Systems biology

Raman, Karthik ; Chandra, Nagasuma

Abstract: Systems biology seeks to study biological systems as a whole, contrary to the reductionist approach that has dominated biology. Such a view of biological systems emanating from strong foundations of molecular level understanding of the individual components in terms of their form, function and interactions is promising to transform the level at which we understand biology. Systems are defined and abstracted at different levels, which are simulated and analysed using different types of mathematical and computational techniques. Insights obtained from systems level studies readily lend to their use in several applications in biotechnology and drug discovery, making it even more important to study systems as a whole

DOI: https://doi.org/10.1007/s12045-010-0015-7

Systems Biology

Karthik Raman and Nagasuma Chandra

Systems biology seeks to study biological systems as a whole, contrary to the reductionist approach that has dominated biology. Such a view of biological systems emanating from strong foundations of molecular level understanding of the individual components in terms of their form, function and interactions is promising to transform the level at which we understand biology. Systems are defined and abstracted at different levels, which are simulated and analysed using different types of mathematical and computational techniques. Insights obtained from systems level studies readily lend to their use in several applications in biotechnology and drug discovery, making it even more important to study systems as a whole.

1. Introduction

Biological systems are enormously complex, organised across several levels of hierarchy. At the core of this organisation is the genome that contains information in a digital form to make thousands of different molecules and drive various biological processes. This genomic view of biology has been primarily ushered in by the human genome project. The development of sequencing and other high-throughput technologies that generate vast amounts of biological data has fuelled the development of new ways of hypothesis-driven research. Development of computational techniques for analysis of the large data, as well as for the modelling and simulation of the complex biological systems have followed as a logical consequence. Simulatable computational models of biological systems and processes form the cornerstone of the emerging science of *systems biology*.

Traditionally, biology has focused on identifying individual genes, proteins and cells, and studying their specific functions. Each of

Karthik Raman recently completed his PhD in computational systems biology from IISc, Bangalore. He is currently a postdoctoral research associate in the Department of Biochemistry, at the University of Zurich. His research interests include the modelling of complex biological networks and the analysis of their robustness and evolvability.

Nagasuma Chandra obtained her PhD in structural biology from the University of Bristol, UK. She serves on the faculty of Bioinformatics at IISc, Bangalore. Her current research interests are in computational systems biology, cell modeling and structural bioinformatics and in applying these to address fundamental issues in drug discovery.

Keywords
Pathway modelling, systems modelling, biological networks, complex systems.

To study an aircraft, focused detailed studies on individual components such as the engine, wings and tail, would not be sufficient to understand how an aircraft can fly.

these is indeed extremely important in understanding the individual molecules, but as individual isolated pieces of information, they are insufficient to provide insights about complex phenomena such as human health and disease. As an analogy, to study an aircraft, focused detailed studies on individual components such as the engine, wings and tail, would not be sufficient to understand how an aircraft can fly. More importantly, it would not provide any understanding of what component influences what other component in what manner and to what extent, an understanding which is very important to effectively set things right when something malfunctions. In the same way, since diseases occur when there is some malfunction in the form or function of one or more of the cellular components, we need an understanding how various molecules in a cell influence each other in health, in order to attempt curing or correcting it to the extent possible.

The scale at which various molecular level studies can now be carried out is providing us systematic data on many fronts enabling us to reconstruct holistic models of larger systems [1]. Systems biology seeks to study biochemical and biological systems from a holistic perspective, promising to transform how biology is done. The goal is for a comprehensive understanding of the system's influence on its individual components, leading to the appearance of complex properties such as robustness, emergence, adaptation, regulation and synchronisation, seen so very often in biological systems. Essentially, systems biology advocates a departure from the reductionist viewpoint, emphasising on the importance of a holistic view of biological systems. It also aims at a departure from the "spherical cow" [1], in trying to encapsulate the enormous complexity of biological systems in greater detail. Systems biology adopts an integrated approach to study and understand the function of biological systems, particularly, the response of such systems to perturbations such as the inhibition of a reaction in a pathway, or the administration of a drug. It can of course be argued that systems biology is just a new name for the conventional disciplines such as physiology and

[1]A euphemism often directed at the severe approximations that characterise modelling.

pharmacology, which are well established for several decades now. Undoubtedly, these disciplines emphasise the need for considering whole systems. Yet, systems biology emerges as a new discipline, since it differs from the conventional disciplines in a fundamental way: the latter treat much of the whole system as a 'black-box', giving us only an idea of the end picture but not enabling us to ask 'why' or 'how' a particular outcome is seen. Systems biology on the other hand aims to reconstruct systems by a bottom-up approach, with detailed knowledge about the individual components that make up the system and how these components interact with each other. Modelling and simulation of complex biological networks form the cornerstone of systems biology; the coupling of *in silico* models with *in vivo* and *in vitro* experimentation, with modelling guiding experimentation and experimentation aiding in model refinement, can provide impetus to improve the understanding of biological systems. Effects and influences of one component on the other are deciphered, providing a greater understanding of how genotypes relate to phenotypes.

Systems biology aims to reconstruct systems by a bottom-up approach, with detailed knowledge about the individual components that make up the system and how these components interact with each other.

2. Elements of Systems Biology

Systems biology, being a holistic approach involves modelling and analysis of metabolic pathways, regulatory and signal transduction networks for understanding cellular behaviour. There are also various levels of abstraction at which these systems are modelled, with a wide variety of techniques that can be employed based on the quality and quantity of data available.

The critical step in the modelling and analysis of these pathways is their reconstruction, involving the integration of diverse sources of data to create a representation of the chemical events underlying biological networks [2]. A variety of high-throughput experiments have been developed to provide extensive data on the proteome, metabolome, transcriptome and the reactome in a cell (see *Box* 1 for glossary of terms). Some of these techniques include microarray analyses of the transcriptome and mass spectrometry analyses that generate proteomics data. It is important to

Box 1. Glossary of Some Terms Used and their Related Concepts

Genomics and other 'omics': The German botanist Hans Winkler coined the term 'genome' in 1920 by combining the words GENe and chromosOME. A precise definition of genome is 'all the DNA in a cell' because this includes not only genes but also DNA that is not part of a gene, or non-coding DNA or in other words, all the genetic material in the chromosomes of a particular organism'; its size is generally given as its total number of base pairs. In the same way, a proteome is a collection of all proteins, coded by a genome. The human proteome is the collection of proteins found in the human body. With the success of large-scale quantitative biology projects such as genome sequencing, the suffix '-ome-' has migrated to a host of other contexts: Omics such as genomics, transcriptomics, proteomics, metabolomics or interactomics. The suffix 'ome' often signifies totality of some sort.

Genome Sequence refers to the sequence of consecutive DNA 'letters' spanning all the chromosomes of a cell from start to finish.

Gene Expression: This refers to the 'turning on' of a gene. Most human genes are active, or turned on, only in certain cells under certain conditions. Genes for eye colour are active in eye cells but not in stomach cells. Similarly, some genes may lie dormant for years and then turn on and become malignant late in life. Transcription is the process of 'turning on', or activating a gene.

Genotype and Phenotype: Genotype refers to the particular form of a gene a person has. The genetic constitution of an organism, as distinguished from its physical appearance (its phenotype). A phenotype on the other hand is the physical trait such as red hair, or behaviour such as anxiety. A phenotype results from the 'expression' of a gene or genes.

Network: A network is generally a collection of related nodes, connected on the basis of interactions. It is essentially a map of 'who interacts with whom'. The most common kinds of networks are protein–protein interaction networks, cataloguing how different proteins interact with one another within the cell.

Hub: Much like the hub of a wheel, highly connected nodes in networks are referred to as hubs. These often are very important to network function and at the same time are the 'Achilles heels' of many real-world networks. In the context of biological networks, hubs may present interesting drug targets.

Data-driven Modelling: Traditional physical modelling is knowledge-driven, where knowledge is derived from first principles; for example, the modelling of the motion of a pendulum based on ODEs from physical laws. Data-driven modelling, in contrast, is employed when certain input–output characteristics of the system are known but there is only limited knowledge on the system in question. Linear regression is a very simple example of a data-driven model.

understand that these experiments generate genome-scale 'omics' data, which cover a majority of the components such as metabolites, transcripts and proteins, in a cell. Another major feature of systems biology is the strong integration of experiment with theory; it is quite common that a model is used to generate one or

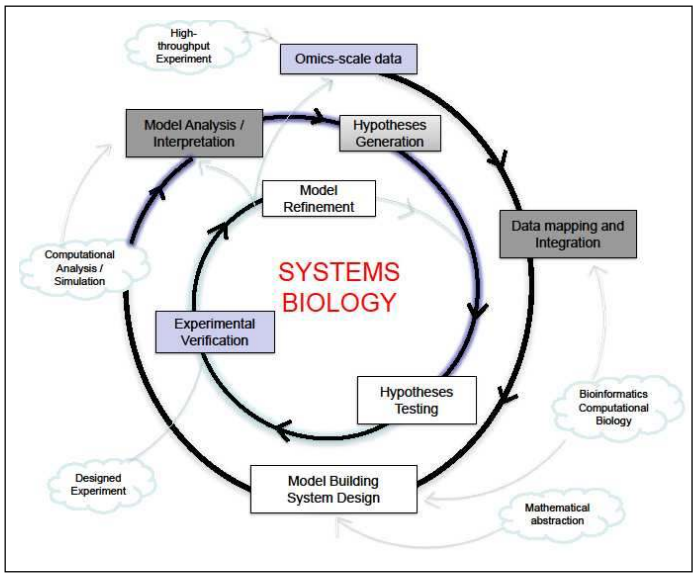

Figure 1. Systems biology process. This process relies on an iterative procedure of model building, experimental verification, model analysis and model refinement. The concepts that underlie these processes have been shown as clouds.

more hypotheses, which are then tested experimentally, and iteratively contribute to model refinement. In essence, the various parts of a systems biology study are (a) define a model system, (b) identify a choice of attributes/parameters to study the system that is appropriate for the problem being addressed, (c) comprehensive experimental measurements, (d) appropriate mathematical abstraction of the system that is computationally tractable and (e) computational simulations that can generate and test various hypotheses, (f) that can later be verified by experimental approaches (*Figure* 1).

2.1 *System Definition*

Systems biology experiments are often characterised by a synergy between theory and experiment (*Figure* 1). As in traditional biological experiments, the chosen model system must be suitable for experimental investigations, and should also be complex enough to capture the biological phenomenon of interest. Simple bacteria such as *E. coli* are often used as model organisms to understand the organisation and behaviour of prokaryotic systems. *Saccharomyces cerevisiae* is the *de facto* standard model organism for understanding eukaryotic systems. Similarly, the fruitfly *Drosophila* and the worm *Caenorhabditis elegans* are

Pathways or sets of pathways function as modules of the larger systems, which provide a practical framework to study the biological processes/ phenomena.

used as models to incrementally understand more and more complex multi-cellular organisms. It is important to note that, although some of these systems are significantly less complex than mammalian systems, several processes are conserved, leading to the possibility of very useful predictions of the behaviour of mammalian systems from the modelling of simpler systems. Very often, it is impractical to consider whole organisms or whole cells, especially to address questions pertaining to the mechanism of a given process. Pathways or sets of pathways function as modules of the larger systems, which provide a practical framework to study the biological processes/phenomena.

Metabolic pathways, signal transduction pathways and regulatory pathways have been studied from a variety of organisms from which a wide range of biological insights have been obtained. Such pathways have also been combined into the larger context of networks, where the abstraction is often a bit less quantitative, again for practical reasons. Studies on transcriptional network of yeast, metabolic network of *E. coli*, serve as examples of studies at this level. Thus, the scale of the system can vary from tens of components to several thousands. The resolution of information can also vary from detailed atomistic information to broad cellular views. For example, a defined model could contain thermodynamic data of the metabolic reactions in a given pathway, that in turn have sound correlations with the three-dimensional structures of the involved enzymes. On the other hand the defined system could simply contain logical connections between different cellular states implying functional correlations without any further details on the cells themselves. A systems biology approach is characterised by a series of iterative experimentation and model refinement (also see *Figure* 1), often using perturbations to the system as a handle to affirm roles of known components as well as to discriminate between alternative models [3]. Another important feature is that most of these components, from computation to high-throughput laboratory experiments are amenable to automation.

2.2 *Modelling in Systems Biology: Model Abstraction*

Models are created to simulate a process or a set of processes observed in the natural world in order to gain insights into process mechanisms and predict outcomes for a given set of specific input parameters. Conceptual and theoretical modelling constructs are expressed as sets of algorithms and implemented as software packages. What constitutes a model depends upon what is understood about a given process and how best it is computationally tractable. For example, in drug discovery, a model can refer to the relationship of the structure of a target molecule to its ability to bind a certain type of ligand at one end of the spectrum, while at the other end, it can refer to a statistically derived relationship of a set of ligands to a particular biological activity, with no explicit consideration of the mechanism or the basis of such activities. Conceptual modelling is an integral part of problem solving in general and in fact an essential component of any activity that attempts to achieve a goal in a systematic way.

Models are created to simulate a process or a set of processes observed in the natural world in order to gain insights into process mechanisms and predict outcomes for a given set of specific input parameters.

The advantages of having a model are manifold: (a) it gives the precise definition of the components of a given system (or the genotype), (b) it allows performing simulations and monitoring the end-effect, which may be the given phenotype in this context, (c) it helps in dissecting the role of every component in the system through the analysis of perturbations, (d) it helps us to interpret complex hard-to-understand problems, (e) it helps in studying systems that are impractical to study through conventional experiments, (f) it helps both in designing minimal systems that can result in a particular phenotype, as well as analysing the effect of the addition of newer components into the framework, and (g) it is highly amenable for high-throughput simulations and highly cost-effective, useful especially in applications such as drug discovery.

Thus, models not only provide significant insights into the underlying biology and application opportunities, but also enable the efficient study of what may be highly impractical, or even impossible through biochemical and molecular biology experiments. It

What constitutes a model depends upon what is understood about a given process and how best it is computationally tractable.

must however be emphasised that a model is only as good as our understanding of what constitutes a system and how it has been built. Model building is thus a critical step in *in silico* analysis and is often iterated and refined with validation steps.

Given that biological systems and processes are understood at many different levels and in many different aspects, it is no wonder that many different kinds of models should exist in practice. *Figure* 2 illustrates that models span a wide range, emanating from the organisational hierarchy in which biological systems are understood.

On one hand, there are structural models at atomic levels implying certain functions, whereas on the other hand, there are whole genome-based mathematical models of either pathways or entire organisms implying functions at very different levels. It is

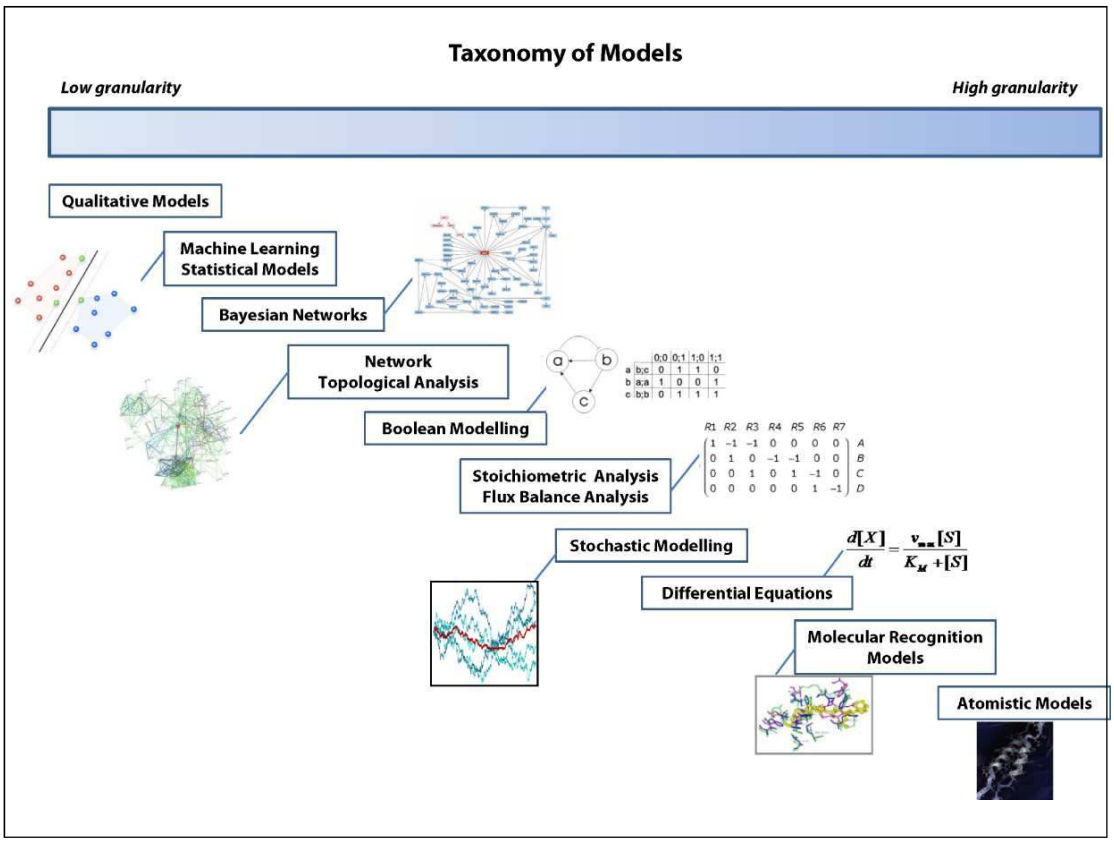

Figure 2. Modelling techniques in systems biology. The various methods have been represented alongside an axis that details the granularity (or resolution) typical for each method.

important to understand the abstraction levels of the models, so that conclusions are drawn at appropriate levels from the analyses. The choice of the method depends upon the type and extent of data available, as well as the objective of the modelling exercise in terms of the level at which the system is desired to be understood.

Models are routinely built from a variety of sources, which vary in the degree of accuracy of experiments recorded and often depend even on the interpretation of data available. Model validation is a critical quality control step that endorses the results obtained through simulation of the model. Typical model validation involves the comparison of model predictions against known biochemical and genetic evidences obtained by various experiments, particularly when experimental data has not been used for tuning the models.

Models are routinely built from a variety of sources, which vary in the degree of accuracy of experiments recorded and often depend even on the interpretation of data available.

3. Key Properties of Biological Systems/Models

Biological systems are characterised by several key properties, which distinguish them from models in other disciplines. Knowledge of these fundamental principles, which characterise biological systems, is important both for understanding their function and for modelling them. Some of these interesting properties are discussed below.

3.1 *Irreducibility*

Irreducibility is an important concept that makes systems thinking important. We may undoubtedly gain significant insight into each of the components of the system by studying them individually, but we will require to study the system as a whole in order to gain a holistic perspective of what these components do when they are all put together in an appropriate manner. An analogy to a book is often drawn, where one cannot understand a book by reading one word at a time. In other words, knowing the meaning of every word in a book does not tell us what the book is about. They have to be placed in context to grasp the story in the book. It is this context that is sought out in systems biology; thus it is not

The larger properties that arise out of that context are often referred to as 'emergent properties'. Emergent properties are thus consequences of the interactions between system components.

In general, the behaviour of a system is quite different from merely the sum of the functions of its various parts.

only the form and function of individual molecules, but rather their functional orchestration in a 'context' in a complex manner that makes a living species. The larger properties that arise out of that context are often referred to as 'emergent properties'. Emergent properties are thus consequences of the interactions between system components.

3.2 *Emergence*

Systems are composed of individual elements or 'parts' that interact in various ways. In general, the behaviour of a system is quite different from merely the sum of the functions of its various parts. As Anderson put it as early as 1972, in his classic paper by the same title, "*More is different*" [4], it is not possible to reliably predict the behaviour of a complex system, despite a good knowledge of the fundamental laws governing the individual components:

The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. The constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity. At each level of complexity entirely new properties appear. Psychology is not applied biology, nor is biology applied chemistry. We can now see that the whole becomes not merely more, but very different from the sum of its parts.

P W Anderson, 1972 [5]

This reinforces the need to develop methods to study biological systems at the systems level, rather than at the level of individual components.

3.3 *Complexity*

The term complexity**,** a concept linked to the concept of systems itself, is often used in a variety of disciplines to characterise a system with a number of components intricately linked to each other, giving rise to behaviours that may not be described by

simple models. Emergent behaviour (described above) is one of the most fundamental features of complex systems. Health or disease are examples of complex systems. These cannot be predicted simply by analyzing the individual ligands or proteins that comprise the cells. A more complete picture of their context would be necessary to achieve it. Biological systems, needless to say, are extraordinarily complex, which is evident in individual prokaryotic cells, let alone multi-cellular organisms. For example an *E. coli* cell has about 4500 genes coding for at least as many proteins. At the outset, trying to understand how these many proteins embedded in less than femtolitre (10^{-15} L) of volume, perform together to enable the many functions of the *E. coli* cell appears to be a daunting task. However, we can understand many aspects of the cell if we divide complexity in hierarchies and then focus on understanding each individual level in a stepwise manner and then re-assimilate them in an appropriate context. In other words, understanding how, at the bottom of the hierarchy, DNA, proteins and metabolites function in individual cells, helps us to understand how different proteins give rise to pathways, how pathways come together to form processes, and how these are organised in a functional cell (*Figure* 3). In higher organisms, we

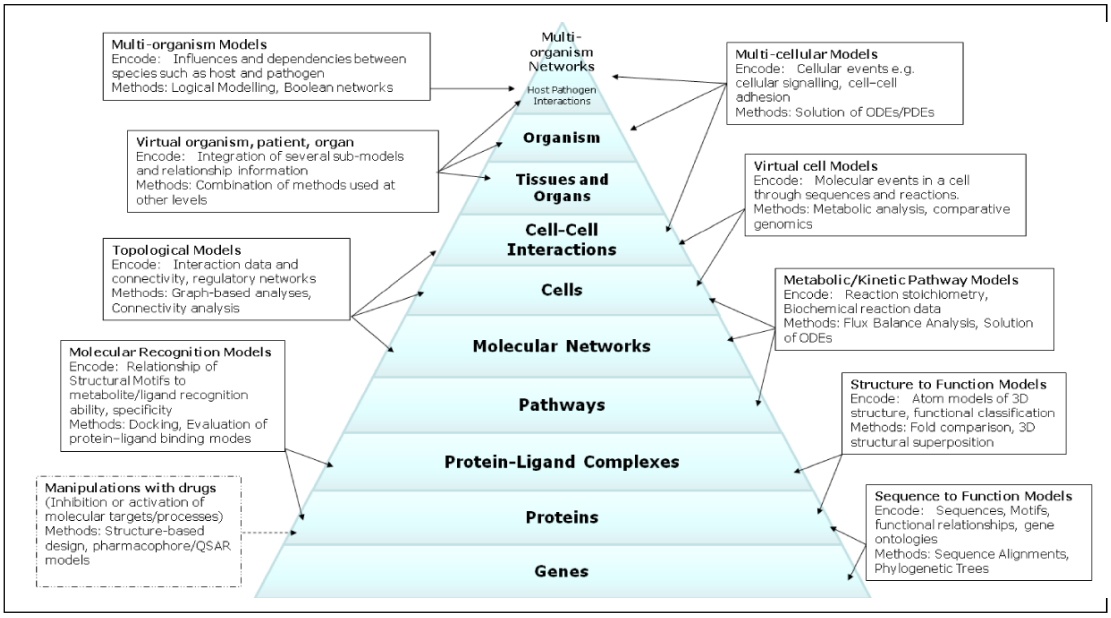

Figure 3. Levels of hierarchies for understanding and modelling biological systems. The figure illustrates different types of models that are appropriate at a given level of hierarchy. The information they encode (abstraction level) are listed for each of them as also the methods that are in current practice to design, build and analyse the models.

would extend this to understanding how different cell types are formed and organised in tissues, how tissues make complex organs, and finally, how many different organs are orchestrated in the top hierarchical level, the organism. However, amidst this complexity, there is modularity with many common mechanisms for a range of biological events. Different cell types and functions use recurrent basic mechanisms of organisation and communication; thus common patterns underlie diverse expressions of life. Understanding single cell types, even when the organisms containing them are evolutionarily distant, such as bacteria and humans, would inevitably provide enormous amount of information to understand other cell types. Complexity has important implications for modelling; the complexity of large systems often makes them intractable for analyses. Therefore, large systems are often broken down into their constituent modules, or sub-systems, which are more amenable for analyses.

3.4 *Modularity*

Bacterial systems may be viewed as comprising modules, which may be insulated or decoupled from one another, or alternatively, connected to one another. Modules constitute semi-autonomous entities with dense internal functional connections and relatively looser external connections with their environment. Modularity or the encapsulation of functions, can contribute to both robustness (by confinement of damage) and to evolvability (by rewiring of modules for new functionality) [5]. An obvious example of a module is a cell in a multi-cellular organism, which interacts with both the environment and other cells. Modules are also commonly organised in a hierarchical fashion: a cell is composed of organelles, while also being a part of higher structures such as tissues and organs (*Figure* 3) . At a different level, a signal transduction system is an extended module that achieves isolation on account of the specificity of the binding of chemical signals to receptor proteins as well as the specificity of the interactions between the signalling proteins within the cell [5].

Bacterial systems may be viewed as comprising modules, which may be insulated or decoupled from one another, or alternatively, connected to one another.

3.5 *Robustness and Fragility*

Robustness may be understood as the relative property of a system to retain stability despite several perturbations, internal or external. No system can be robust to all kinds of perturbations. Robustness in biological systems is achieved using several complex mechanisms involving feedback, alternative (fail-safe) mechanisms featuring redundancy and diversity (heterogeneity), structuring of complex systems into semi-autonomous functional units (modularity), and their reliable co-ordination via establishment of hierarchies and protocols. Sensitivity or fragility, however, characterises the ability of organisms to respond adequately to a stimulus. Robustness and fragility have been described in the literature as inseparable; the '*robust, yet fragile*' nature of complex systems is thought to exhibit 'highly optimised tolerance'.

Robustness may be understood as the relative property of a system to retain stability despite several perturbations.

Complex engineered systems (and biological systems) are often quite resistant to designed-for uncertainties, but quite susceptible to other perturbations. For example, modern aeroplanes, vis-à-vis the Wright brothers' aeroplane, are quite stable to atmospheric perturbations, but are fundamentally sensitive to complete electrical failure, due to the tight dependence of the control on a wide variety of electrical systems. Several biological systems are quite sensitive to what may be quantitatively small perturbations. There are several examples of networks which exhibit high insensitivity to attacks on nodes in random, but high sensitivity showing high disruption when there is a targeted attack on a few highly connected (hub) nodes.

4. Practice of Systems Biology

Systems biology primarily involves the building of models of systems, detailing metabolism, regulation, signalling and protein–protein interactions (*Figur*e 4). A variety of modelling techniques encompassing a wide spectrum of resolution and accuracy are used. *Figure* 2 shows some of these methods, also indicating the level of detail that the method usually deals with. The levels of biological organisational hierarchy at which such methods can

Systems biology primarily involves the building of models of systems, detailing metabolism, regulation, signalling and protein–protein interactions.

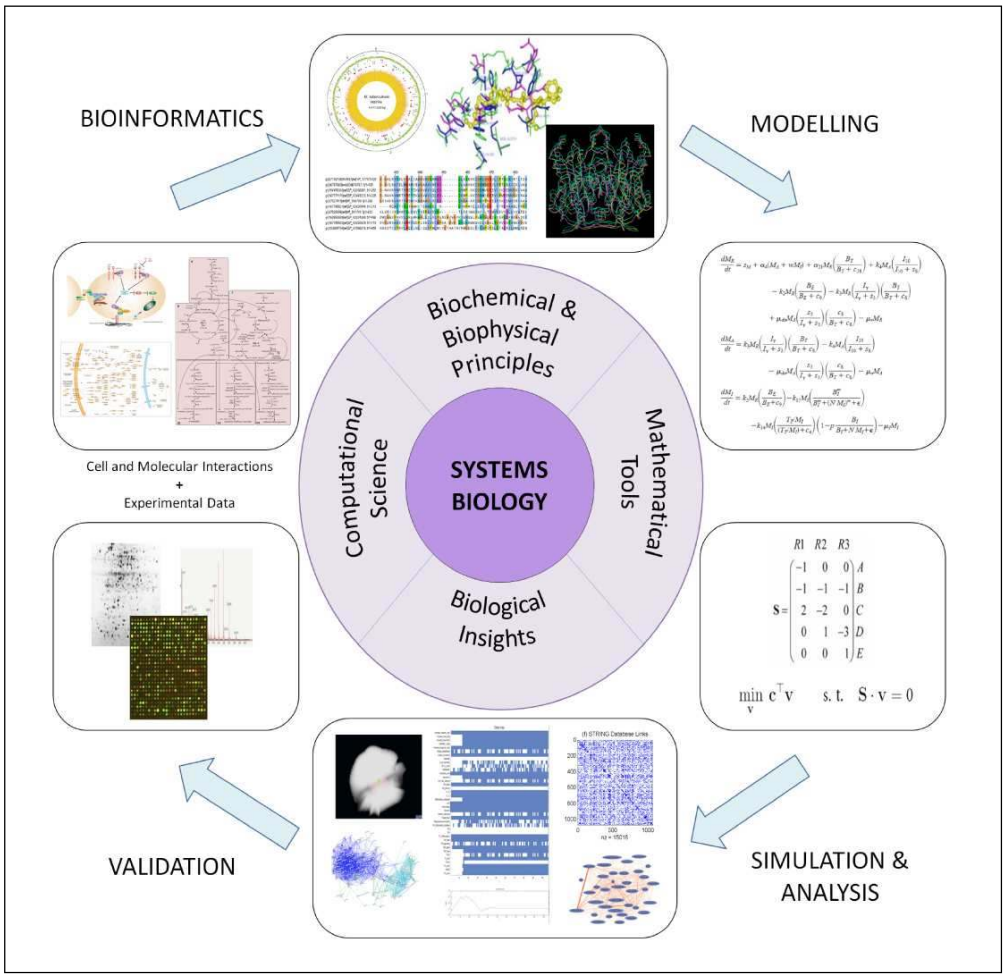

Figure 4. An overview of modelling in systems biology. This figure illustrates the various components of the systems biology modelling cycle, of how various types of experimental data are translated to a mathematical model, followed by simulation. The simulation results are then used to infer predictions (system behaviour), which are often compared against experimental results, leading to further improvements/enhancements to the mathematical model.

be used have already been illustrated in *Figure* 3. Some of the tools and resources useful for systems level modelling and simulation of biological systems are listed in *Table* 1.

At the highest level of resolution, there are atomistic models, followed by molecular recognition models, incorporating details at the lowest atomic level. These are followed by mechanistic models of molecular networks, which are usually realised using

Models for Simulation	
Biomodels.net	http://www.biomodels.org/
CellML	http://www.cellml.org/
Panther	http://www.pantherdb.org/
Pathway Databases	
BioCyc	http://biocyc.org/
BioChemWeb	http://www.biochemweb.org/
KEGG Pathway	http://www.genome.jp/
Reactome	http://www.reactome.org/
Quantitative Data	
BioPAX	http://www.biopax.org/
BRENDA	http://www.brenda-enzymes.info/
Systems Biology Standards	
LibSBML (API)	http://sbml.org/
MathML	http://sbml.org/
SBML	http://sbml.org/
little b	http://www.littleb.org/
MIRIAM	http://www.biomodels.org/
Pathway Design and Network Based Tools	
Cell Designer	http://www.celldesigner.org/
Cytoscape	http://cytoscape.org/
JDesigner	http://www.sys-bio.org/software/jdesigner.htm
Metatool	http://www.biocyc.org/
SBGN	http://www.sbgn.org/Main_Page
Teranode	http://www.teranode.com/
GUI-Modelling and Simulation Tools	
E-Cell	http://www.e-cell.org/
Gepasi	http://www.gepasi.org/
MATLAB	http://www.mathworks.com/
Maple	http://www.maplesoft.com/
SBML Toolbox	http://sbml.org/
Systems Biology Workbench (SBW)	http://sbml.org/
Non-GUI Modelling and Simulation Tools	
Pathway Analyzer	http://sourceforge.net/projects/pathwayanalyser
COBRA	http://gcrg.ucsd.edu/
JigCell	http://jigcell.biol.vt.edu/

Table 1. This table outlines some of the important resources for systems biology, from pathway databases, to databases of kinetic parameters, as well as modelling and simulation tools.

differential equations detailing kinetic parameters and stochastic modelling, to account for inherent noise in biochemical systems. At a lower level of resolution are the constraint-based modelling techniques such as flux balance analysis (FBA) and stoichiometric analyses, which rely more on global properties of networks, such as stoichiometry and mass conservation, rather than the intricate kinetic parameters.

Boolean networks thrive with lesser data, where interactions between network components are represented by means of Boolean functions such as 'OR', 'AND', 'AND NOT' and so on. Such discrete modelling techniques have applications in several areas. Topological analyses of networks, which are constructed predominantly based on knowledge of association or causality, can also provide interesting insights into the organisation and properties of biological systems. At a further lower level of resolution are Bayesian networks and other statistical learning models, as well as qualitative models of biological systems.

The choice of methods for modelling and simulation is predominantly determined by the quality and quantity of data that are available, as well as the desired objective of the modelling exercise. When well-characterised kinetic parameters are available for a set of reactions in a given pathway, a kinetic model consisting of differential equations describing the rate of change of concentration of each of the metabolites can be constructed. Such a system of equations can then be solved to obtain insights about the essentiality of each component. For example, a mathematical model of glycolysis in *T. brucei* has been built, based on *in vitro* enzyme kinetic data [6].

When kinetic parameters are not available, constraint-based models of reaction networks can be constructed and analysed, obtaining insights into the metabolic capabilities of systems as well as gene essentiality [2, 5]. At a lower level of resolution, interaction networks of metabolites or more importantly, proteins, can be constructed and analysed, obtaining fundamental insights into centrality, and consequently lethality (or essentiality) [7].

4.1 *Kinetic Modelling*

When sufficient mechanistic details are available for cellular processes such as metabolism, signal processing and gene regulation, detailed quantitative predictions about cellular dynamics can be made. Typically, ordinary differential equations (ODEs) are used for this purpose. ODE-based simulations involve a mechanistic representation of the reaction network, with all the involved association/dissociation constants, rate constants and affinities or appropriate approximations. Since such data is not always available, this method has limited applicability. Biochemical reactions are regularly represented by differential equations that indicate the rate of consumption and production of various species involved in the reactions. The system of differential equations so generated can be solved and the system can be simulated. An important caveat is that even where kinetic parameters are available, they have often been determined *in vitro*, rather than *in vivo*, which again significantly impacts the accuracy of simulations. Genome-scale kinetic modelling of biological systems is an interesting challenge that lies ahead in systems biology.

When sufficient mechanistic details are available for cellular processes such as metabolism, signal processing and gene regulation, detailed quantitative predictions about cellular dynamics can be made.

4.2 *Constraint-Based Modelling*

Kinetic data available for the simulation of networks are quite scarce, rendering the kinetic modelling of metabolic networks a challenging task. An approach used often to overcome the limitation of data, is to add appropriate 'constraints' on the systems, so as to make it feasible to find meaningful solutions. Constraints are generally in the form of rules, which define the upper and lower limits of the acceptable values for a given variable or in the form of some well-known laws of chemistry that must be upheld while solving for the system. Constraint-based analyses of reconstructed metabolic networks have proved to be quite effective in various applications such as metabolic engineering, prediction of outcomes of gene deletions, and in the elucidation of cellular regulatory networks .

Genome-scale kinetic modelling of biological systems is an interesting challenge that lies ahead in systems biology.

One specific example of metabolic modelling using a constraint-

Gene deletion studies can be performed by constraining the reaction flux(es) corresponding to the gene(s) (and therefore, of their corresponding proteins(s)), to zero.

based approach is Flux-Balance Analysis (FBA) [8], which uses linear optimisation to determine the steady-state reaction flux distribution in a metabolic network by maximising an objective function, such as ATP production or growth rate. FBA involves carrying out a steady state analysis, using the stoichiometric matrix for the system in question. An important assumption is that the cell performs optimally with respect to a metabolic function, such as maximisation of biomass production or minimisation of nutrient utilisation, on the premise that selection pressures during evolution, guide systems towards optimality. Once an objective function is fixed, the system of equations can be solved to obtain a steady state flux distribution. This flux distribution is then used to interpret the metabolic capabilities of the system.

FBA has the capabilities to address the effects of gene deletions and other types of perturbations on the system. Gene deletion studies can be performed by constraining the reaction flux(es) corresponding to the gene(s) (and therefore, of their corresponding proteins(s)), to zero. Effects of inhibitors of particular proteins can also be studied in a similar way, by constraining the upper bounds of their fluxes to any defined fraction of the normal flux, corresponding to the extents of inhibition. FBA gives a general idea of the metabolic capabilities of an organism; gene deletion studies using FBA yield information on the criticality of genes for the growth/survival of an organism. The analysis of perturbations using flux balance models of metabolic networks provides a handle to analyse the lethality of individual gene deletions, as well as double knock-outs, to identify pairs of genes that are indispensable, as well as to determine and analyse synthetic genetic interactions.

4.3 *Pathway Models*

A pathway model is the lowest level of abstraction in system-based models.

A pathway model is the lowest level of abstraction in system-based models. It looks at only the reactions in the metabolome of an organism and accounts for several of the interactions between the gene products of an organism and its metabolites. However, this is a significant improvement on the mere sequence data that

is often employed for modelling and analysis. Several paradigms exist for pathway modelling and they are reviewed in the literature [9] . Based on the availability of data, a suitable paradigm can be chosen for modelling; this affects the accuracy of the simulations performed on the systems. Some examples of the use of pathway models are illustrated in later sections.

Systems biology finds application in several fields, including metabolic engineering and drug discovery.

4.4 *Network-Based analysis*

Barabási and Oltvai [10] have shown that tools from network theory may be adapted to biology, providing profound insights into cellular organisation and evolution. Hubs which are heavily connected components in a graph may be identified and targeted to 'knock out' a system. In a typical interaction-based modelling of metabolic pathways, connections between the various proteins and metabolites in a system are obtained. When further analysed, specific hubs emerge to be more connected. These hubs may serve as interesting targets as they have the potential to affect several other connections in the system. The advantage of interaction-based modelling is that the amount of data required is relatively less and it is possible to generate interaction networks from existing databases. There is a need for more such derived databases, which would be of immense use in applications such as drug discovery.

5. Promise of Systems Biology

Systems biology finds application in several fields, including metabolic engineering and drug discovery. It has an immense potential to improve our fundamental understanding of biological systems. Biology has itself immensely benefited from building on the study of 'model' organisms such as *Arabidopsis thaliana*, *Drosophila melanogaster*, *C. elegans* and *Escherichia coli*. Systems approaches have been successfully applied for the study of model organisms such as *Escherichia coli*, where the metabolic capabilities have been predicted *in silico* and verified experimentally. Systems-level studies of organisms such as *S. cerevisiae* are expected to significantly impact the study of more complex organisms such as humans.

Systems-level studies of organisms such as *S. cerevisiae* are expected to significantly impact the study of more complex organisms such as humans.

Another field where excellent progress has been made is in the modelling of the heart as a virtual organ.

An excellent application of systems biology in metabolic engineering, with commercial potential, has been illustrated by Stephanopoulos and co-workers, for improving lysine production. Stephanopoulos and co-workers have also reported a genome-wide FBA of *Escherichia coli* to discover putative genes impacting network properties and cellular phenotype, for re-engineering lycopene synthesis [11]. Metabolic fluxes were calculated such as to optimise growth, followed by scanning the genome for single and multiple gene knockouts that yield improved product yield while maintaining acceptable overall growth rate. For lycopene biosynthesis in *Escherichia coli*, such targets were identified and subsequently tested experimentally by constructing the corresponding single, double and triple gene knockouts. A triple knockout construct ($\Delta gdhA\Delta aceA\Delta fdhF$) was identified, which exhibited a 37% increase over an engineered, high producing parental strain.

Another field where excellent progress has been made is in the modelling of the heart as a virtual organ, at various levels [12]. Models of different cell types in the heart have led to the creation of the first virtual organ, which is being used in drug discovery and testing and in simulating the action of devices such as cardiac defibrillators. The culmination of systems modelling lies in the modelling of complete systems, accounting for all component reactions, the localisation of these components and their interactions. The interaction between these organelles or compartments and the interface with the physical world, in terms of external temperature, pH and other effects becomes more relevant in highest levels of biological hierarchy (*Figure* 3). Computational models of human physiology come into play both to relate to whole animal models used in traditional pharmacology and more importantly, to build integrated data-driven models that can be refined to mimic the human physiology more closely.

The IUPS Physiome project (http://www.physiome.org.nz/) is a project that is aimed at describing the human organism quantitatively, to understand key elements of physiology and pathophysi-

ology. The salient features of the project are the databasing of physiological, pharmacological and pathological information on humans and other organisms and integration through computational modelling. The models span a wide range, from diagrammatic schema suggesting relationships among system components, to fully quantitative computational models describing the behaviour of physiological systems and response of an organism to environmental change. Each mathematical model is an internally self-consistent summary of available information and thereby defines a working hypothesis about how a system operates. Predictions from such models are subject to tests, with new results leading to new models. The goal is to understand the behaviour of complex biological systems through a step-by-step process of building upon and refining existing knowledge.

Each mathematical model is an internally self-consistent summary of available information and thereby defines a working hypothesis about how a system operates.

Efforts are underway to extend these concepts further to virtual patients. Entelos' PhysioLab (http://www.entelos.com/) has developed models of human physiology that supplement animal model systems. For example, Entelos' Diabetes PhysioLab has more than 60 virtual patients, each one representing a hypothesis of the pathophysiology of diabetes, constrained by the pathway networks and consistent with validation experiments. Such models have the potential for performing patient profiling, classifying patient types and even to tailor-design treatment regimes, with a long-term goal of making personalised medicine, a reality.

The possibility of drug discovery based on systems biology is exciting – it holds promise for the discovery of more efficacious drugs with fewer adverse effects. Often, adverse drug reactions might emerge on account of the binding of the drug to proteins other than the intended targets. By considering larger systems and accounting for such possibilities, it is possible that such problems may be identified by *in silico* analyses. It is envisaged that the complete understanding of a system in terms of all the components present and their complex interaction network would assist in discovering the *ideal* drug, which has high specificity and effectiveness.

The predictive power provided by data-driven computation has long been a critical component in product development and safety testing in other industries, from aerospace engineering to circuit design.

6. Future Perspectives

The enormous progress in the development of new methods in different branches of biology and their abstraction through computational models that we are currently witnessing, has already shown enormous potential, both for understanding biological processes at a better level, as well as for any application opportunities. These opportunities, which are expected to increase even more in the coming years, promise to make the mission of creating data-driven models and simulations a reality, leading to fundamental changes in the way we discover drugs. The predictive power provided by data-driven computation has long been a critical component in product development and safety testing in other industries, from aerospace engineering to circuit design.

With the current rate of advances in systems biology, we can also expect significant enhancements in pathway models, process models and indeed in entire system models, both in terms of mathematically representing complex phenomena as well as in terms of mimicking and simulating the biological events. The success of the virtual heart project [12] and creation of virtual patients representing different pathophysiologies are suggestive of this trend. We can also envisage that the use of pharmacogenomics and tailor-made medicines could be distinct possibilities in the near future. In short, the stage is all set for the integration and application of skills from mathematics and computer science to address complex problems in biology and medicine, in a big way.

Suggested Reading

[1] **A Aderem, Systems biology: its practice and challenges, *Cell*, Vol.121, No.4, pp. 511–3, 2005.**

[2] **J A Papin, *et al*., Reconstruction of cellular signalling networks and analysis of their properties, *Nat. Rev. Mol. Cell Biol.*, Vol.6, No.2, pp.99–111, Vol.3, 2005.**

[3] **T Ideker, T Galitski, and L Hood, A new approach to decoding life: systems biology, *Annu. Rev. Genomics Hum. Genet.*, Vol.2, pp.343–72, 2001.**

[4] P W Anderson, More Is Different, *Science*, Vol.177, No.4047, pp.393–396, 1972.

[5] L H Hartwell *et al*, From molecular to modular cell biology, *Nature*, Vol.402, No.6761 Suppl, pp.C47–52, 1999.

[6] B M Bakker *et al*, Glycolysis in bloodstream form Trypanosoma brucei can be understood in terms of the kinetics of the glycolytic enzymes, *J. Biol. Chem.*, Vol.272, No.6, pp.3207–15, 1997.

[7] H Jeong *et al*, Lethality and centrality in protein networks, *Nature*, Vol.411, No.6833, pp.41–2, 2001.

[8] K J Kauffman, P Prakash, and J S Edwards, Advances in flux balance analysis, *Curr. Opin. Biotechnol.*, Vol.14, No.5, pp.491–496, 2003.

[9] J Stelling, Mathematical models in microbial systems biology, *Curr. Opin. Microbiol.*, Vol.7, pp.513–518, 2004.

[10] A L Barabasi and Z N Oltvai, Network biology: understanding the cell's functional organization, *Nat. Rev. Genet.*, Vol.5, No.2, pp.101–13, 2004.

[11] H Alper *et al*, Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli, *Metab. Eng.*, Vol.7, No.3, pp.155–164, 2005.

[12] D Noble, From the Hodgkin-Huxley axon to the virtual heart, *J. Physiol.*, Vol.580, (Pt 1), pp.15–22, 2007.

Address for Correspondence
Karthik Raman
Department of Biochemistry
University of Zurich
Winterthurerstrasse 190,
CH-8057, Switzerland
Email:
k.raman@bioc.unizh.ch

Nagasuma Chandra
Bioinformatics Centre
Raman building
Indian Institute of Science
Bangalore 560 012, India
Email: nchandra@
physics.iisc.ernet.in