drug targets [Pavlopoulos et al., 2008]. Defining a system by the dynamic state of its metabolome relies on the effective integration of omics data because the metabolic state of the system is largely derived from the global expression of its genome and proteome [Davidov et al., 2003]. Recent advances enable the production of metabolic network models reconstructed from genome sequences, as well as experimental measurements of much of the metabolome [Douglas B. Kell, 2006]. The network-based representation and analysis of biological systems contributes to a greater understanding of their structures and functions at different levels of complexity. These techniques can also be used to identify potential novel therapeutic targets based on the characterisation of vulnerable or highly influential network components [Azuaje et al., 2010]. Recently developed computational tools enable rational and graphical composition of genetic circuits for standard parts, and subsequent simulation for testing the predicted functions *in silico* [Marchisio and Stelling, 2009]. Model construction typically involves the translation of prior knowledge into a list of reactants and reactions [Aldridge et al., 2006]. These models require information about the kinetics of each of the reactions participating in the network, such as the kinetic laws describing the dynamics of the reactions with their respective parameters. Thereafter, kinetic equations are mapped to the reaction model through either stochastic or deterministic approaches, and at times a hybrid model may be adopted. Several modelling approaches are derived from assuming different simplistic kinetic mechanisms: Convenience rate law, mass-action, Michaelis-Menten, power-law, LinLog [Sergio Grimbs, 2009]. There are several modes of modeling these metabolic networks: Structural modeling, Flux balance Analysis and Metabolic control Analysis. Thus, using the concept of mathematical modeling and applying the science of systems biology, the simulation of metabolic pathway is carried out based on certain kinetic laws and their parameters.

Once a target is identified, whole cell screening assays can be designed, using synthetic biology strategies [Weber and Fusseneger, 2009]. Temporal information from computer simulations are validated by those done on simplified versions of yeast cell cycle and have shown to be in good agreement with experiment [Tyson et al., 2001]. The conversion of a reconstruction into a mathematical format facilitates a myriad of computational biological

studies, including evaluation of network content, hypothesis testing and generation, analysis of phenotypic characteristics and metabolic engineering [Thiele and Bernhard , 2010].

2.3 Enzymes unique to *L.major* and absent in *H.sapiens*

It was noted that these distinctive enzymes listed below restricted only to *L.major* catalyzed reactions are involved in some of the metabolic pathways (Table 1).

Sr. No.	Enzymes	Metabolic Pathways
1	side-chain/ phosphoglycan β-1,3 galactosyltransferase 7	LPG,GIPL,GPI Biosynthetic Pathway
2	side-chain/ phosphoglycan β-1,3 galactosyltransferase 6	
3	side-chain/ phosphoglycan β-1,3 galactosyltransferase 5	
4	side-chain/ phosphoglycan β-1,3 galactosyltransferase 4	
5	side-chain/ phosphoglycan β-1,3 galactosyltransferase 3	
6	side-chain/ phosphoglycan β-1,3 galactosyltransferase 2	
7	side-chain/ phosphoglycan β-1,3 galactosyltransferase 1	
8	side-chain/ phosphoglycan β-1,3 galactosyltransferase related protein 6	
9	side-chain/ phosphoglycan β-1,3 galactosyltransferase related protein 5	
10	side-chain/ phosphoglycan β-1,3 galactosyltransferase related protein 4	
11	side-chain/ phosphoglycan β-1,3 galactosyltransferase related protein 3	
12	side-chain/ phosphoglycan β-1,3 galactosyltransferase related protein 2	
13	side-chain/ phosphoglycan β-1,3 galactosyltransferase related protein 1	
14	phosphoglycan β-1,2 arabinosyltransferase	
15	side- chain β-1,2 arabinosyltransferase	
16	mannosyltransferase (GPI14)	
17	GIPL galf transferase	
18	β- galactofuranosyltransferase like protein (LPG1)	
19	galactofuranosyltransferase like protein lpg1 - like protein	
20	β- galactofuranosyltransferase	
21	GPI 18/ PIG V	
22	GPI 10/ GPI anchor biosynthetic protein	
23	gpi transamidase complex	
24	Sterol C-24 reductase,putative	Ergosterol Biosynthesis
25	C-5 sterol desaturase,putative	
26	Sterol 24-C methyltransferase	
27	Δ6 desaturase	Fatty Acid Elongation
28	Δ4 desaturase	
29	Δ6 elongase	
30	Δ5 elongase	
31	Δ5 desaturase	
32	Δ12 desaturase	
33	stearic acid desaturase	
34	enoyl-[-ACP] reductase, putative	
35	fatty acid elongase	Fatty Acid Biosynthesis
36	β- ketoacyl ACP synthase	
37	cardiolipin synthase, putative	Ester Biosynthesis

Table 1. Enzymes restricted to *L.major* and the respective Metabolic Pathways belonging to Lipid and Fatty Acid Biochemical Network.

Thus, of the overall 20 metabolic pathways involved in Lipid and Fatty Acid Metabolism, there are 13 such pathways which are exclusively restricted to *L.major*, of which only 7 pathways had the enzymes and reactions unique to *L.major*. From the Table 1 it is found that the maximum number of the genes and the resultant gene products (enzymes), distinctive to the *L.major* are specifically engaged in the LPG and GIPL Biosynthesis Pathway followed by the fatty acid metabolism.

2.4 Gene Regulatory Network (GRN) and lipid metabolism

The prior mentioned metabolic pathways (Fig.4) are taken into account to form an interacting regulatory network of genes in order to lay an emphasis upon the clustering of the genes into metabolic pathways of related functions.
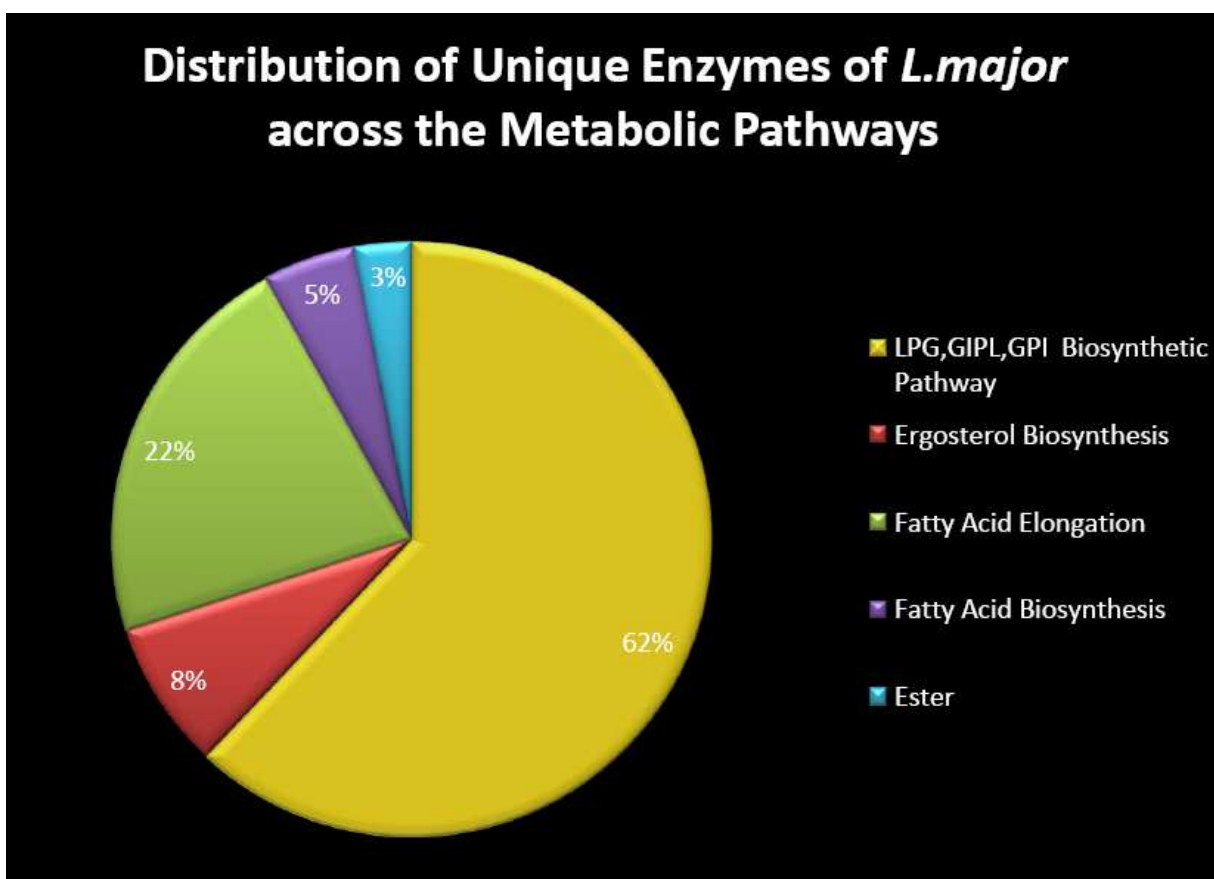
Fig. 4. Pie chart showing the abundance of unique enzymes present in LPG, GPI anchor and GIPL biosynthetic pathway.

A gene regulatory network (GRN) aims to capture the dependencies between these molecular entities and is often modeled as a network composed of nodes which represents genes, proteins and/or metabolites) and edges representing molecular interactions such as protein–DNA and protein–protein interactions or rather indirect relationships between genes. Network model architectures can be distinguished by the type of model: stochastic or deterministic, static or dynamic and the type of relationships between the variables (directed or undirected; linear or non-linear function or relation table). Although many undirected network representations exist, the focus of this chapter is on directed networks.

From this network it is clearly visible, that LPG Biosynthetic Pathway, GIPL Biosynthesis Pathway, GPI anchor Biosynthesis Pathway, and Dolichyl-diphosphooligosaccharide biosynthesis Pathway are inter-connected through one or more genes involved commonly in the regulation of the proteins which act as enzymes catalyzing the reactions in these pathways (Fig.5).
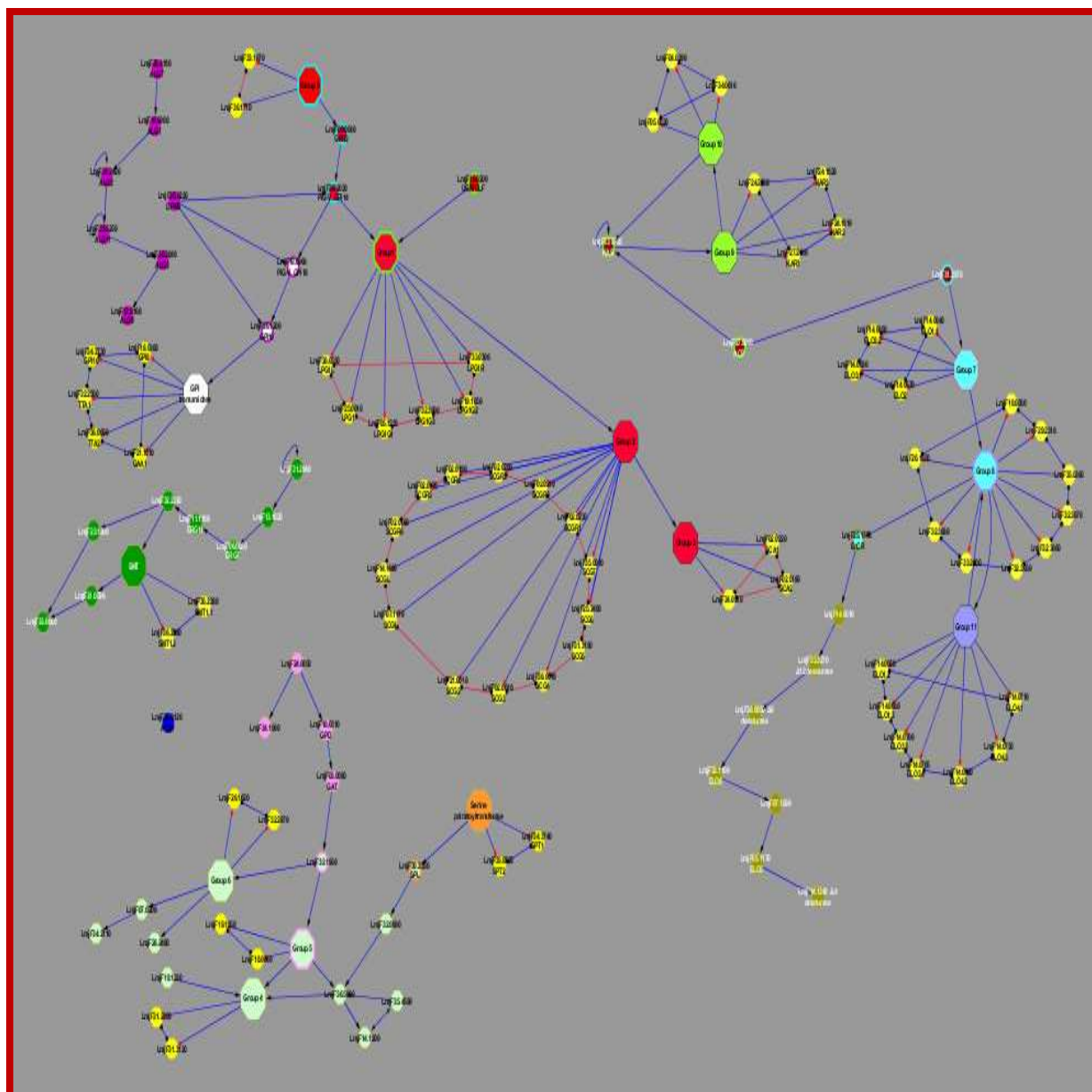

Fig. 5. Gene Regulatory Network constructed using Cytoscape for visualizing the inter-connection of the Lipid Metabolic Pathways restricted to *L.major *.

*Legend overleaf

LEGEND:
- Lipophosphoglycan Pathway
- LPG + GIPL Biosynthesis Pathway
- LPG + GIPL + GPI- anchor Biosynthesis Pathway

○ GPI- anchor Biosynthesis Pathway
◉ GPI - anchor + GIPL Biosynthesis Pathway
● Dolichyl-diphosphooligosaccharide Biosynthesis
● Ergosterol Biosynthesis
● Sphingolipid Metabolism
○ Ester Phospholipid Biosynthesis
○ Sphingolipid Metabolism + Ester Phospholipid Biosynthesis
● Glycerolipid biosynthesis- initial steps
○ Ester Phospholipid Biosynthesis + Glycerolipid biosynthesis- initial st
● Ether Phospholipid Biosynthesis
● Fatty acid Biosynthesis Initiation
○ Fatty acid biosynthesis-- elongase pathway
● Very Long Chain Fatty Acid Biosynthesis
● Fatty Acid Elongation − Unsaturated
● Fatty Acid Elongation - Saturated
◉ F.A. elongase + F.A. initiation
◉ F.A. elongation saturated + F.A. initiation
◉ F.A. elongase + V.L.C. F.A. Biosynthesis
◉ F.A. elongase + F.A. elongation saturated
◉ Group of genes involved in single reaction

Similarly, the pathways Fatty Acid Elongation (Saturated and Unsaturated), Fatty acid biosynthesis Initiation, Very Light Chain Fatty Acid Biosynthesis and Fatty acid biosynthesis-elongase pathway; and the pathways Ester phospholipid biosynthesis, Sphingolipid Metabolism and Glycerolipid biosynthesis- initial steps are respectively inter-related through common intervention of the genes. But, the pathways Ergosterol Biosynthesis and Ether Phospholipid Biosynthesis are the only ones that are not related with any of the other Metabolic Reactions, and are self-contained.

Hence, the gene regulatory network gives an entire overview about the reacting species in the same dimension. The reactions and pathways that are inter-connected and have common enzymes and compounds taking part in the same reactions can further be investigated. Also, it gives an outline and a general idea related to the cluster of genes taking part in the reactions.

From the Cytoscape Regulatory Network of the Lipid Metabolism Pathway, it is found that the maximum number of the genes and the resultant gene products (enzymes), distinctive to the *L.major*, ae specifically clustered in the LPG and GIPL Biosynthesis Pathway, followed by the fatty acid metabolism.

Thereby, from the Cytoscape network and Fig 6, it is established that majority of the enzymes in a large number are present only in the LPG, GIPL and GPI anchor bioynthesis pathways, thus giving a wide scope for the enzymatic studies to be undertaken. Consequently, the Lipophosphoglycan Pathway in conjunction with the GIPL and GPI anchor biosynthesis is of paramount importance.

Fig. 6. Enlarged portion of the Gene Regulatory Network focusing on the large number of genes involved in LPG, GIPl and GPI anchor biosynthesis pathway [Pathways marked by red octagon]. Thus it can be concluded that the gene GPI 14 [marked by off-white] can play a crucial role in the pathway since it is at a junction leading to the other 2 pathways- GIPL and GPI. Also, the cluster of genes assembled into Group1 [marked by pink] is at a vital cross-road leading to GIPL and LPG Biosynthetic Pathway.

2.5 Compartmentalization

The sub-cellular localization of the target enzymes plays a very significant role in order to introduce an inhibitor in a cellular system. Hence, the location of the enzymes involved in LPG Biosynthetic Pathway is determined and the presence of the transmembrane domains is also checked (Table 2).

Cellular overview was created using this data to simplify the complexity and for elaborative interpretation of the data. The compartments included are Nucleus, Golgi apparatus, Endoplasmic Reticulum, and the Cell Membrane. The fluxes and the choke-point reactions are also marked to indicate the critical reactions in the pathway (Fig.7).

Sr. No.	Enzymes	TMHMM domains	Sub-cellular Localization	Signal Peptide	Signal Anchor
1	phosphatidylinositol-N-acetylglucos-aminyltransferase subunit c, putative	7	integral to Cell membrane	-	-
2	n-acetylglucosaminyl-phosphatidyl inositol biosynthetic protein,putative	-	cytoplasm	-	-
3	N-acetyl-D-glucosaminyl-phosphatidylinositol deacetylase	1	integral to cell membrane	Present	-
4	*mannosyl transferase,putative*	8	integral to ER	-	-

Sr. No.	Enzymes	TMHMM domains	Sub-cellular Localization	Signal Peptide	Signal Anchor
			membrane		
5	*beta galactofuranosyl transferase*	1	integral to cell membrane	-	Present
6	*GIPL galf transferase, putative*	1	integral to cell membrane	Present	-
7	*beta-galactofuranosyltransferase-like protein*	1	integral to cell membrane	-	Present
8	galactofuranosyltransferase lpg1-like protein	1	integral to cell membrane	-	Present
9	*GIPL galf transferase, putative*	1	integral to cell membrane	-	Present
10	*GIPL galf transferase, putative*	-	cytoplasm	-	-
11	*phosphoglycan beta 1,3 galactosyltransferase 4, scβGalT*	1	Golgi Membrane	Present	-
12	*phosphoglycan beta 1,3 galactosyltransferase 3, scβGalT*	1	Golgi Membrane	-	-
13	*phosphoglycan beta 1,3 galactosyltransferase 2, scβGalT*	1	Golgi Membrane	-	-
14	*phosphoglycan beta 1,3 galactosyltransferase 1, scβGalT*	1	Golgi Membrane	-	-
15	*side chain β1,3 galactosyltransferase-like protein*	1	Golgi Membrane	-	-
16	*side chain β1,3 galactosyltransferase-related protein 6*	1	Golgi Membrane	-	-
17	*side chain β1,3 galactosyltransferase-related protein 5*	1	Golgi Membrane	-	-
18	*side chain β1,3 galactosyltransferase-related protein 4*	1	Golgi Membrane	-	-
19	*side chain β1,3 galactosyltransferase-related protein 3*	1	Golgi Membrane	-	Present
20	*side chain β1,3 galactosyltransferase-related protein 2*	1	Golgi Membrane	-	-
21	*side chain β1,3 galactosyltransferase-related protein 1*	1	Golgi Membrane	-	-
22	*phosphoglycan beta 1,3 galactosyltransferase 7, scβGalT*	1	Golgi Membrane	-	-
23	*phosphoglycan beta 1,3 galactosyltransferase 6, scβGalT*	1	Golgi Membrane	-	-
24	*phosphoglycan beta 1,3 galactosyltransferase 5, scβGalT*	1	Golgi Membrane	-	-
25	*phosphoglycan beta 1,2 arabinosyltransferase, scβAraT-2*	1	Golgi Lumen	-	-
26	*phosphoglycan beta 1,2 arabinosyltransferase, scβAraT*	1	Golgi Lumen	-	-
27	*phosphoglycan beta 1,2 arabinosyltransferase*	1	Golgi Lumen	-	-

Sub-cellular Localization of enzymes involved in the LPG Metabolic Pathway

Table 2. Sub-cellular Localization of the Enzymes, through Literature Survey, GeneDB and UniProtKB. The enzyme names marked in bold and italicized are the ones that are present exclusively in the *L.major* pathogen.
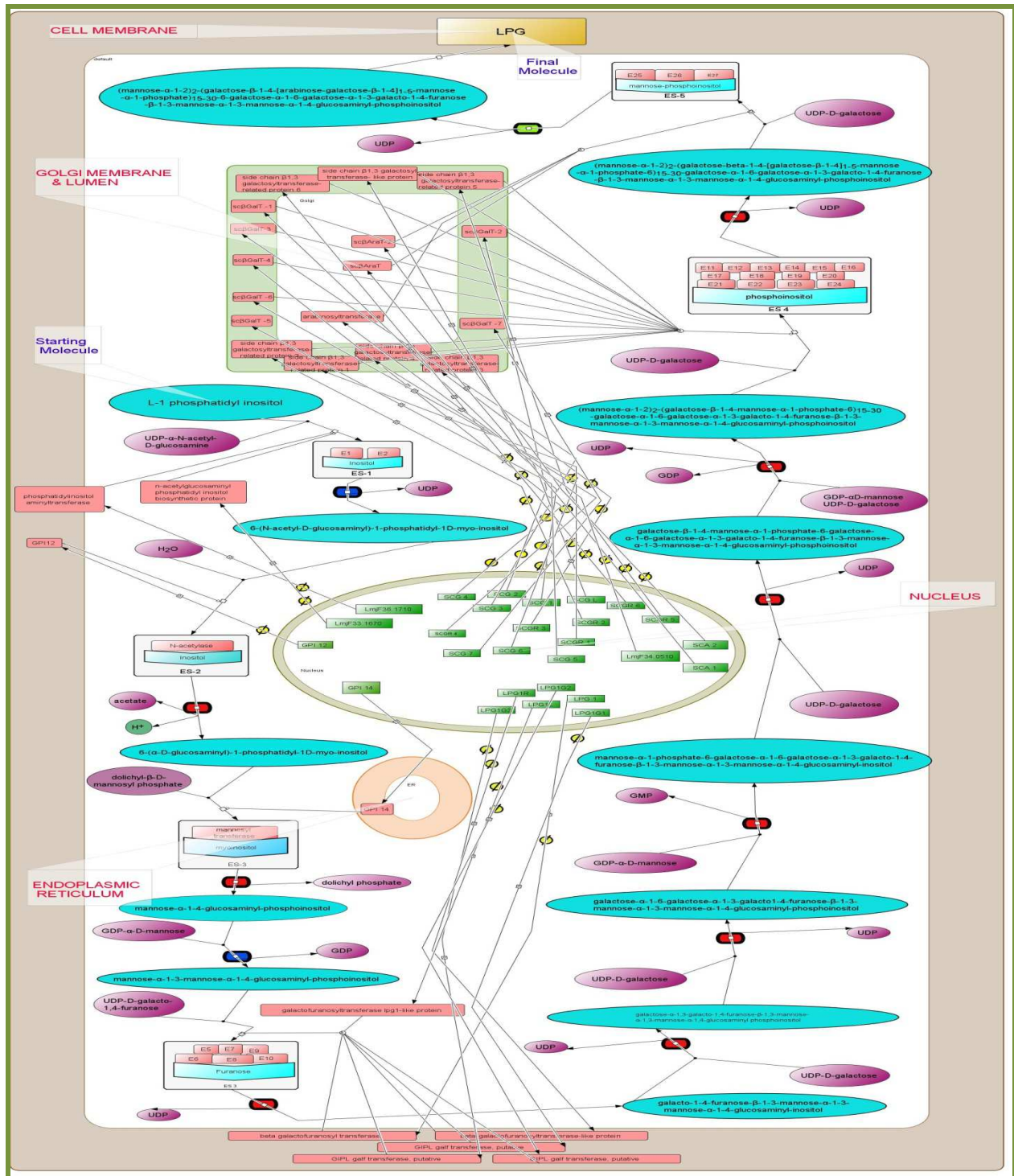
Fig. 7. Cellular Overview of the LPG Biosynthetic Pathway with respect to the Sub-cellular Localization, constructed using Cell Designer. The enzymes, for which the literature data was also not available, were considered to be within the cytosol.*

* Legend overleaf

LEGEND:

Reactants & Products

Metabolites

Enzymes

Genes

Receptor

Enzymes within Complex

Enzyme-Substrate Complex

Flux

Choke Point Reactions

Choke Point Reactions- On consumption side

Choke Point Reactions- On production side

2.6 Biochemical modeling

Since it's established that the **Lipophosphoglycan Biosynthetic Pathway** posed a vital role in the Leishmania Parasite, and is a significant Pathway for identifying the drug targets, it is reconstructed in Cell Designer for generating a Mathematical Model of the same (Fig.8).

The genes coding for the enzymatic proteins, substrates, products, Enzyme-substrate complex, and the metabolites – all the entities are taken into account during the rebuilding of the pathway.

From this reconstruction, an outline of the change in the fluxes through the pathway is created. The fact file of all the contributing moieties is described in Table 3 and Table 4.

LPG Biosynthesis Reaction Summary		
Sr. No.	Species	Figures
1	Enzymes	27
2	Genes	27
3	Metabolites	27
	Consumed	13
	Produced	14
4	Compounds (Reactants & Products)	14
5	Enzyme-Substrate Complexes	6
6	Total no of Reactions	13
	Enzyme Catalyzed Reactions	6
	Non- Enzyme Catalyzed Reactions	7

Table 3. Summary of the LPG Biosynthesis Reaction.
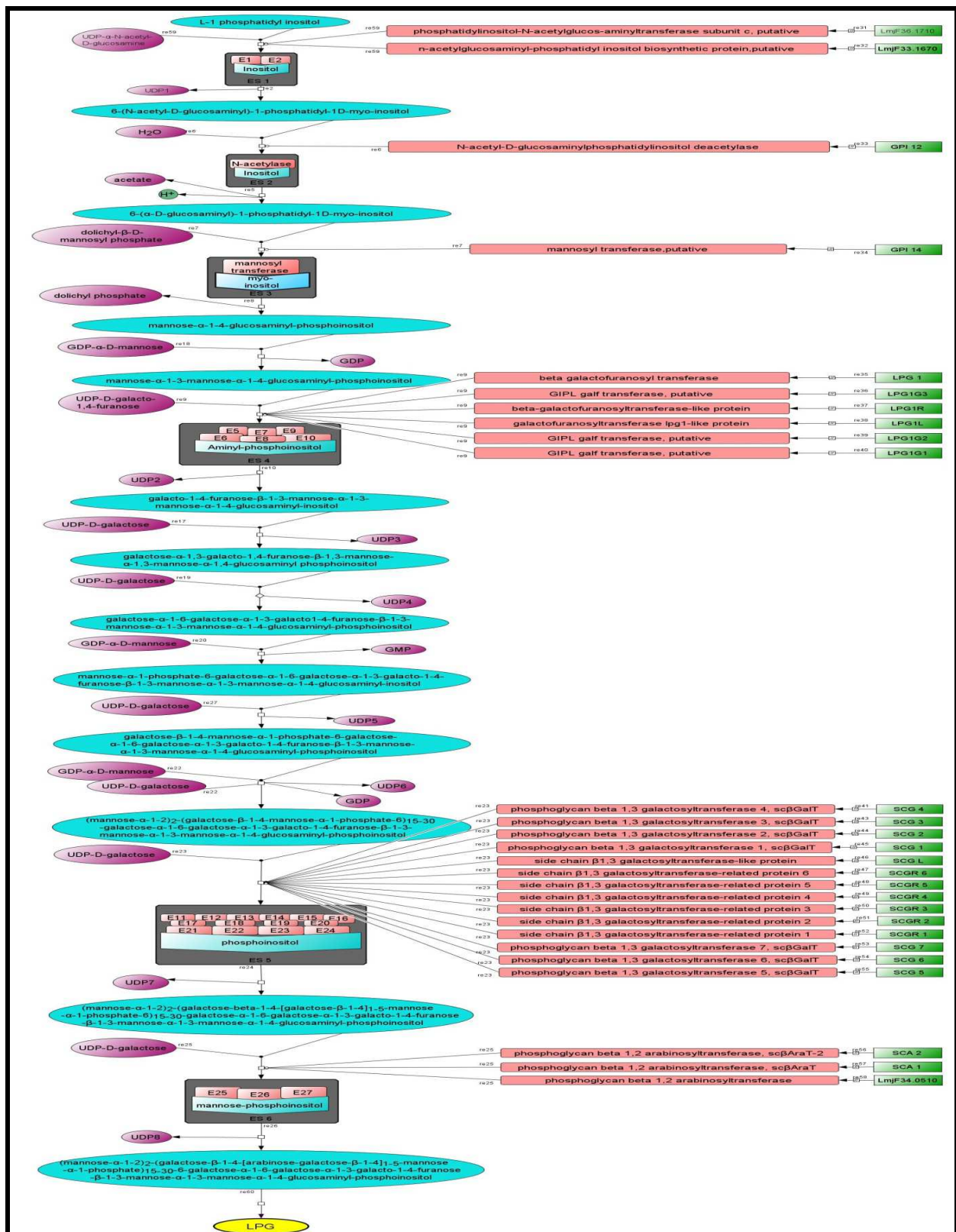
Fig. 8. Mathematical modeling of LPG Biosynthetic Pathway using Cell Designer.

Breakdown of Reactions		
1	Total no of Irreversible Reactions	67
	Enzyme Catalyzed Reactions	54
	Non- Enzyme Catalyzed Reactions	13
2	Transport Reactions	51
	Gene Reactions	27
	Non-gene reactions	24
3	Non- Transport Reactions	16

Table 4. Analysis of the individual reactions occurring in the LPG Biosynthesis Pathway.

Translating a known metabolic network into a dynamic model requires rate laws for all chemical reactions. The mathematical expressions depend on the underlying enzymatic mechanism; they can become quite involved and may contain a large number of parameters. Rate laws and enzyme parameters are still unknown for most enzymes. Convenience kinetics is used to translate a biochemical network – manually into a dynamical model with plausible biological properties. It implements enzyme saturation and regulation by activators and inhibitors, covers all possible reaction stoichiometries, and can be specified by a small number of parameters. Its mathematical form makes it especially suitable for parameter estimation and optimization. In general, the convenience kinetics applies to arbitrary reaction stoichiometries and captures biologically relevant behavior such as saturation, activation, inhibition with a small number of free parameters. It represents a simple molecular reaction mechanism in which substrates bind rapidly and in random order to the enzyme.

2.6.1 Convenience mass action kinetics

For the enzyme catalyzed reaction, i.e. from substrate to the ES-complex formation, the convenience rate law is taken into account. Rate laws and enzyme parameters are still unknown for most enzymes. For reactions with two or more catalysts, one individual rate law is generated for each catalyst. The total rate law for this particular reaction is given as the sum of the individual rates of all participating catalysts. The equation used here for convenience kinetics is:

$$\frac{E}{defult} \times \frac{kcat\_re \times \dfrac{\dfrac{R}{default}}{kcm\_re\_R} \times \dfrac{\dfrac{M}{default}}{kmc\_re\_M}}{1 + \dfrac{\dfrac{R}{default}}{kcm\_re\_R} \times 1 + \dfrac{\dfrac{M}{default}}{kmc\_re\_M}}$$

where, E is the Enzyme concentration, kcat_re is the catalysis constant, R is the Reactant concentration, M is the Metabolite constant, kmc_re_R and kmc_re_M are the reactant and metabolite Michelis constants respectively, default indicates the default initial quantity value defined by us. For reactions with more than 1 enzyme (catalysts), the equation is summed up for each value of the enzymes.

In general, the convenience kinetics applies to arbitrary reaction stoichiometries and captures biologically relevant behavior such as saturation, activation, inhibition with a small number of free parameters. It represents a simple molecular reaction mechanism in which substrates bind rapidly and in random order to the enzyme.

2.6.2 Generalized mass action

For non-enzyme catalyzed reaction, i.e. from the ES to the product formation, generalised mass-action kinetics is used. Mass action kinetics has numerous analytical properties that are of inherent interest from a dynamical systems perspective. These kinetic laws fail to describe enzyme saturation at high substrate concentrations, which is a common and relevant phenomenon.

$$kass\_re \times \frac{ES - Co,plex}{default}$$

where, kass_re is the association constant, ES-complex is the concentration of the ES-complex, default is the initial quantity value defined by us.

2.6.3 Hill-Hinze equation

SBMLsqueezer offers the general Hill equation as a kinetic law for the formation of gene products:

$$V_{re2} = vmax\_re2 \cdot \frac{[R1]^{hic\_re2\_s2}}{[R1]^{hic\_re2\_2} + ksp\_re2\_s2^{hic\_re2\_s2}}$$

Since it is assumed that the transition from a gene to protein as "a known transition omitted", and ignored the RNA, the Hill equation is modified by the SBML Squeezer as per the reaction"s convenience and the equation is given as: Vmax_re where, vmax is the maximum velocity attained. Now, according to Quassi Steady State rule, d/dt [mRNA] =0 , i.e. for the intermediate compound, and moreover the concentration of the mRNA reaches steady state very quickly, compared to protein concentration, it is considered that [mRNA] concentration is always considered at steady state.

2.7 Numerical simulations

On the basis of these kinetic laws, the numerical simulations are carried out with certain default values of the parameters and the initial quantities. By means of the numerical simulations, certain graphs are obtained which display the interactions of the molecules with each other and the effect of minute variations in the kinetic parameters. The graphs

also generate visualization for the fluxes. For each of the enzyme catalyzed reactions the simulation process is carried out separately for reducing the complications in interpretation of the graphs generated.

In each of the graphs, it is postulated that there was a constant decrease in the concentration of the substrate, while the product had a steady increase in the concentration, with the ES complex initially forming up but then gradually decreasing with time when the product formation began each time. Therefore, on an overall basis, from the above generated graphs, it is concluded that as and when the downstream reaction is considered, there is a potential fall in the flux, though not in the reaction 1, indicating that the flux increased from reaction 1 to reaction 2. Moreover, there is a consistent drop in the concentration of the ES complex as well, which signified that the product is gradually being formed. The concentration of gene G1 almost remain unchanged during this process as the transcription does not change the state of the gene. The principal result of this work is laid onto the idea that it shows a simply identifiable class of kinetic expressions, including the familiar detailed balanced kinetics as a proper subclass, which ensures consistency with the extended thermodynamic conditions. The rate laws applied provided a very good fit to the data. With the network model it is identified which steps have strong effect on other parts of the network using sensitivity analysis, and those steps that do indeed have high levels of control are then chosen further for kinetic analysis. This helped to further postulate that GPI 14 (encoding for mannosyltransferase) and Galf (encoding for Galf transferase) could be examined further since they showed a strong effect on the rest of the network model (Fig.9 A,B,C,D,E and F). The most important feature for building a bottom-up biochemical network model is that the relation between the concentration of the effectors and the rate be accurate. The objective is to learn more about how cells work via the means of computational models and not all about the mechanisms of catalysis, except when they are found themselves of importance to cellular function. The concept of "gene function" becomes synonym with the kinetics of its protein product embedded in the cellular biochemical network.

The network-based representation and analysis of biological systems contributes to a greater understanding of their structures and functions at different levels of complexity. By defining the properties of a system as a whole, the approach aids in improvising decision making in the therapeutic drug development, which has a limitation otherwise, to be predicted from the parts. The goal of modeling in systems biology is to provide a background for hypothesis formulation and prediction based on *in silico* simulation of parasite biology across the multiple distance and time scales. The metabolic reconstruction of biochemical pathway serves as a tool for clarifying inconsistencies between data sources, and provides a platform for data integration and hypothesis generation for further scientific research in infectious disease. By means of the Cell Designer, utilizing its SBML language platform, we articulate a succinct representation of the biochemical properties adopted by a biological system and the underlying metabolic networks, and consequently translate them into systems compatible to stochastic simulations. Ensuing such model construction, and subsequent simulations we introduce a time component in the interaction reactions to determine certain physiochemical conditions of the biological system. Simulations aid in transforming the static data into the dynamic sphere of physiological timescales, by revealing insights into the flexibility of energy metabolism. These simulated models of
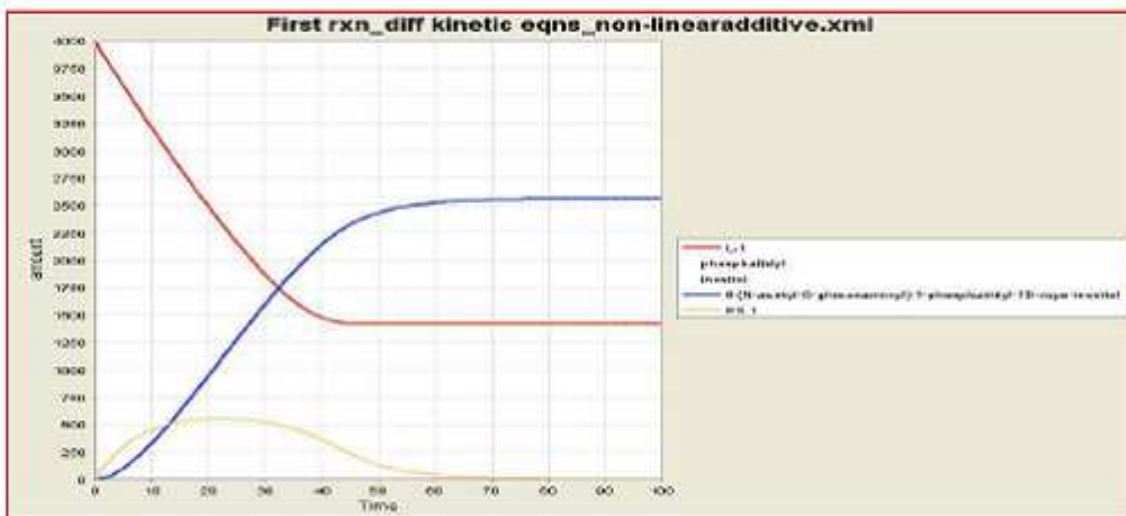
A

<u>First Reaction :</u>

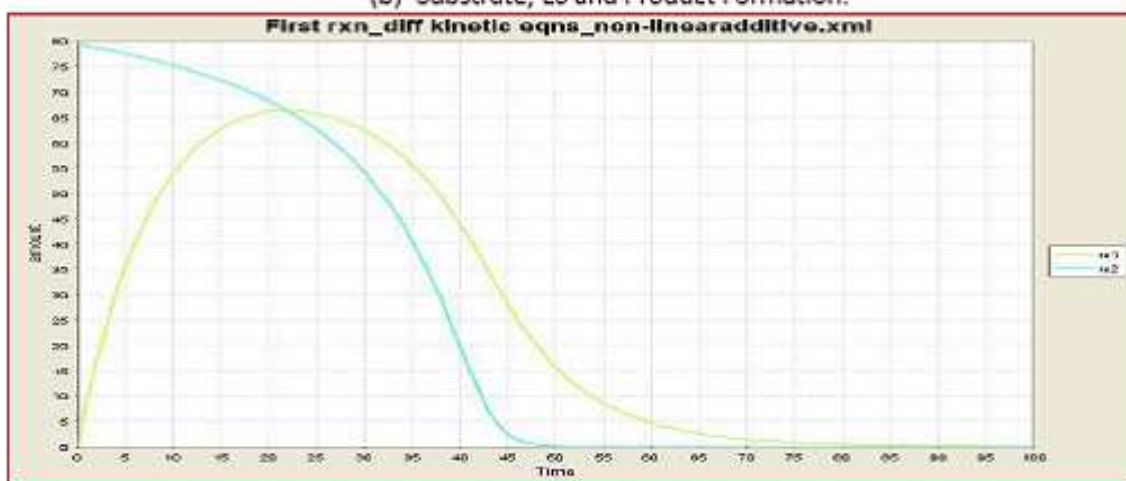### Simulation graphs of Reaction 1
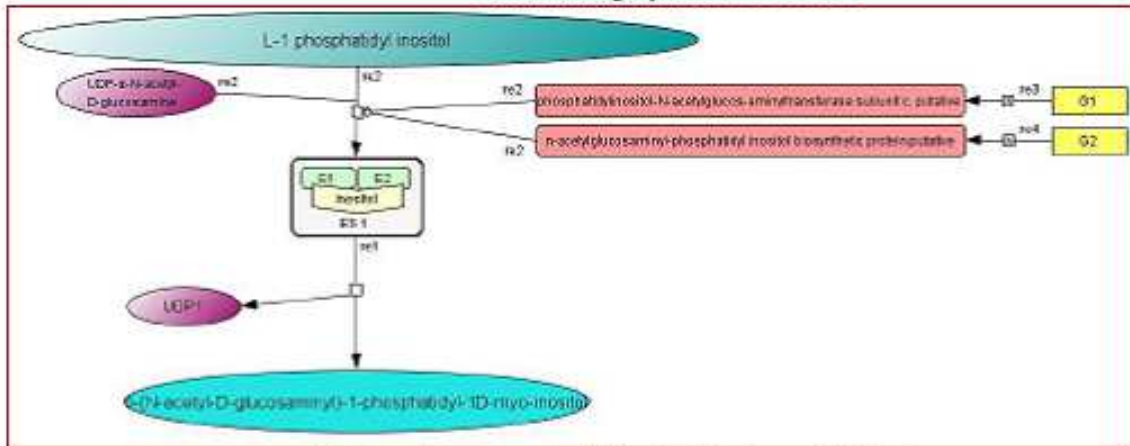
(a) Reaction represented by Cell Designer indiviually

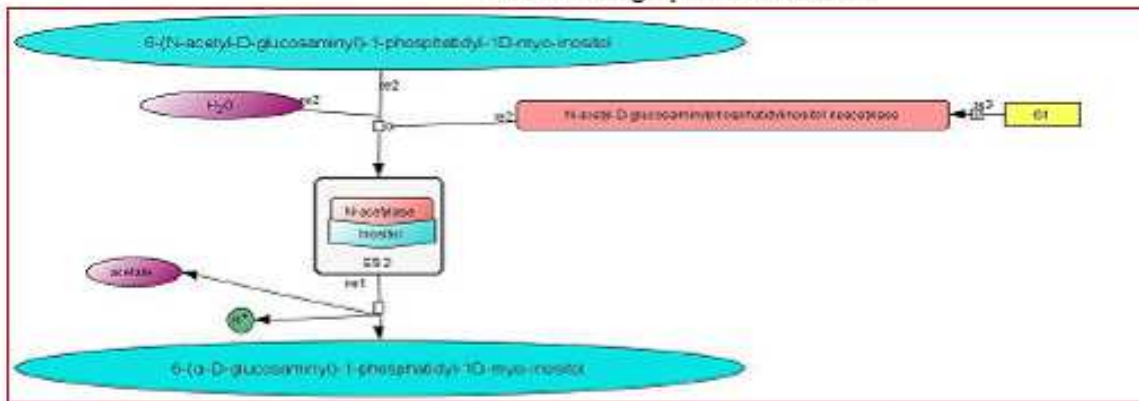
(b) Substrate, ES and Product Formation.
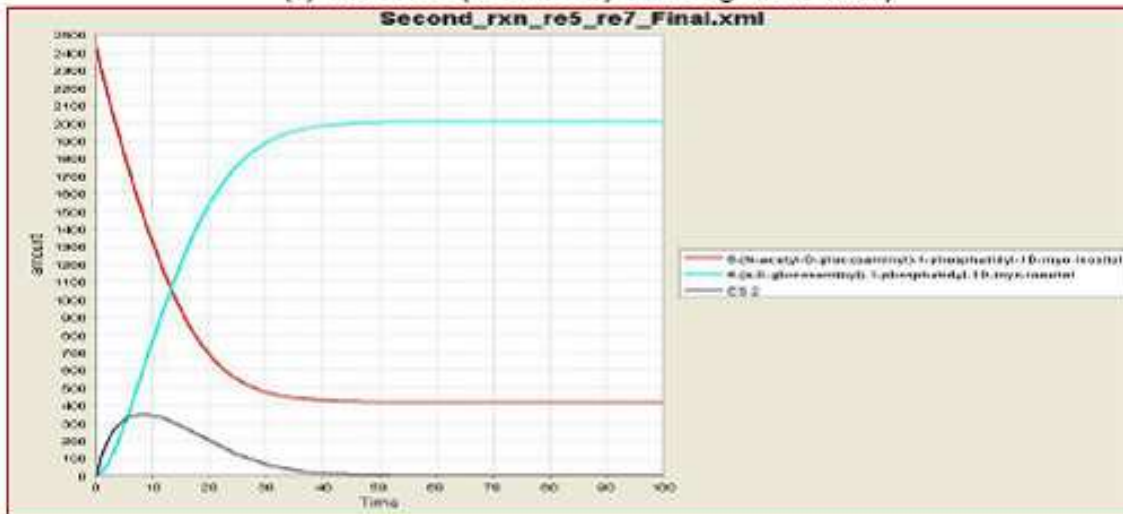

(c) Fluxes of the sub- reaction 1 and sub- reaction 2.
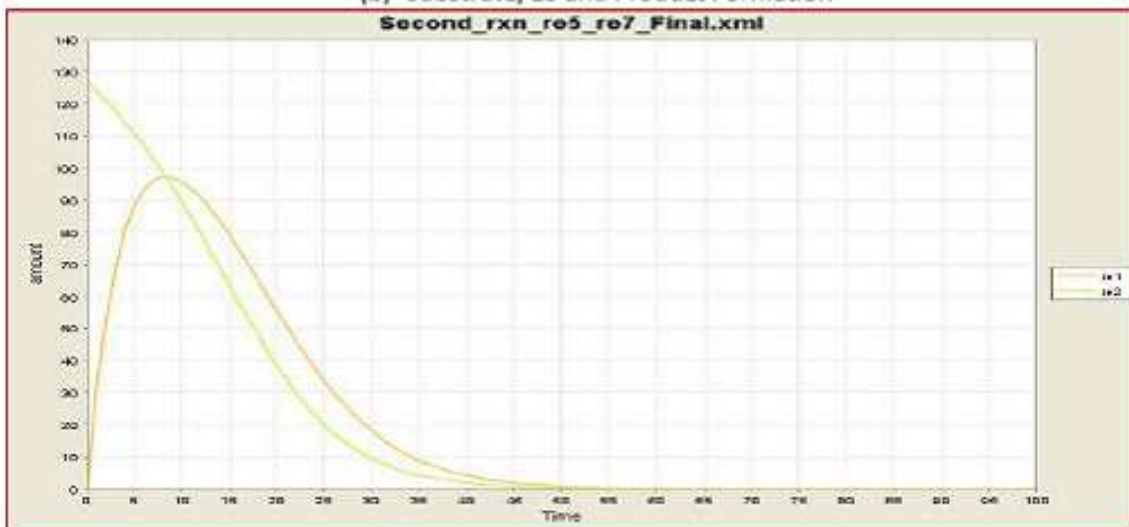
B

<u>Second Reaction:</u>

Simulation graphs of Reaction 2

(a) Reaction represented by Cell Designer indiviually
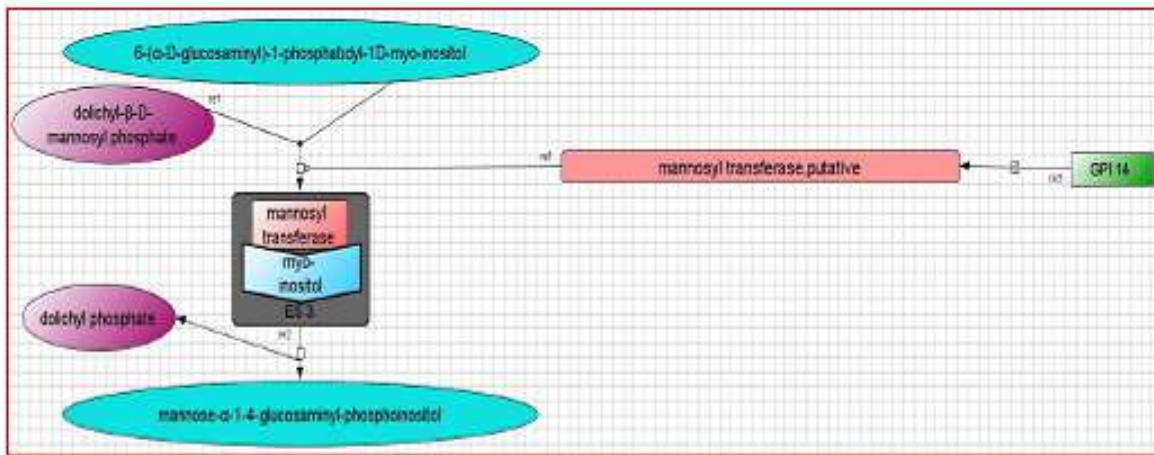

(b) Substrate, ES and Product Formation

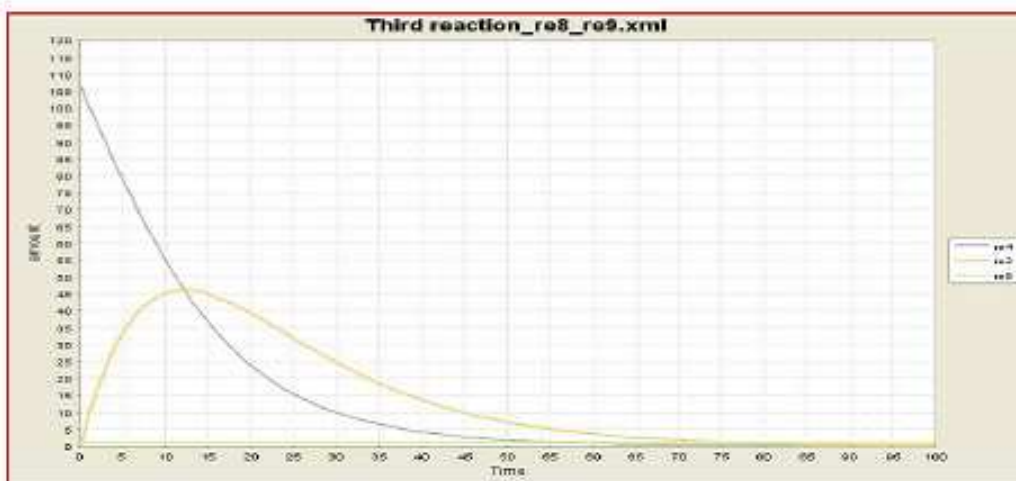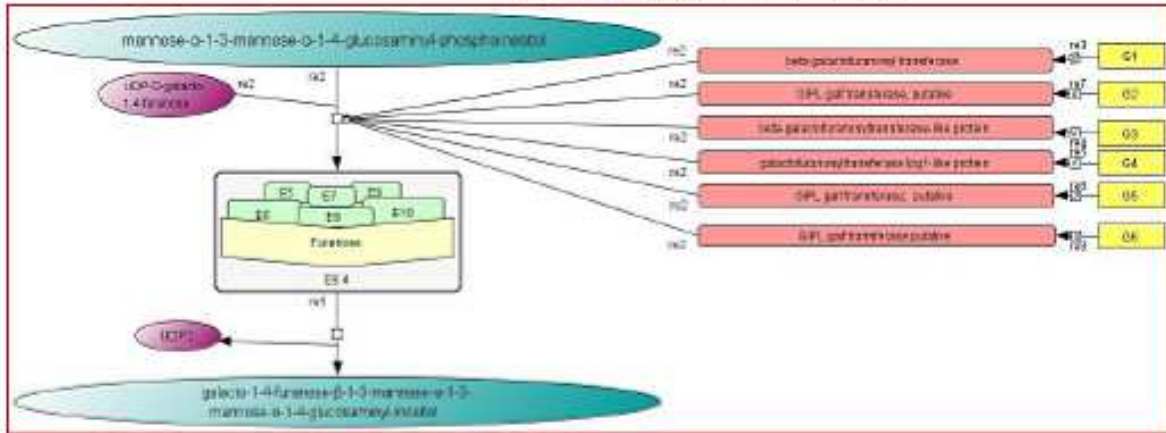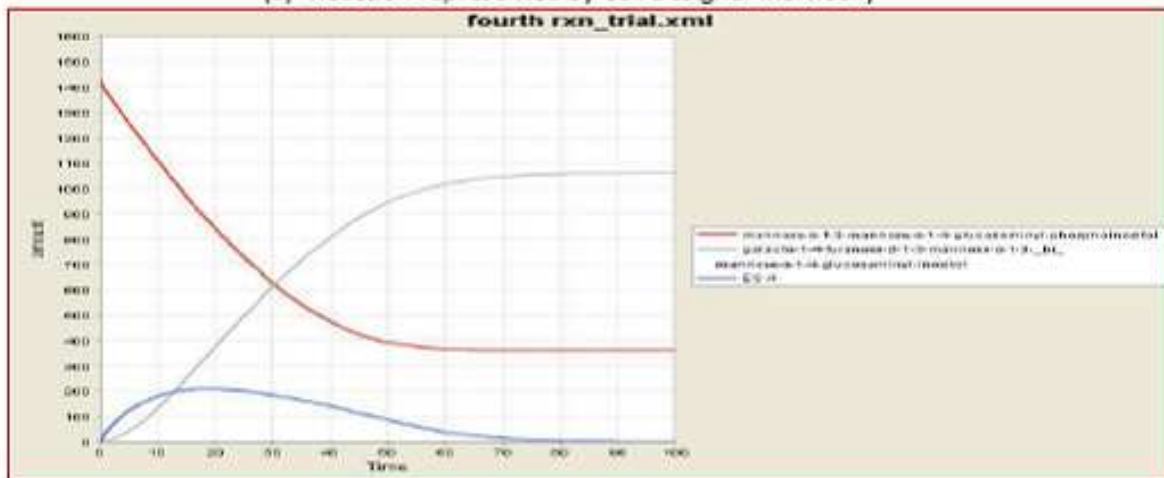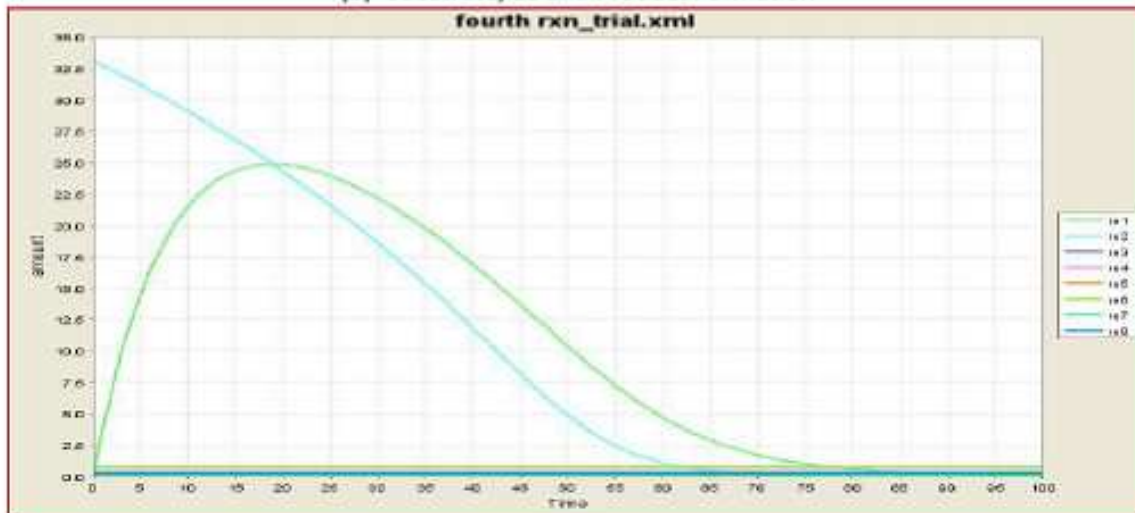(c) Fluxes of reaction sub-1 and sub-reaction 2.

C

__Third Reaction:__

Simulation graphs of Reaction 3

(a) Reaction represented by Cell Designer indiviually

(b) Substrate, ES and Product Formation : the Graph shows a steady formation of the product

(c) Fluxes of sub- reaction 1,sub-reaction 2 and sub-reaction 3

D

 Fourth reaction:

(a) Reaction represented by Cell Designer indiviually

(b) Substrate, ES and Product Formation
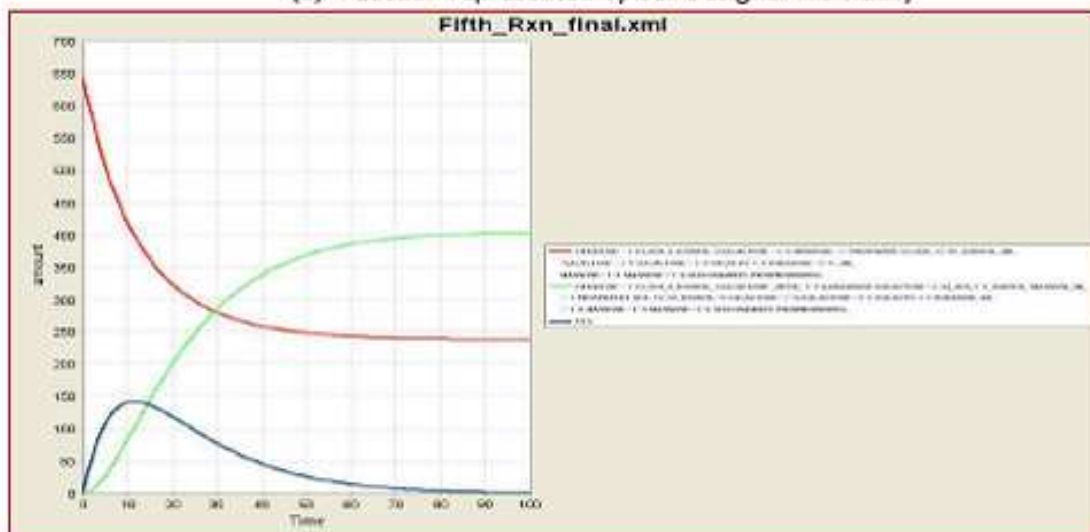

(c) Fluxes of sub- reaction 1,sub-reaction 2, sub-reaction 3, sub- reaction 4,sub-reaction 5, sub- reaction 6,sub-reaction 7, and sub- reaction 8.
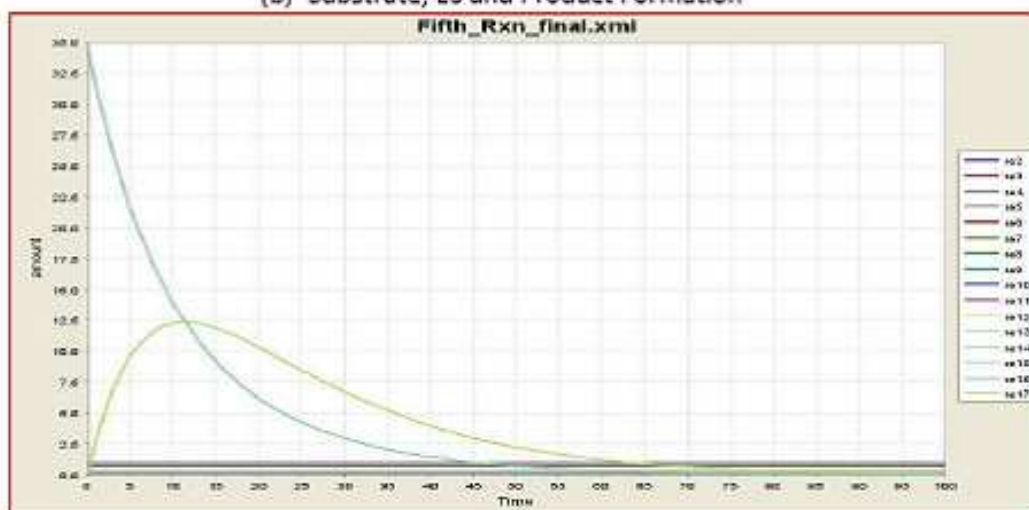
E

Fifth Reaction:
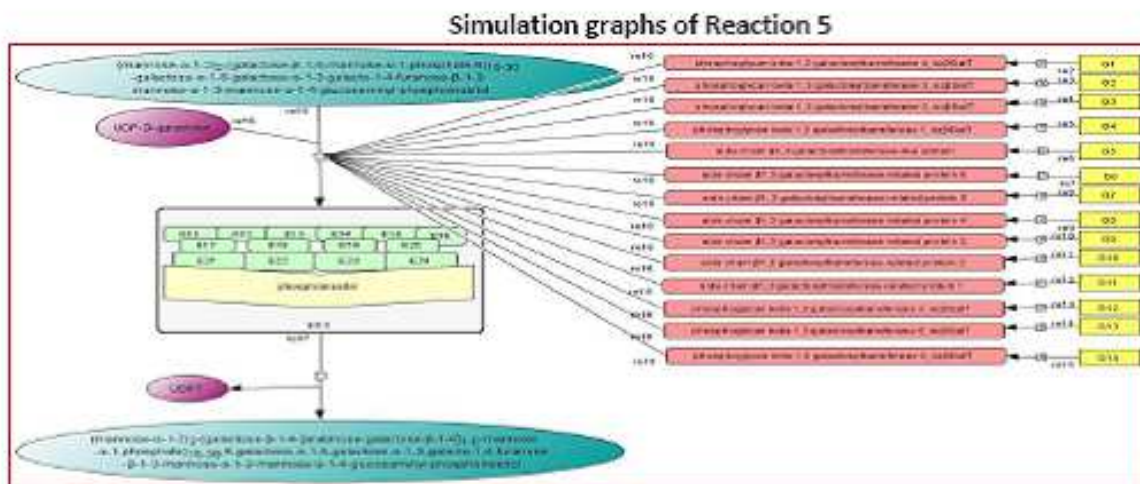
Simulation graphs of Reaction 5

(a) Reaction represented by Cell Designer indiviually

(b) Substrate, ES and Product Formation

(c) Fluxes of sub- reaction 1 to sub-reaction -17

F

<u>Sixth Reaction:</u>

(a)   Reaction represented by Cell Designer indiviually

(b) Substrate, ES and Product Formation

(c) Fluxes of sub-reaction 2, sub-reaction 3, sub- reaction 4,sub-reaction 5 and sub- reaction 6.

Fig. 9. Fluxes across six set of reactions in LPG biosynthetic pathway of *L. major* predicted through discrete stochastic simulation event.

biochemical reaction pathways are quite enthralling since their mathematics provides a means to amalgamate prior knowledge with postulated data and the underlying physical principles, through the graphical visualizations. Thus, we say graphical approaches simplify the process of transforming a network diagram into a set of linked equations. [Aldridge et al., 2006] Thereby, the prediction of potential novel targets can aid in the identification of revolutionary approach for controlling the level of specific, therapeutically relevant genes or proteins. We can test the efficacy of drug targets in the presence of enzyme inhibitors, through the metabolic reconstructions. These techniques can also be used to identify potential novel therapeutic targets based on the characterization of vulnerable or highly influential network components. [Azuaje et al., 2010] To summarize in a nutshell, the *in vivo* and *in silico* understanding of the "Omics" field and the subsequent assimilation of genome, proteome, metabolome and interactome sciences in cellular and multicellular systems is essential for drug discovery.

3. References

Aldridge et al., 2006, Physicochemical modeling of cell signaling pathways, Nature Cell Biology, 8:1195-1203.

Antonio del Sol et al., 2010, Diseases as network perturbations, Current Opinion in Biotechnology, 21:566-571.

Azuaje et al., 2010, Identification of potential targets in biological signaling systems through network perturbation analysis, Biosystems, 100:55-64.

Bergter and Vermelho, 2010. Structures of Glycolipids Found in Trypanosomatids: Contribution to Parasite Functions, The Open Parasitology Journal, 2010, 4:84-97.

Cavalieri and Filippo, 2005, Bioinformatic methods for integrating whole-genome expression results into cellular networks, DrugDiscoveryToday, 10:727-734.

Chavali et al., 2008. Systems analysis of metabolism in the pathogenic trypanosomatid Leishmania major, Molecular Systems Biology, 4:1-19.

Davidov et al., 2003. Advancing drug discovery through systems biology, Drug Discovery Today, 8:175-183.

Douglas B. Kell, 2006. Systems biology, metabolic modeling and metabolomics in drug discovery and development, Drug discovery Today, 11:1085-1092.

Falkow S, 1988, Molecular Koch's postulates applied to microbial pathogenicity, Rev Infect Dis, 10:S274-S276.

Funahashi et al., 2003, Cell Designer: a process diagram editor for gene regulatory and biochemical network, Biosilico, 1: 159-162.

Haughan et al., 1995, Effects of an azasterol inhibitor of sterol 24-transmethylation on sterol biosynthesis and growth of Leishmania donovani promastigotes, Biochem J 308: 31–38.

G.F. Spath, 2000. Lipophosphoglycan is a virulence factor distinct from related glycoconjugates in the protozoan parasite Leishmania major, PNAS, 97:9258–9263.

Horn and Jackson 2004. General mass Action Kinetics. Archive for rational Mechanics and Analysis 47:81-116.

Karplus and McCammon, 2002. Molecular dynamics simulations of biomolecules, Nature structural Biology, 9:646-788.

Lipka, G., Chowdhry, B.Z. and Sturtevant, J.M. 1984 A comparison of the phase transition properties of 1,2-diacylphosphatidylcholines and 1,2-diacylphosphatidylethanolamines in water and deuterium oxide. *J. Phys. Chem.*, 88, pp 5401–5406.

Lee et al., 2006. Flux balance analysis in the era of metabolomics, Brief Bioinform, 7:140-150.

Liebermeister and Klipp. 2006. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints, Theoretical Biology and Medical Modelling, 3:1186- 1192.

Marchisio and Stelling, 2009. Computational design tools for synthetic biology,Current Opinion in biotechnology, 20:479-485.

Niemela et al., 2009. Bioinformatics and Computational method for lipidomics, Journal of Chromatography B, 877:2855-2862.

Oppenheimer et al., 2011. Isolation and characterization of functional Leishmania major virulence factor UDP-galactopyranose mutase, Biochemical and Biophysical Research Communications, 407:552-556.

Perez Victoria,F.J., Sanchez-Canete,M., Seifert,K., Croft, S.L., Sundar,S., Castanys,S., and Gamarro, F. 2006 Mechanisms of experimental resistance of *Leishmania* to miltefosine: Implications for clinical use. Drug Resistance Updates 9, 26–39.

Reed et al., 2006. Towards multidimensional genome annotation, Nat Rev Genet, 7:130-141.

Roberts et al., 2009, Proteomic and network analysis characterize stage-specific metabolism in *Trypanosoma cruzi* ,BMC Systems Biology, 3: 1- 15.

Sergio Grimbs, 2009. Towards structure and dynamics of metabolic networks, Institutional Repository of the University of Potsdam.

Singh 2011 Membrane permeability in Biological Systems: A Systems Biology Perspective. Journal of Computer Science and Systems Biology 4:027-032.

Thiele and Bernhard, 2010, A protocol for generating a high-quality genome-scale metabolic reconstruction,Nature Protocols, 5:93-121.

Tyson et al., 2001, Network dynamics and cell physiology, Nat.Rev.Mol.Cell.Biol. 2:908-916.

Weber and Fusseneger, 2009, The impact of synthetic biology on drug discovery, Drug discovery Today, 14:956-963.

Zhang and Beverley, 2010, Phospholipid and sphingolipid metabolism in Leishmania, Molecular & Biochemical Parasitology, 170:55–64.

Zhang et al., 2004, The LPG1 gene family of *Leishmania major*, Molecular & Biochemical Parasitology, 136:11–23.

Zhang K. et al., 2003, Sphingolipids are essential for differentiation but not growth in Leishmania, EMBO J 22: 6016–6026.

Zufferey and Mamoun 2005. The initial step of glycerolipid metabolism in Leishmania major promastigotes involves a single glycerol-3-phosphate acyltransferase enzyme important for the synthesis of triacylglycerol but not essential for virulence, Molecular Microbiology, 56:800–810.

Zufferey et al. 2003. Ether phospholipids and glycosylinositolphospholipids are not required for amastigote virulence or for inhibition of macrophage activation by Leishmania major, J Biol Chem 278: 44708–44718.
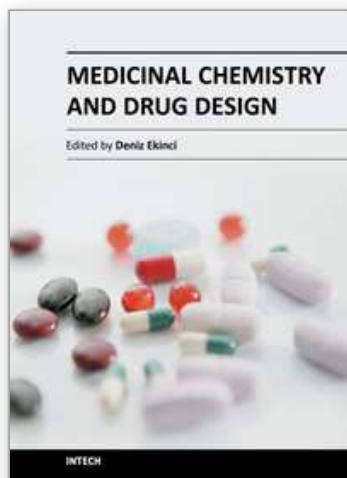
Medicinal Chemistry and Drug Design

Edited by Prof. Deniz Ekinci

Over the recent years, medicinal chemistry has become responsible for explaining interactions of chemical molecules processes such that many scientists in the life sciences from agronomy to medicine are engaged in medicinal research. This book contains an overview focusing on the research area of enzyme inhibitors, molecular aspects of drug metabolism, organic synthesis, prodrug synthesis, in silico studies and chemical compounds used in relevant approaches. The book deals with basic issues and some of the recent developments in medicinal chemistry and drug design. Particular emphasis is devoted to both theoretical and experimental aspect of modern drug design. The primary target audience for the book includes students, researchers, biologists, chemists, chemical engineers and professionals who are interested in associated areas. The textbook is written by international scientists with expertise in chemistry, protein biochemistry, enzymology, molecular biology and genetics many of which are active in biochemical and biomedical research. We hope that the textbook will enhance the knowledge of scientists in the complexities of some medicinal approaches; it will stimulate both professionals and students to dedicate part of their future research in understanding relevant mechanisms and applications of medicinal chemistry and drug design.

How to reference

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

Shailza Singh and Sonali Shinde (2012). Stochastic Simulation for Biochemical Reaction Networks in Infectious Disease, Medicinal Chemistry and Drug Design, Prof. Deniz Ekinci (Ed.), ISBN: 978-953-51-0513-8, InTech, Available from: http://www.intechopen.com/books/medicinal-chemistry-and-drug-design/stochastic-simulation-for-biochemical-reaction-networks-in-infectious-disease

INTECH
open science | open minds