

# Chapter 8

## Systems Approaches to Study Infectious Diseases

Priyanka Baloni, Soma Ghosh and Nagasuma Chandra

**Abstract** Exposure to infectious agents can either lead to active disease or containment or killing of the pathogen. Outcome of an infectious disease is determined by the complex interplay between the host and the pathogen. Therefore, understanding the crosstalk between the host and the pathogen during infection is crucial to identify molecules that are important for the spread or suppression of the disease and for identification of drug targets. Both the host and the pathogen have several mechanisms for countering each other thereby adding layers of complexity to the host-pathogen interplay. Reconstructing mathematical models of complex processes such as cell regulations, signal transductions and host-pathogen interactions provide a detailed understanding of the various interactions and crosstalks occurring in a biological system and thus form a platform to study the system as a whole. Various experimental methods in functional genomics and proteomics as well as computational approaches have been developed over the years that help in building and modeling the biological systems. These approaches have proved quite helpful in identifying drug targets, generating hypotheses rationalizing and finally predicting the cause and final outcome of diseases.

**Keywords** Infectious disease · Host-pathogen interactions · Flux-balance analysis · Stoichiometric matrix · Response networks · Gillespie's algorithm · Boolean modelling · Networks · Tuberculosis · Malaria · *Plasmodium falciparum* · plasmODB · Cholera · targetTB

---

N. Chandra (✉) · S. Ghosh · P. Baloni  
Department of Biochemistry, Indian Institute of Science, Bangalore, India  
e-mail: nchandra@biochem.iisc.ernet.in

S. Ghosh  
e-mail: soma@mbu.iisc.ernet.in

P. Baloni  
e-mail: baloni@mbu.iisc.ernet.in

## 8.1 Introduction

Infectious diseases are directly responsible for about a third of all deaths occurring worldwide. Tuberculosis, pneumonia, malaria, cholera are among the most fatal infectious diseases, responsible for 58 % child mortality in developing nations (WHO 2012). These infectious diseases can be categorized depending upon their frequency of occurrence into sporadic, endemic, epidemic or pandemic diseases. Although several anti-infective drugs are available for these diseases, they continue to be a burden to human health, a problem further compounded by the emergence of drug resistant varieties of the pathogens (Spellberg et al. 2008), (MacPherson et al. 2009). Discovery of newer, safer and robust drugs require the formulation of new strategies that involve innovative ways of tackling the diseases. It has now become increasingly clear that strategies stemming from holistic system approaches may hold the key for effective and sustained management of infectious diseases (Aderem et al. 2011). A wealth of molecular level data has been gathered over the years on several causative microorganisms, which has increased substantially due to the advances in genomics and other high-throughput technologies. The scale and the complexity of each piece of data, is indeed quite high and requires computational analysis to help in comprehending and making useful inferences from it.

Systems biology is the study of large scale systems, reconstructed from many small scale interactions. This approach is based on the premise that the ‘whole is greater than the sum of its parts’ (Hood and Perlmutter 2004). It provides a holistic understanding of the biological function from molecular and cellular level to an entire organism and serves as a platform to study and correlate the processes occurring in a complex living system at different scales to understand a biological phenomenon. Application of such computational methods is evident in the field of drug discovery. Simulations using reconstructed models further aid in knowledge based drug target identification, discovery of biomarkers as well as for rational design of vaccines. Overall, studying a system as a whole rather than individual molecular characterizations performed in isolation would be required to understand the phenotypic behaviour of a given system.

With advances in techniques such as high-throughput sequencing, microarrays, nuclear magnetic resonance and mass spectrometry, it is now possible to get better insights into the field of transcriptomics, proteomics and metabolomics, and the data generated using these techniques serve as direct inputs into development of systems level models. The large scale omics data are analyzed using computational methods to derive essential molecular interactions. These molecular interactions are used to build a detailed mathematical model to represent the biological system being studied. Once validated, these models are used to simulate a range of scenarios to predict the behaviour of the system under various conditions. The hypotheses generated can be taken back to the bench again and validated using focused experimental studies (Aderem et al. 2011; Vodovotz et al. 2008). Systems biology, thus, along with different ‘omics’ studies is being increasingly used to identify pathways involved in specific disease conditions, establish interconnectedness of different pathways

and understand cellular responses to various certain conditions including physiological stress and exposure to a pathogenic organism (da Hora Junior et al. 2012; Day et al. 2010; Kitano 2002; Weckwerth 2003; Weston and Hood 2004).

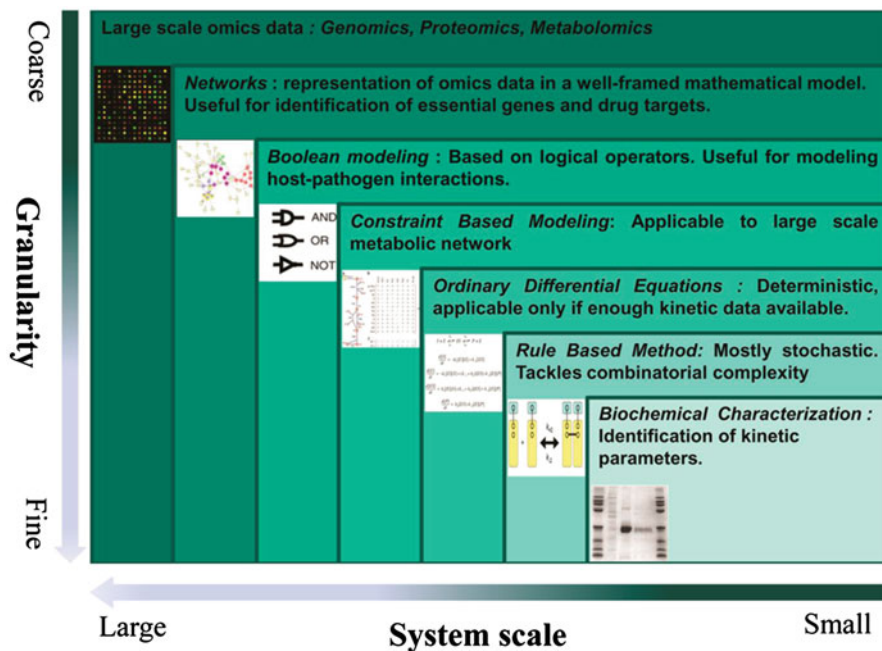
The study of host-pathogen interactions focuses upon the interactions between microbial or viral pathogens and their plant or animal hosts. The interactions are multi-faceted and form a complex network including moves and counter-moves from both species leading to one of two broad outcomes, either clearance or proliferation of bacteria (Forst 2006; Johanns et al. 2010). Using systems biology approaches it has become feasible to study various phenomena such as recognition of the pathogen by the host immune system, mechanism of virulence, pathogenesis, mechanisms of antibiotic resistance, persistence of disease all as aspects of the complex host-pathogen interplay, the knowledge ultimately useful for biomarker and drug target identification (Weston and Hood 2004; Wang et al. 2010a). Systems biology as a discipline, in fact utilizes both experimental and computational approaches to build computationally amenable mathematical models of complex biological processes. This chapter provides an overview of various systems biology approaches available for studying causative organisms that cause infectious diseases and also the interplay between host and pathogen. In particular, the chapter focuses on the various modeling approaches that are available and being utilized for such studies and summarizes various insights obtained for a few important infectious diseases.

## 8.2 Modeling Methods

Deciphering functions of individual components even at a genome scale is not sufficient to understand the complexity of the organism or the complex interplay between the host and pathogen. Availability of large scale genomics, proteomics and metabolomics data have led to advances in obtaining pair-wise interactions between pairs of molecules. Different pieces of data are required to be pooled together using mathematical formalisms to build up a biological system, which can be used to address various biological questions. This also provides a handle to the experimentalist to prioritize the proteins for functional studies. Various modeling methods that are commonly used in the field of systems biology are described briefly here and are also depicted in Fig. 8.1. The models are ordered according to the level of granularity in the figure.

### 8.2.1 Networks

The parts lists obtained from individual *omics* level experiments starting from the genome sequencing are assembled based on various molecular interactions obtained experimentally through a number of studies documented in literature. The list of protein-protein interactions are augmented substantially through a variety of



**Fig. 8.1** The different modeling methods used in Systems biology. Methods are colour coded based on granularity

knowledge-based predictions using methods based on Rosetta stone concept (Marcotte et al. 1999), phylogenetic profiling (Pellegrini et al. 1999), gene-neighbourhood and its conservation (Dandekar et al. 1998). The set of pair wise interactions and genome-wide functional linkages (Strong et al. 2003) thus identified, ultimately lead to network reconstructions. Databases such as STRING (Szklarczyk et al. 2011) in fact make this available to the community in a comprehensive manner.

Individual molecular constituents in the cell form nodes, while interactions between them form edges, put together forming large complex graphs. Graph theory can then be used to understand and explore various aspects of the cell in different conditions (Albert 2007). Depending on the system being reconstructed, directed (eg. signalling networks), undirected (protein-protein interactions) or bipartite networks (metabolite-enzyme) can be generated. The edges can be further weighed if appropriate experimental data is available. Protein-protein interaction networks representing interactomes serve to understand the dynamics of a biological cell. Shortest path analysis has been used to identify criticality of particular nodes in the network (Ravasz et al. 2002). Through systematic knock-outs or node or edge deletions, nodes leading to significant number of broken paths and hence their relative importance in the network is assessed. These networks can be further divided into sub networks based on the intra and inter connectivity and represent the different functional modules present in the system.

Although network analyses helps in identifying important and influential molecules in a system and study the communication between the molecules in detail, it is mostly static in nature and captures a single condition in most cases. Static networks do not provide a complete understanding of the system, but reflect a single snapshot of the numerous possible interactions that can occur as a result of the various adaptive and environmental changes at that instant of time. One approach to overcome this limitation is reported by Ideker et al., who integrated mRNA expression data into a yeast protein-protein and protein-DNA interaction network, to identify subnetworks that were most active under different conditions (Ideker et al. 2002). Active sub networks were identified by calculating the significant fold change of each gene in that subnetwork as a result of changing conditions. The high scoring subnetworks correlated well with known regulatory mechanism. Such active subnetworks that convey a systems response given an experimental condition are termed as response networks (Forst 2006).

Reconstruction of signaling networks, where nodes are signaling components and directional edges are the regulations, helps understand the signaling cascading events taking place inside a cell. Interactions can be tagged as positive or negative or stimulatory or inhibitory (Wang and Albert 2011). Importance of a node is determined by studying the effect of that node's deletion on the propagation of the signal. Minimal set of nodes that can perform signal transduction independently have also been identified using this method.

Organism specific metabolic networks have been constructed and studied using methods such as flux-balance analysis. This requires three basic types of data; (a) enzyme, corresponding substrates and products, (b) stoichiometric matrix of all reactions which gives the ratio in which the substrates and products participate in the reaction and (c) cellular location of the reaction (Feist et al. 2008). Biochemical pathways can be represented using different network types. In a metabolite network, metabolites form nodes and two nodes are connected if they share a substrate-product relationship. In a reaction network nodes represent reactions and two reactions are connected if the product of one forms a substrate for the other. Bipartite networks are useful representations to capture biochemical pathways. A bipartite network contains two types of nodes and an edge can only be drawn between two different types of nodes. In case of biochemical pathways, enzymes form one set of nodes, while metabolites form another set of nodes and a connection can be made only between an enzyme and a metabolite (Raman et al. 2006). Detailed networks can also be built where kinetic information is incorporated as weights in the network. Metabolic networks are analysed using the graph theory tools to identify hubs and cluster the reactions based on their functions. Other tools such as Petri-nets (Pinney et al. 2003) have also been used to study various properties of an organism. Cytoscape (Shannon et al. 2003) is used widely to visualize as well as perform basic network analysis. The Boost Graph Library (Siek et al. 2002) implementation of MATLAB is also frequently used to perform network analysis.

## 8.2.2 Constraint Based Modeling

Constraint based modeling approaches are being used widely for studying metabolism in a cell. Metabolic reactions are represented using a stoichiometric matrix of size  $m \times n$ , where rows represent metabolites ( $m$ ) and columns represent all the reactions ( $n$ ) present in an organism. Entries in the matrix represent the stoichiometric coefficients of the metabolites in the reaction (Orth et al. 2010; Raman and Chandra 2009). Given the stoichiometric matrix ( $S$ ), FBA aims to calculate the flux ( $v$ ) through each reaction at steady state, such that  $S \cdot v = 0$ . These models are further constrained to mimic biological systems such that a unique flux distribution for the organism is obtained using linear optimization. An interesting feature of FBA is its ability to perform single and multiple gene deletion knockouts. This is done by constraining the bounds of all the reactions coded by that gene to zero. This analysis helps in identifying essential genes and drug targets (Raman et al. 2005). Effect of inhibitors can also be studied by constraining the required reaction to a fraction of the wild type bounds. Segre et al. developed a variant of FBA known as MoMA (Segre et al. 2002), which unlike FBA is not solely based on optimizing the objective function. The idea being that any genetically modified organism may not achieve optimality since the mutant strains are not subjected to long term evolutionary pressures and may perhaps attempt to attain biological function via minimal changes in the flux distribution.

A major advantage of constraint based modeling is that they do not require a detailed understanding of the reaction mechanism or other kinetic parameters to perform *in silico* simulations. Many modifications to the original methods have been reported to incorporate gene expression data (Colijn et al. 2009) and other *omics* data (Schellenberger et al. 2011) to obtain a better mimic of the biological system under investigation. Various tools such as FAME (Boele et al. 2012), FASIMU (Hoppe et al. 2011), COBRA toolbox (Schellenberger et al. 2011), MetaFlux (Latendresse et al. 2012) have been developed over the years to perform FBA and its variants (Lakshmanan et al. 2012).

## 8.2.3 Kinetic Modeling Using Ordinary Differential Equations

Biochemical reactions have classically been represented as differential equations that define the rate of consumption or production of metabolites. Given the kinetic details of any set of reactions, one can build a mathematical model by forming a system of ordinary differential equations (de Jong 2002). Simulations from ordinary differential equations (ODEs) are much more reliable and precise as they are built and analysed using detailed kinetic parameters. An obvious advantage of this method over FBA is that the time evolution of the model can be studied to obtain a detailed understanding of the system, instead of only analysing the steady state behaviour. However, non-availability of kinetic data limits the broad applicability of this method. MATLAB is widely used to solve the system of ODEs contained in these models. Other software packages such as JDesigner (Sauro 2004), Cell Designer (Funahashi et al. 2003), and Copasi (Hoop et al. 2006) are also commonly used for this purpose.

### 8.2.4 *Boolean Modeling*

Boolean modeling also called as logic modeling is being used to model complex biochemical systems and capture the qualitative behaviour of the biological system. Each component in the model can exist in two states, either *on* or *off*. Transition from one state to another is encoded using logical operators. One of the major advantages of logic modeling is the ease with which complex molecular interactions can be represented and therefore these are widely used to model complex biological phenomenon such as apoptosis (Schlatter et al. 2009) or host-pathogen interactions (Raman et al. 2010). New methodologies are being continually developed that transforms Boolean models into a continuous model so as to study the time course evolution of a biological system. State transition rates of each nodes are calculated using mathematical tools such as Markov processes and multivariate polynomial interpolation (Wittmann et al. 2009; Stoll et al. 2012).

### 8.2.5 *Rule Based Modeling*

In a rule based model, the biological system is defined using a set of rules. These rules use the notation of a simple chemical reaction and describe the local events taking place inside a cellular system that eventually leads to the emergence of a global property. This method is based on the principle of Gillespie's algorithm (Gillespie 1977), according to which a cell is considered as a well-mixed system and interaction between any two molecules in the cell is dependent on the rate of interaction between the two and the abundances of each molecule interacting. This method is particularly useful when modeling any regulatory system as these systems are inherently complex in nature and have the potential to generate a variety of distinct species as a result of the cascading events that occur in such systems. Formally, due to combinatorial complexity arising from the set of possible interactions in the system, a large number of distinct species are generated, which can all be systematically studied and outcomes of specific scenarios predicted (Hlavacek and Faeder 2009). Rule based methods are also being explored as tools for multi-level modeling of biological systems (Maus et al. 2011). Software tools such as BioNetGen (Blinov et al. 2004), Kappa (Danos et al. 2008), RuleMonkey (Colvin et al. 2010) have been used for rule based modeling. These methods are generally stochastic in nature; however the rules can be rewritten as ODEs to build deterministic models.

### 8.2.6 *Models of Host-Pathogen Interactions*

Understanding the outcome of an infectious disease not only requires a detailed study of the host and pathogen system individually, but more importantly, the communication and the crosstalk that occurs between the two systems. Individual models of

host and pathogens describing different biological processes are widely available and can be easily manipulated to obtain a host-pathogen model. Such models provide a detailed description of the crosstalk that exists between the two systems as well as the individual processes. This provides a realistic picture of the biological phenomenon being studied and also helps in extrapolating the influence of such crosstalk on host and pathogen.

Host-pathogen interactions have been modeled using several approaches, ranging from simpler models for the prediction of protein—protein interactions between the host and pathogen, to complex models for the metabolic and signal transduction networks. Kirschner and co-workers have developed a virtual model of the host immune response to *M.tb* using agent-based modeling methods (Marino et al. 2011). Numerous insights about critical factors and parameters governing host-pathogen interactions can be obtained through these studies. Integrating the host and pathogen FBA models and further modification of the optimization function have also been used to study host-pathogen interactions (Bordbar et al. 2010).

Different types of approaches can be integrated each of which best describes different aspects of a biological system to obtain overall mechanistic insights. For example, FBA is used for studying metabolic networks while Boolean modeling is used for regulatory networks and the approaches can be clubbed to obtain a metabolic as well as a regulatory model. This is important because the different modules of a biological system interact with each other and influences the functioning of the modules. Covert et al. (2008) have developed a method, iFBA, also known as integrative FBA that integrates FBA with Boolean logic and ODEs to model the dynamics of networks related to the carbohydrate uptake mechanism. They compared the predictions of the integrated model with the individual model and showed that an integrated model is a significant improvement over the individual models. The applications of these methods are described using case studies of different infectious diseases and are presented in the succeeding sections.

### 8.3 Tuberculosis

According to the sixteenth global report on tuberculosis (TB), published by WHO, an estimated 8.5–9.2 million new cases of TB have emerged in the year 2010, while 0.9–1.2 million of the HIV-negative people have succumbed to the disease, and an additional 0.35 million deaths have occurred from the HIV-associated TB cases. Threat from this disease increases drastically with the advent of multidrug resistant (MDR), extremely-drug resistant (XDR) and totally drug resistant (TDR) strains. Unfortunately, no new drugs have come up in the last five decades and the drugs available in the market have their inadequacies. It is thus important to think of newer strategies and develop new classes of drugs to counter the spread of this disease.

The etiological agent of TB, *Mycobacterium tuberculosis* (*M.tb*), enters the host primarily via aerosols containing the bacilli, and on reaching the lungs they are internalized by the alveolar macrophages and undergo phagocytosis. Pathogenesis starts



after formation of the phagosome, wherein *M.tb* prevents maturation of the infected macrophage and in this niche the pathogen is able to survive and reproduce. The widespread nature of this disease depends upon its ability to spread easily by aerosol transmission, which is further facilitated by immune-dependent tissue-damaging inflammation (Pieters 2008).

Upon infection, a dynamic interplay occurs between the host and pathogen leading to either of the four outcomes: (a) the initial host response may be completely effective and kill the bacilli; (b) the organisms can grow and multiply immediately after infection resulting in active TB, (c) the bacilli may become dormant and never cause disease at all and (d) the latent bacilli can eventually become active and progress to disease condition (Schluger and Rom 1998). Needless to say, the difference between the outcomes is enormous and results in extreme phenotypes between disease and health. Various experimental as well as computational tools have been used to study the pathogenesis of this disease and its interaction with the host, briefly summarized here.

Deciphering the whole genome sequence of *M.tb* has been an important landmark in tuberculosis research (Cole et al. 1998). The genome sequence provided a first comprehensive parts-list of the molecular constituents of the cell. This triggered extensive amount of downstream research leading to detailed biochemical and biophysical characterizations of a number proteins (Lew et al. 2011; Galagan et al. 2010). More importantly perhaps, it has provided an impetus for systems level studies. Genome sequence has helped tremendously in completing the gaps in knowledge from decades of biochemical and molecular biology studies of individual molecules in the organism. It has revealed complete lists of proteins belonging to many biochemical pathways, transcription factors, two-component signalling systems (Tyagi and Sharma 2004). It has led to comparative genomics studies through gene and protein sequence comparisons and further to several functional genomics studies (Tucker et al. 2007). Proteins responsible for cellular metabolism are identified comprehensively; indicating that, *M.tb* indeed has most of the standard pathways present in other bacteria such as glycolysis, citric acid cycle, pyruvate, fatty acid, amino acid metabolism to list a few (Cole et al. 1998). There are also interesting differences, for example, presence of mycolic acid and arabinogalactan pathways, the glyoxylate shunt and beta oxidation pathway for fatty acid metabolism. Identification of such unique features has been useful to obtain direct explanations for phenotypic characteristics of the organism such as the presence of a thick waxy outer cover.

Advances in high-throughput 'omics' technologies, that has resulted in a large amount of omics data in the last few years, help significantly in functional characterizations (Kirschner et al. 2010) of both host and pathogen's genomes. Global gene expression profiles of *M.tb* under different conditions are available. The set of genes in *M.tb* required for optimal growth have been characterized by using the transposon site hybridization (TraSH) method which provides a comprehensive idea about functional significance and essentiality of each gene (Sasseti et al. 2003). The proteome of *M.tb* has also been analyzed by 2D gel electrophoresis and mass spectrometry and also by the isotope-coded affinity tag reagent method coupled with mass spectrometry (Schmidt et al. 2004). Using a guinea pig model of tuberculosis, the bacterial proteome during the early and chronic stages of disease has been examined

(Kruh et al. 2010) by liquid chromatography-mass spectrometry. The study identified numerous *M.tb* proteins, from essential kinases to products involved in metal regulation and cell wall remodeling, present throughout the course of infection. Cell wall processes, intermediary metabolism and respiration were found to be major functional classes of proteins represented in the infected lung. Recently, protein-protein interactions in *M.tb* have been determined experimentally in a high-throughput manner using a bacterial two-hybrid system (Wang et al. 2010a).

Genome scale studies are being carried out for the host systems as well. Several gene expression profiles under different conditions of exposure to *M.tb*, disease and treatment with anti-tuberculars have been obtained, which identify genes that show maximal changes in their expression under different conditions (Boshoff et al. 2004). siRNA screens have been used to systematically knock-out various genes and infer their importance for survival, pathogenesis and stress response (Kumar et al. 2010). Recently many techniques have been developed to visualize spatial features of such interactions inside tissues, which include intravital multiphoton microscopy and four dimensional FRET (Konjufca and Miller 2009; Hoppe et al. 2009). Although these techniques are in their incipient stages of development, they offer promising results and greater understanding of host-pathogen interactions.

The data thus obtained from the above described *omics*-data can be further used to build computational models. One way of incorporating such large scale data is to build a protein-protein interaction network. A comprehensive reconstruction using crowd sourcing based curation from literature and available databases together, capture as many as 71086 interactions in 3967 proteins (Vashisht et al. 2012) adding substantially to the existing resources. Incorporating drug-specific gene-expression fold changes in the network as node weights, Padiadpu et al. (2010) captured the effect of drugs on *M.tb* interactome and the mechanism of triggering resistance. Another study by Kauffman et al. (Rachman et al. 2006) identified genes that are important for the survival and persistence of *M.tb* in a macrophage cell by using a combination of approaches. Using a reconstructed protein-protein interaction network and incorporating genome-wide DNA array into this network, pathways such as iron metabolism, cell wall synthesis, DNA damage repair and fatty acid degradation were identified as important to the pathogen (Rachman et al. 2006).

Yet another method of using experimental data to build computational models is constraint based modeling. Details of this modeling method are provided in the methods section. This method serves as an excellent tool to study genome scale metabolic models. McFadden and co workers (Beste et al. 2007) reconstructed the first genome scale metabolic model for *M.tb*, capturing all known biosynthetic pathways operational in a cell for synthesis of major macromolecular components. This model was calibrated using data from chemostat cultivations of *M.bovis* BCG in continuous culture and measurement of steady state growth parameters. Almost at the same time, an independently reconstructed genome scale network model of *M.tb* H37Rv named iNJ661 was reported by Palsson and coworkers (Jamshidi and Palsson 2007). The authors grew this bacterial model *in silico* on various media, and observed that growth rates were comparable to experimental observations of doubling times in the range of 12–24 h in different media. Using these models, reaction

fluxes indicating substrate consumption rates were measured, which correlated well with experimentally determined values. Raman et al. have identified putative drug targets using *in silico* gene deletions for the mycolic acid pathway model in *M.tb* (Raman et al. 2005).

Another classical method to study the dynamics of a cellular system is ordinary differential Equations (ODE), wherein time courses of metabolic reactions are mathematically represented by ODEs. Singh et al. (Singh and Ghosh 2006) built a kinetic model of the tricarboxylic acid cycle and the glycolytic pass of *E.coli* and *M.tb* to compare the two systems and study the effect of enzyme inhibition and thus identify potential drug targets. Kinetic modeling has also been carried out to study the host immune system upon TB infection to reveal the existence of a non-infected steady state and an endemically infected steady state, which can lead to latency or activation of the disease (Ibargüen-Mondragón et al. 2011)

Signalling interactions in a cell can be easily represented by Boolean modeling, also described in the methodology section. Raman et al. built a Boolean model of the host—pathogen interactome (Raman et al. 2010), accounting for several mechanisms of invasion by the pathogen, defense of the host, as well as the defense mechanisms of the pathogen and was simulated under a variety of conditions. The model consisted of 75 nodes that represented the molecules involved in host and pathogen and different states of the molecules and events were governed by logical operators or Boolean rules. This provides a framework to understand the conditions and parameters that favour clearance versus those that favour either active disease or contain the bacteria in a dormant state.

Rule based modeling have also been used to represent signalling processes, especially for those events, wherein the molecule can take up different states depending on its environment. Such models are known to best capture the environmental dependencies. An et al. (An and Faeder 2009) built a rule based model of the Toll-like receptor 4 signal transduction cascade. Simulation of the original model and ‘knock-out’ were performed to study the behaviour of the system. Ghosh et al., have reported a rule based model to study host-pathogen interaction for TB infection and the role of iron for both host and pathogen during the course of infection has been studied. Regulating the concentration of mycobactin was discussed as one of the strategies to control bacterial infection (Ghosh et al. 2011).

Boolean network models of immunological components of the interplay of various mechanisms of attack and defense in the host and pathogen with respect to *M.tb* have been developed and provides insights into the immune responses as well as the different outcomes of *M.tb* infections under different conditions (Raman et al. 2010). Kirschner and co-workers have worked on several mathematical models for the interaction of *M.tb* with the human immune system, some examples of which are a virtual model of the immune response to *M.tb* that characterises the cytokine and cellular network during infection, two compartmental models capturing the important processes of cellular activation and priming capable of reproducing typical disease progression scenarios, agent-based models for simulating granuloma formation (Marino et al. 2011) and a mathematical model describing macrophage biochemical processes based on activation, killing and iron regulation. Host-pathogen FBA models

enable studying the metabolic states of the system in an infected condition. Gene essentiality studies were performed and the predictions were shown to be much more accurate in the combined model. The models were further integrated with gene expression data for the different forms of the disease, such as latency, meningal and pulmonary tuberculosis, to study the subtle metabolic differences amongst the different forms and therefore to have much more accurate perturbation studies for the different forms (Bordbar et al. 2010).

The above methodologies have helped in successfully identifying the different aspects of *M.tb* infection. Protein-protein interactome analyses have helped in identifying highly influential proteins that can form potential drug targets (Padiadpu et al. 2010). Metabolic reconstructions of the host and pathogen as well as the combined models have provided useful insights into genes essential for the survival of the pathogen using FBA (Jamshidi and Palsson 2007). Further, integrating host and pathogen FBA models have provided useful insights into the metabolic changes that occur in the host upon bacterial infection (Bordbar et al. 2010). Host-pathogen interaction studies guide in identifying factors important for virulence, the different immune responses and most importantly understanding the emergence of resistance (Raman et al. 2010). A new concept of *co-targets* was proposed by Raman et al. that inhibited two targets simultaneously to deal with resistance. All these analyses have been integrated into a rational pipeline called targetTB to identify potential drug targets for *M.tb* (Raman et al. 2008), which has yielded a list of about 450 high confidence drug targets.

## 8.4 Malaria

Malaria caused by *Plasmodium* parasites, is transmitted through the bite of infected Anopheles mosquito. In 2011, an estimated number of 216 million cases of malaria were reported and 655000 deaths were caused by malaria in 2010 (World malaria report 2011), indicating that it is one of the major contributors to global morbidity and mortality rates. Although malaria is curable, it is still a life-threatening disease, and with the emergence of antimalarial resistant strains it has become difficult to tackle this disease efficiently.

Whole genome sequencing of *Plasmodium falciparum* was accomplished in 2002 (Gardner et al. 2002) and it has revealed that approximately 35% of the proteins encoded have identifiable function and the remaining are uncharacterized. With the availability of genomic sequence of *P.falciparum* it has become easier to identify unique enzymes involved in pathways, which are different from the humans, such that inhibitors can be synthesized against them, thus disrupting the pathway in pathogen. Mass spectrometric studies have been performed in order to understand the mechanism by which the parasite modulates the level of different metabolites taking part in various metabolic processes of the host so as to survive inside the host cell and proliferate (Olszewski et al. 2009). Due to the complex life cycle of the pathogen, it becomes necessary to identify genes expressed at different stages of infection such

that they can be used as targets (Winzeler 2005). A combination of genomics and proteomics methods were employed by Hall et al. (2005) in order to identify a conserved set of genes in *Plasmodium spp.* and also emphasize upon genes which have been chosen under selective pressure at different stages of pathogenesis. Flux balance model for *P.falciparum* was constructed in order to study the metabolic state of the pathogen upon perturbation and also predict the essential genes which can also be used as targets (Plata et al. 2010). The model consisted of 1001 reactions and 616 metabolites, of which enzyme-gene associations were reported for 366 genes and 75 % of the total enzymatic reactions known. Models were enriched by incorporating gene-expression data and also the accuracy of the predictions to experimental results was high indicating that *in silico* models can be used for studying the complex pathogen. An open access database called PlasmoDB has been developed which provides information about the transcriptome and protein expression data of *Plasmodium spp.* at different stages of their life cycle, which can be used to investigate the involvement of a gene in a defined process by correlating with gene expression profiles or proteomics or protein-protein interactions data of the species (Aurrecochea et al. 2009).

*Plasmodium spp.* is capable of surviving inside the host by synthesizing different chemical compounds during various stages of its life cycle. Although these compounds have been used as targets for vaccine development, not much success has been achieved in eradicating malaria. Due to the complex host-pathogen interaction and prevalence of resistance to antimalarial drugs, efforts have been made to discover newer drugs using a systems biology approach. The immune response of the host plays a complicated role in malaria as it not only helps in evading the pathogen but is also responsible for causing complications in the host (McNicholl et al. 2000). Jomaa et al. reported a non-mevalonate pathway of isoprenoid biosynthesis, located in the apicoplast region of *Plasmodium*, and the drugs effective against the metabolites involved in this pathway as potent antimalarials (Jomaa et al. 1999). Reverse vaccinology approach has been employed to search for antigens in *Plasmodium spp.* which when targeted will appropriately, aid in vaccine development. Systems biology has been used to anticipate the immune response of the host cells upon the interaction with the antigen and also understand the complex life cycle of the parasite (Rappuoli and Aderem 2011). Bioinformatics approaches have been used to annotate the genome of *Plasmodium spp.*, majority of which is still uncharacterized. Fed into systems biology models, simulations help in discovering newer therapies for malaria as the parasite has acquired resistance against known drugs. Number of potent antimalarials (artemisinin and its derivatives) has been synthesized and systems biology based approaches will aid in characterizing the mechanism of action of these newly discovered antimalarial compounds (Dharia et al. 2010).

## 8.5 Cholera

Reports from WHO indicate that 3.5 million suffer from diarrhoeal infections, the causative agent being *Vibrio cholerae*, capable of secreting the potent cholera toxin (Nelson et al. 2009). This acute intestinal infection is transmitted through contaminated food and water and if left untreated can lead to death of the patients. Although it

is curable if treated on time, severe symptoms are observed in immune-compromised patients. The strains of *V.cholerae* have been classified either as classical or El Tor. Two sero groups, *V. cholerae* O1 and *V. cholerae* O139, are mostly responsible for the outbreak of cholera. Multidisciplinary approaches are being used to find new drugs to reduce the number of deaths caused by cholera.

Top down approaches have been used to identify additional genes that are involved in *V.cholerae* virulence and colonization inside host intestine (Kaper et al. 1995). Apart from the enterotoxin produced by *V.cholerae*, Asaduzzaman et al. have also narrowed down on other essential virulence factors present in the bacterium such as toxin-coregulated pilus that functions as a receptor for the bacteriophage and encoding cholera toxin genes (Asaduzzaman et al. 2004). A regulator-centric approach has been used to focus upon LysR-type transcriptional regulators (LTTRs), one of the most diverse families of transcriptional factors in prokaryotes having role in wide range of processes. A few LTTRs were found to be involved in intestinal colonization as well as metabolic regulation *in vivo* (Bogard et al. 2012). Mathematical models have been developed to understand the dynamics of pathogen colonization and indicate the contribution of host and pathogen towards bacterial gut density (Spagnuolo et al. 2011). Such studies are essential to understand pathogenesis of the disease. By performing a high-throughput phenotypic screen of 50,000-compound small molecule library, Hung and coworkers tried to identify inhibitors of *V.cholerae* virulence factor expression (Hung et al. 2005). The authors have reported a compound named virstatin, which is capable of inhibiting virulence expression, ToxT regulation (part of ToxR regulon, responsible for virulence) post-transcriptionally, and also preventing colonization in the intestine of the animal model to an extent.

Although cholera is a re-emerging disease, till date no simple assay has been developed to diagnose this disease efficiently. Oral or IV rehydration are recommended treatment and thus administering immediate oral rehydration therapy, rapid recovery of the patients can be observed. Since the late nineteenth century till 1970s, injections of inactivated whole bacteria were used as a vaccine. However, the limitation of these is that they are effective only for short durations. Oral vaccines against cholera were developed to overcome the shortcomings of parenteral vaccines. Till date two major classes of oral cholera vaccines namely killed WC- based and genetically attenuated live vaccines are used to treat cholera (Shin et al. 2011). Although newer vaccines such as Dukoral and Shanchol have received WHO prequalification, these vaccines also have their own limitations, thus keeping the problem of vaccine discovery as an open challenge (WHO 2012).

Systems biology approaches have been used in order to analyze gene expression of *V.cholerae* to identify virulence genes, which may provide a better insight to the infectious process. Using gene-expression data, comparison of the dynamic transcriptomes was carried out for the pathogen growing in different media at various stages of growth. A set of regulatory interactions for genes involved in virulence were identified (Kanjilal et al. 2010). Using information from different sources regarding the pathogen, gene response network has been constructed which is expected to aid in design of biomarkers and therapeutics. A metabolomics approach has been used to measure the extracellular changes in the flux of certain metabolites upon the

administration of cholera toxin in cell lines, and this approach can be extended to study spatial and temporal changes in the metabolites flux, thus providing a clear picture of the metabolic activity in the cell in the presence of toxin (Eklund et al. 2006). Thus, using systems biology approaches it has become possible to identify the genes involved in virulence, interaction of the pathogen with the host, discover new biomarkers for the disease and also develop newer vaccines to overcome the limitations of the already existing vaccines (Hill et al. 2006).

## 8.6 *Staphylococcus aureus* Infection

*Staphylococcus aureus* (*S.aureus*), causative agent of nosocomial infection, is a life threatening pathogen to human population due to the wide range of diseases it causes, especially hospital acquired infections. Apart from the number of infections that this microbe is responsible for, it has also been observed that *S.aureus* is acquiring resistance against multiple antibiotics (Kaatz et al. 2005). In some parts of the world, methicillin resistant strains of *S.aureus* (MRSA) have been reported, which is posing a major health problem. Thus, it has become essential to understand the mechanism of pathogenesis of *S.aureus* and also its interaction with the host.

The global transcriptional profile of the pathogen aids in the study of regulatory genes and also gives insight into the expression profile of the genes under different conditions such as exposure to antibiotics (Kuroda et al. 2003) and stress (Anderson et al. 2006). Plikat et al. have constructed a protein expression map to study proteomes of *S. aureus* Mu50 and its mutants. Using GSEA (Gene set enrichment analysis), they have carried out studies to determine the virulence factors and pathways affected in mutants. Capsular polysaccharide of *S.aureus* had been earlier regarded as putative protective antigen and hence as possible vaccine candidate. However, subsequent studies noted that the clinical isolates lack a capsule, hence rendering the vaccine ineffective in the clinical trials. They have also reported that multivalent-antigen vaccine is capable of eliciting both cell-mediated and humoral immunity and in turn induce protection against *S.aureus* thus preventing infections at various anatomical sites (Plikat et al. 2007). Systems biology approaches have been used to identify targets in order to develop multivalent-antigen vaccine and also determine host-microbe interaction which helps in understanding the pathogenesis mechanism and ultimately finding a solution for preventing as well as curing the disease.

## 8.7 Applications of Systems Biology in ‘Anti-Infective’ Drug Discovery

With the advent of large scale omics data and the development of various modeling tools, it is possible to build large scale biological models. Although, the reductionist approach provides detailed insights into the molecules responsible for a particular

disease, inhibition of a given protein molecule in isolation is insufficient to provide insights into the effect of this inhibition on the system as a whole. Existence of biologically feasible alternate paths may render this inhibition useless. Systems biology provides a mathematical framework to understand the physiological effect of inhibition in a network of interacting components. In the classical drug discovery regime, a major part of it was a black box and a target was selected based on the end result obtained. Mathematical models obtained can be used to study the effect of inhibition of the targets or exposure of the system to the drug, so that a rational behind the working of each drug is understood. TargetTb (Raman et al. 2008) is one such attempt wherein a comprehensive target identification pipeline is developed for *M.tb*. Many known targets were identified, thus validating the model and many more new targets have been suggested. A total of 451 high confidence potential drug targets were listed. The success rates from such pipelines are likely to be high as target selections are knowledge driven. Methods such as FBA have also been successful in identifying set of essential enzymes in *P.falciparum* and form a starting point for antimalarial drug targets (Huthmache 2010). *Systems vaccinology* is a branch of systems biology that helps in predicting the efficacy of vaccines in a biological system. It is also useful in studying the immunological responses after vaccinations thus helping in vaccine development (Trautmann and Sekaly 2011). Figure 8.2 describes the various applications of Systems biology.

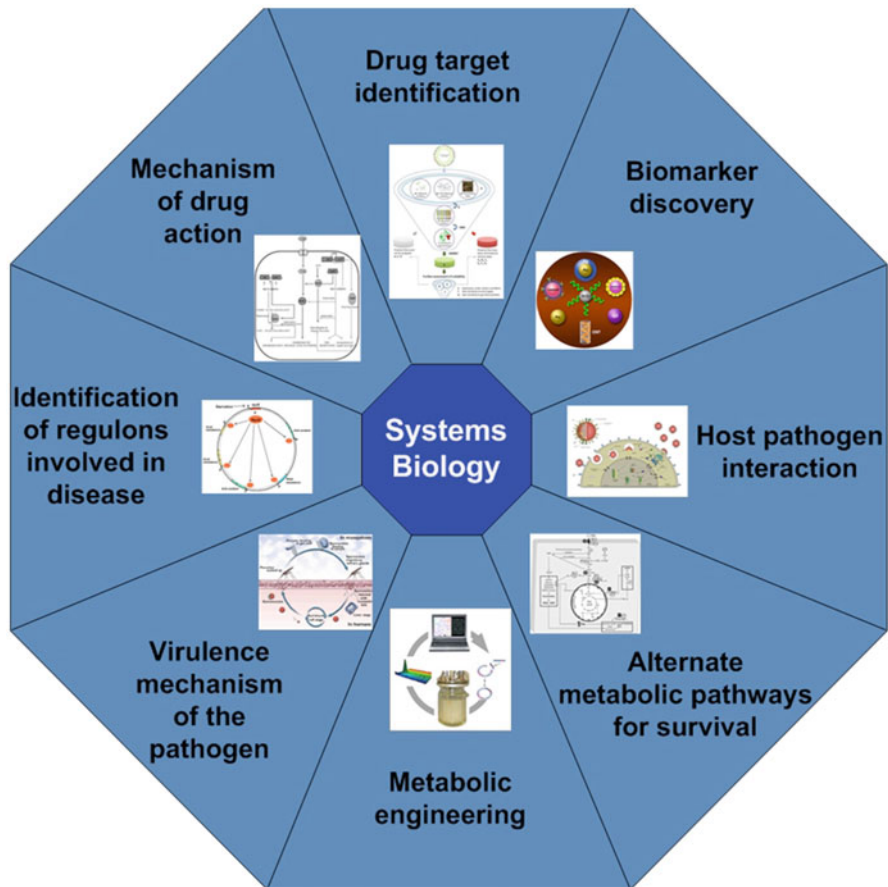
## 8.8 Conclusion

Understanding a biological phenomenon involves studying the system as a whole rather than as parts. Systems biology provides us with the tools to examine different biological aspects, such as protein-protein interactions, protein-metabolites interactions, regulatory mechanisms, signaling cascades using computational means. This is crucial because a continuous interaction exists between different biological processes and therefore studying these processes individually, as carried out in a reductionist approach, do not provide a holistic view of the system under study. Over the years many computational as well as experimental tools have been developed that help in collation, reconstruction and analysis of large-scale data.

The scale at which various molecular level studies are currently being carried out, is yielding genome-scale and systems level data on many fronts, leading to ready reconstructions of large systems. These can then be integrated with the deep insights already available about individual components. Although a complete systems view of the disease has still not been deciphered, it seems that we have at the least a coarse grained map of the pathogen in many of these cases, helpful for obtaining an aerial view of the disease that can be used for addressing a variety of questions. The map of course is sufficiently fine-grained in parts enabling a more detailed zoomed in version in some pathways especially with respect to intermediary metabolism.

Reconstruction of large scale models encompassing various processes of the bacterium and simulation will be extremely valuable in identifying best strategies for





**Fig. 8.2** Various applications of systems biology

intervention. Methods to study biological systems at multiple scales and levels and *virtual cells* are not as yet standardized. Nor are the methods required to generate comprehensive omics scale data from multiple perspectives, particularly when it comes to quantitative profiling. Thus, reports in literature of such cellular level models not only for *M.tb*, but in general for any organism are few and far between. Nevertheless, it is quite clear that the virtual cell approach, especially when quantitative aspects are incorporated, holds a lot of promise for picking an efficient or even an optimal strategy for killing the pathogen.

## References

- Aderem A, Adkins JN, Ansong C, Galagan J, Kaiser S, Korth MJ, Law GL, McDermott JG, Proll SC, Rosenberger C et al (2011) A systems biology approach to infectious disease research: innovating the pathogen-host research paradigm. *mBio* 2(1):e00325–00310
- Albert R (2007) Network inference, analysis, and modeling in systems biology. *Plant Cell* 19(11):3327–3338
- An GC, Faeder JR (2009) Detailed qualitative dynamic knowledge representation using a BioNetGen model of TLR-4 signaling and preconditioning. *Math Biosci* 217(1):53–63
- Anderson KL, Roberts C, Disz T, Vonstein V, Hwang K, Overbeek R, Olson PD, Projan SJ, Dunman PM (2006) Characterization of the *Staphylococcus aureus* heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J Bacteriol* 188(19):6739–6756
- Asaduzzaman M, Ryan ET, John M, Hang L, Khan AI, Faruque A, Taylor RK, Calderwood SB, Qadri F (2004) The major subunit of the toxin-coregulated pilus TcpA induces mucosal and systemic immunoglobulin A immune responses in patients with Cholera caused by *Vibrio cholerae* O1 and O139. *Infect Immun* 72(8):4448–4454
- Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res* 37(Suppl 1):D539–D543
- Beste DJV, Hooper T, Stewart G, Bonde B, Avignone-Rossa C, Bushell ME, Wheeler P, Klamt S, Kierzek AM, McFadden J (2007) GSMN-TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biol* 8(5):R89
- Blinov ML, Faeder JR, Goldstein B, Hlavacek WS (2004) BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics* 20(17):3289–3291
- Boele J, Olivier BG, Teusink B (2012) FAME, the Flux Analysis and Modeling Environment. *BMC Syst Biol* 6:8
- Bogard RW, Davies BW, Mekalanos JJ (2012) MetR-regulated *Vibrio cholerae* metabolism is required for virulence. *MBio* 3(5):e00236–12
- Bordbar A, Lewis NE, Schellenberger J, Palsson BØ, Jamshidi N (2010) Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol* 6(1):422
- Boshoff HIM, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE 3rd (2004) The transcriptional responses of *Mycobacterium tuberculosis* to inhibitors of metabolism novel insights into drug mechanisms of action. *J Biol Chem* 279(38):40174–40184
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd et al (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393(6685):537–544
- Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, Cheng TY, Moody DB, Murray M, Galagan JE (2009) Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. *PLoS Comput Biol* 5(8):e1000489
- Colvin J, Monine MI, Gutenkunst RN, Hlavacek WS, Von Hoff DD, Posner RG (2010) RuleMonkey: software for stochastic simulation of rule-based models. *BMC Bioinformatics* 11(1):404
- Covert MW, Xiao N, Chen TJ, Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* 24(18):2044–2050
- da Hora Junior BT, Poloni Jde F, Lopes MA, Dias CV, Gramacho KP, Schuster I, Sabau X, Cascardo JC, Mauro SM, Gesteira Ada S et al (2012) Transcriptomics and systems biology analysis in identification of specific pathways involved in cacao resistance and susceptibility to witches' broom disease. *Mol BioSyst* 8(5):1507–1519
- Dandekar T, Snel B, Huynen M, Bork P (1998) Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci* 23(9):324
- Danos V, Feret J, Fontana W, Harmer R, Krivine J (2008) Rule-based modelling, symmetries, refinements. *Form Methods Syst Biol* 5054:103–122

- Day J, Schlesinger LS, Friedman A (2010) Tuberculosis research: going forward with a powerful “translational systems biology” approach. *Tuberculosis* 90(1):7–8
- de Jong H (2002) Modeling and simulation of genetic regulatory systems: a literature review. *J Comput Biol* 9(1):67–103
- Dharia N, Chatterjee A, Winzeler E (2010) Genomics and systems biology in malaria drug discovery. *Curr Opin Investig Drugs* 11(2):131 (London, England: 2000)
- Eklund SE, Snider RM, Wikswa J, Baudenbacher F, Prokop A, Cliffel DE (2006) Multianalyte microphysiometry as a tool in metabolomics and systems biology. *J Electroanal Chem* 587(2):333–339
- Feist AM, Herrgård MJ, Thiele I, Reed JL, Palsson BØ (2008) Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol* 7(2):129–143
- Forst CV (2006) Host-pathogen systems biology. *Drug Discov Today* 11(5–6):220–227
- Funahashi A, Morohashi M, Kitano H, Tanimura N (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico* 1(5):159–162
- Galagan JE, Sisk P, Stolte C, Weiner B, Koehrsen M, Wymore F, Reddy TB, Zucker JD, Engels R, Gellesch M et al (2010) TB database 2010: overview and update. *Tuberculosis* 90(4):225–235
- Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S et al (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419(6906):498–511
- Ghosh S, Prasad KVS, Vishveshwara S, Chandra N (2011) Rule-based modelling of iron homeostasis in tuberculosis. *Mol Biosyst* 7(10):2750–2768
- Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81(25):2340–2361
- Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, Florens L, Janssen CS, Pain A, Christophides GK et al (2005) A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307(5706):82–86
- Hill DR, Ford L, Lalloo DG (2006) Oral cholera vaccines: use in clinical practice. *Lancet Infect Dis* 6(6):361–373
- Hlavacek WS, Faeder JR (2009) The complexity of cell signaling and the need for a new mechanics. *Sci Signal* 2(81):pe46
- Hood L, Perlmuter RM (2004) The impact of systems approaches on biological problems in drug discovery. *Nat Biotechnol* 22(10):1215–1217
- Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, Xu L, Mendes P, Kummer U (2006) COPASI—a complex pathway simulator. *Bioinformatics* 22(24):3067–3074
- Hoppe AD, Seveau S, Swanson JA (2009) Live cell fluorescence microscopy to study microbial pathogenesis. *Cell Microbiol* 11(4):540–550
- Hoppe A, Hoffmann S, Gerasch A, Gille C, Holzhütter HG (2011) FASIMU: flexible software for flux-balance computation series in large metabolic networks. *BMC Bioinformatics* 12(1):28
- Hung DT, Shakhnovich EA, Pierson E, Mekalanos JJ (2005) Small-molecule inhibitor of *Vibrio cholerae* virulence and intestinal colonization. *Science* 310(5748):670–674
- Huthmacher C, Hoppe A, Bulik S, Holzhütter HG (2010) Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst Biol* 4:120
- Ibargüen-Mondragón E, Esteve L, Chávez-Galán L (2011) A mathematical model for cellular immunology of tuberculosis. *J Math Biosci Eng* 8(4):973–986
- Ideker T, Ozier O, Schwikowski B, Siegel AF (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18(suppl 1):S233–S240
- Jamshidi N, Palsson B (2007) Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. *BMC Syst Biol* 1(1):26
- Johanns TM, Ertelt JM, Rowe JH, Way SS (2010) Regulatory T cell suppressive potency dictates the balance between bacterial proliferation and clearance during persistent *Salmonella* infection. *PLoS Pathog* 6(8):e1001043

- Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, Hintz M, Türbachova I, Eberl M, Zeidler J, Lichtenthaler HK (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* 285(5433):1573–1576
- Kaatz GW, McAleese F, Seo SM (2005) Multidrug resistance in *Staphylococcus aureus* due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. *Antimicrob Agents Chemother* 49(5):1857–1864
- Kanjilal S, Citorik R, LaRocque RC, Ramoni MF, Calderwood SB (2010) A systems biology approach to modeling *Vibrio cholerae* gene expression under virulence-inducing conditions. *J Bacteriol* 192(17):4300–4310
- Kaper JB, Morris JG, Jr, Levine MM (1995) Cholera. *Clin Microbiol Rev* 8(1):48–86
- Kirschner DE, Young D, Flynn JAL (2010) Tuberculosis: global approaches to a global disease. *Curr Opin Biotechnol* 21(4):524–531
- Kitano H (2002) Systems biology: a brief overview. *Science* 295(5560):1662–1664
- Konjufca V, Miller MJ (2009) Two-photon microscopy of host–pathogen interactions: acquiring a dynamic picture of infection in vivo. *Cell Microbiol* 11(4):551–559
- Kruh NA, Trout J, Izzo A, Prenni J, Dobos KM (2010) Portrait of a pathogen: the *Mycobacterium tuberculosis* proteome in vivo. *PLoS One* 5(11):e13938
- Kumar D, Nath L, Kamal MA, Varshney A, Jain A, Singh S, Rao KVS (2010) Genome-wide analysis of the host intracellular network that regulates survival of *Mycobacterium tuberculosis*. *Cell* 140(5):731–743
- Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K (2003) Two-component system *VraSR* positively modulates the regulation of cell-wall biosynthesis pathway in *Staphylococcus aureus*. *Mol Microbiol* 49(3):807–821
- Lakshmanan M, Koh G, Chung BK, Lee DY (2014) Software applications for flux balance analysis. *Brief Bioinform* 15(1):108–122
- Latendresse M, Krummenacker M, Trupp M, Karp PD (2012) Construction and completion of flux balance models from pathway databases. *Bioinformatics* 28(3):388–396
- Lew JM, Kapopoulou A, Jones LM, Cole ST (2011) TubercuList—10 years after. *Tuberculosis* 91(1):1–7
- MacPherson DW, Gushulak BD, Baine WB, Bala S, Gubbins PO, Holtom P, Segarra-Newnham M (2009) Population mobility, globalization, and antimicrobial drug resistance. *Emerg Infect Dis* 15(11):1727–1732
- Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D (1999) Detecting protein function and protein–protein interactions from genome sequences. *Science* 285(5428):751–753
- Marino S, El-Kebir M, Kirschner D (2011) A hybrid multi-compartment model of granuloma formation and T cell priming in tuberculosis. *J Theor Biol* 280(1):50–62
- Maus C, Rybacki S, Uhrmacher A (2011) Rule-based multi-level modeling of cell biological systems. *BMC Syst Biol* 5(1):166
- McNicholl JM, Downer MV, Udhayakumar V, Alper CA, Swerdlow DL (2000) Host–pathogen interactions in emerging and re-emerging infectious diseases: a genomic perspective of tuberculosis, malaria, human immunodeficiency virus infection, hepatitis B, and cholera. *Annu Rev Public Health* 21:15–46
- Nelson EJ, Harris JB, Morris JG Jr, Calderwood SB, Camilli A (2009) Cholera transmission: the host, pathogen and bacteriophage dynamic. *Nat Rev Microbiol* 7(10):693–702
- Olszewski KL, Morrissy JM, Wilinski D, Burns JM, Vaidya AB, Rabinowitz JD, Llinás M (2009) Host–parasite interactions revealed by *Plasmodium falciparum* metabolomics. *Cell Host Microbe* 5(2):191–199
- Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? *Nat Biotechnol* 28(3):245–248
- Padiadpu J, Vashisht R, Chandra N (2010) Protein–protein interaction networks suggest different targets have different propensities for triggering drug resistance. *Syst Synth Biol* 4(4):311–322
- Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO (1999) Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci U S A* 96(8):4285–4288

- Pieters J (2008) *Mycobacterium tuberculosis* and the macrophage: maintaining a balance. *Cell Host Microbe* 3(6):399–407
- Pinney JW, Westhead DR, McConkey GA (2003) Petri Net representations in systems biology. *Biochem Soc Trans* 31(Pt 6):1513–1515
- Plata G, Hsiao TL, Olszewski KL, Llinas M, Vitkup D (2010) Reconstruction and flux-balance analysis of the *Plasmodium falciparum* metabolic network. *Mol Syst Biol* 6:408
- Plikat U, Voshol H, Dangendorf Y, Wiedmann B, Devay P, Muller D, Wirth U, Szustakowski J, Chirn GW, Inverardi B et al (2007) From proteomics to systems biology of bacterial pathogens: approaches, tools, and applications. *Proteomics* 7(6):992–1003
- Rachman H, Strong M, Schaible U, Schuchhardt J, Hagens K, Mollenkopf H, Eisenberg D, Kaufmann SHE (2006) *Mycobacterium tuberculosis* gene expression profiling within the context of protein networks. *Microbes Infect* 8(3):747–757
- Raman K, Chandra N (2009) Flux balance analysis of biological systems: applications and challenges. *Brief Bioinform* 10(4):435–449
- Raman K, Bhat AG, Chandra N (2010) A systems perspective of host–pathogen interactions: predicting disease outcome in tuberculosis. *Mol Biosyst* 6(3):516–530
- Raman K, Rajagopalan P, Chandra N (2005) Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS Comput Biol* 1(5):e46
- Raman K, Rajagopalan P, Chandra N (2006) Principles and practices of pathway modelling. *Curr Bioinform* 1(2):147–160
- Raman K, Yeturu K, Chandra N (2008) targetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol* 2(1):109
- Rappuoli R, Aderem A (2011) A 2020 vision for vaccines against HIV, tuberculosis and malaria. *Nature* 473(7348):463–469
- Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL (2002) Hierarchical organization of modularity in metabolic networks. *Science* 297(5586):1551–1555
- Sasseti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48(1):77–84
- Sauro H (2004) An introduction to biochemical modeling using JDesigner. Keck Graduate Institute, Claremont
- Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2. 0. *Nat Protoc* 6(9):1290–1307
- Schlatter R, Schmich K, Vizcarra IA, Scheurich P, Sauter T, Borner C, Ederer M, Merfort I, Sawodny O (2009) ON/OFF and beyond-a boolean model of apoptosis. *PLoS Comput Biol* 5(12):e1000595
- Schluger NW, Rom WN (1998) The host immune response to tuberculosis. *Am J Respir Crit Care Med* 157(3):679–691
- Schmidt F, Donahoe S, Hagens K, Mattow J, Schaible UE, Kaufmann SHE, Aebersold R, Jungblut PR (2004) Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol Cell Proteomics* 3(1):24–42
- Segre D, Vitkup D, Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* 99(23):15112–15117
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504
- Shin S, Desai SN, Sah BK, Clemens JD (2011) Oral vaccines against cholera. *Clin Infect Dis* 52(11):1343–1349
- Siek JG, Lee LQ, Lumsdaine A (2002) The boost graph library: user guide and reference manual. Addison-Wesley Longman Publishing Co., Inc. Boston

- Singh VK, Ghosh I (2006) Kinetic modeling of tricarboxylic acid cycle and glyoxylate bypass in *Mycobacterium tuberculosis*, and its application to assessment of drug targets. *Theor Biol Med Modelling* 3(1):27
- Spagnuolo AM, Dirita V, Kirschner D (2011) A model for *Vibrio cholerae* colonization of the human intestine. *J Theor Biol* 289:247–258
- Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM, Bartlett JG, Edwards J Jr (2008) The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. *Clin Infect Dis* 46(2):155–164
- Stoll G, Viara E, Barillot E, Calzone L (2012) Continuous time Boolean modeling for biological signaling: application of Gillespie algorithm. *BMC Syst Biol* 6(1):116
- Strong M, Graeber TG, Beeby M, Pellegrini M, Thompson MJ, Yeates TO, Eisenberg D (2003) Visualization and interpretation of protein networks in *Mycobacterium tuberculosis* based on hierarchical clustering of genome-wide functional linkage maps. *Nucleic Acids Res* 31(24):7099–7109
- Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 39(Suppl 1):D561–D568
- Trautmann L, Sekaly RP (2011) Solving vaccine mysteries: a systems biology perspective. *Nat Immunol* 12(8):729–731
- Tucker PA, Nowak E, Morth JP (2007) Two-component systems of *Mycobacterium tuberculosis*—structure-based approaches. *Methods Enzymol* 423:477–501
- Tyagi JS, Sharma D (2004) Signal transduction systems of mycobacteria with special reference to *M. tuberculosis*. *Curr Sci* 86(1):93–102
- Vashisht R, Mondal AK, Jain A, Shah A, Vishnoi P, Priyadarshini P, Bhattacharyya K, Rohira H, Bhat AG, Passi A et al (2012) Crowd sourcing a new paradigm for interactome driven drug target identification in *Mycobacterium tuberculosis*. *PLoS One* 7(7):e39808
- Vodovotz Y, Csete M, Bartels J, Chang S, An G (2008) Translational systems biology of inflammation. *PLoS Comput Biol* 4(4):e1000014
- Wang RS, Albert R (2011) Elementary signaling modes predict the essentiality of signal transduction network components. *BMC Syst Biol* 5(1):44
- Wang JH, Byun J, Pennathur S (2010a) Analytical approaches to metabolomics and applications to systems biology. *Seminars Nephrol* 30(5):500–511
- Wang Y, Cui T, Zhang C, Yang M, Huang Y, Li W, Zhang L, Gao C, He Y, Li Y (2010b) Global protein–protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *J Proteome Res* 9(12):6665–6677
- Weckwerth W (2003) Metabolomics in systems biology. *Annu Rev Plant Biol* 54:669–689
- Weston AD, Hood L (2004) Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. *J Proteome Res* 3(2):179–196
- WHO (2012) Fact sheets: infectious diseases. World Health Organization
- Winzeler EA (2005) Applied systems biology and malaria. *Nat Rev Microbiol* 4(2):145–151
- Wittmann DM, Krumsiek J, Saez-Rodriguez J, Lauffenburger DA, Klamt S, Theis FJ (2009) Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC Syst Biol* 3(1):98