Published in IET Systems Biology
Received on 14th January 2009
doi: 10.1049/iet-syb.2009.0005

ISSN 1751-8849

SysBioMed report: Advancing systems biology for medical applications

O. Wolkenhauer[1] D. Fell[2] P. De Meyts[3] N. Blüthgen[4] H. Herzel[5] N. Le Novère[6] T. Höfer[7] K. Schürrle[8] I. van Leeuwen[9]

[1]*University of Rostock, Albert Einstein Str. 21, 18051 Rostock, Germany*
[2]*School of Life Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, UK*
[3]*Receptor Systems Biology Laboratory, Hagendorn Research Institute, Novo Nordisk A/S, Niels Steensens Vej 6, DK-2820 Gentofte, Denmark*
[4]*Institute of Pathology, Universitaetsmedizin, Charité, Chariteplatz 1, 10117 Berlin, Germany*
[5]*Institute for Theoretical Biology, Humboldt University, Invalidenstraße 43, 10115 Berlin, Germany*
[6]*Computational Neurobiology, EMBL-EBI, Wellcome-Trust Genome Campus, Hinxton CB10 1SD, UK*
[7]*German Cancer Research Center (DKFZ), Research Group Modeling of Biological Systems, Im Nuenheimer Feld 267, 69120 Heidelberg, Germany*
[8]*DECHEMA e.V., AVBTL Biotechnologie, Theodor-Heuss-Allee 25, D60486 Frankfurt am Main, Germany*
[9]*Department of Mathematics, University of Dundee, Welcome-Trus/MSI-Complex, Old Hawkhill, Dundee DD1 5EH, UK*
E-mail: olaf.wolkenhauer@informatik.uni-rostock.de

Abstract: The following report selects and summarises some of the conclusions and recommendations generated throughout a series of workshops and discussions that have lead to the publication of the Science Policy Briefing (SPB) Nr. 35, published by the European Science Foundation. (Large parts of the present text are directly based on the ESF SPB. Detailed recommendations with regard to specific application areas are not given here but can be found in the SPB. Issues related to mathematical modelling, including training and the need for an infrastructure supporting modelling are discussed in greater detail in the present text.)

The numerous reports and publications about the advances within the rapidly growing field of systems biology have led to a plethora of alternative definitions for key concepts. Here, with 'mathematical modelling' the authors refer to the modelling and simulation of subcellular, cellular and macro-scale phenomena, using primarily methods from dynamical systems theory. The aim of such models is encoding and testing hypotheses about mechanisms underlying the functioning of cells. Typical examples are models for molecular networks, where the behaviour of cells is expressed in terms of quantitative changes in the levels of transcripts and gene products. Bioinformatics provides essential complementary tools, including procedures for pattern recognition, machine learning, statistical modelling (testing for differences, searching for associations and correlations) and secondary data extracted from databases.

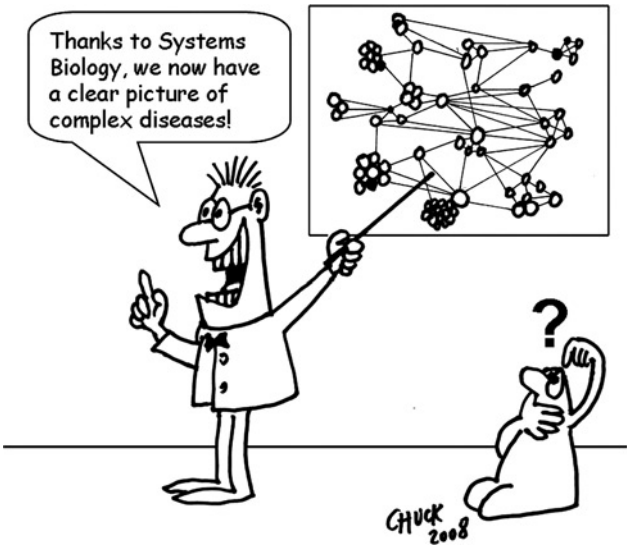
Dynamical systems theory is *the* natural language to investigate complex biological systems demonstrating nonlinear spatio-temporal behaviour. However, the generation of experimental data suitable to parameterise, calibrate and validate such models is often time consuming and expensive or not even possible with the technology available today. In our report, we use the term 'computational model' when mathematical models are complemented with information generated from bioinformatics resources. Hence, 'the model' is, in reality, an integrated collection of data and models from various (possibly heterogeneous) sources. The present report focuses on a selection of topics, which were identified as appropriate case studies for medical systems biology, and adopts a particular perspective which the authors consider important. We strongly believe that

mathematical modelling represents a natural language with which to integrate data at various levels and, in doing so, to provide insight into complex diseases:

- Modelling necessitates the statement of explicit hypotheses, a process which often enhances comprehension of the biological system and can uncover critical points where understanding is lacking.
- Simulations can reveal hidden patterns and/or counter-intuitive mechanisms in complex systems.
- Theoretical thinking and mathematical modelling constitute powerful tools to integrate and make sense of the biological and clinical information being generated and, more importantly, to generate new hypotheses that can then be tested in the laboratory.

Medical Systems Biology projects carried out recently across Europe have revealed a need for action:

- While the need for mathematical modelling and interdisciplinary collaborations is becoming widely recognised in the biological sciences, with substantial implications for the training and research funding mechanisms within this area, the medical sciences have yet to follow this lead.
- To achieve major breakthroughs in Medical Systems Biology, existing academic funding schemes for large-scale projects need to be reconsidered.
- The hesitant stance of the pharmaceutical industry towards major investment in systems biology research has to be addressed.
- Leading medical journals should be encouraged to promote mathematical modelling.



Introduction

Why systems biology for medical applications?

Conventional modes of medical and biological explanation rely primarily on verbal reasoning and are only suited for dealing with mechanisms that involve small numbers of components and short chains of causality. The diseases most relevant for humankind, however, involve a large number and variety of components interacting through complex networks. New approaches are therefore required to fuel further advances in modern medicine. We strongly believe that systems biology provides a particularly promising avenue to tackle complex systems through an interdisciplinary approach that combines experimental work with mathematical modelling. In the medical sciences, it has the potential to make important contributions, among others, to refining existing treatment protocols; identifying new drugs and therapies; designing and testing novel medical devices (e.g. artificial implants); understanding the aetiology, progression and symptomatology of various diseases, facilitating early diagnosis; and improving personalised prognosis and decision-making among others.

In terms of drug discovery, there will be a large number of compounds entering trials in the coming years. Traditional methods to evaluate the compounds' effectiveness will no longer be feasible and, consequently, the pharmaceutical industry will face a rapid increase in research costs in conjunction with a steady decline in the number of drugs

doi: 10.1049/iet-syb.2009.0005

being approved. Strategies are thus required to reduce the fallout rate by setting priorities to decide which compounds should be fully tested. Furthermore, when compounds do reach the market, clinicians require criteria that enable them to choose between suitable drugs and decide upon their optimal combination and administration schedules. A major breakthrough in systems biology is imminent – medical research is at the crossroads; conventional approaches will no longer work.

Value of mathematical modelling

The value and potential of systems biology is best illustrated by examples in which the formulation of a mathematical model was key for major research advances. A widely known and conceptually influential example is the work of Hodgkin and Huxley on nerve impulses. Their groundbreaking findings, which led to the award of a Nobel Prize in 1963, would not have been possible without computations being performed based on a mathematical model. However, we do not need to look back 50 years. Recent success stories of mathematical modelling include applications to HIV, Hepatitis A and B, as well as heart disease.

The group led by Dennis Noble as part of the Human Physiome Project constructed a human heart model, connecting intracellular dynamics of electrical currents, receptors and channels with organ function. This model was successfully used to predict side effects of drugs and to design Ranolazine, an FDA-approved drug for treatment of chronic angina. HIV has also been subject to mathematical modelling. Thanks to the combination therapy of HIV, quantitative data was generated on the decay of the virus load. The corresponding mathematical models suggested a high-turnover rate of the virus and made it possible to estimate the decay rates of the free virus and infected cells. Such models can also be used to describe the HIV dynamics below detection levels and to predict the re-emergence of the virus after treatment. Similar models by F. Michor and colleagues succeeded in characterising the treatment of chronic myelocytic leukaemia in the presence of Bcr-Abl fusion protein.

From the above examples, we can learn the following:

- Success was achieved when quantitative data became available.
- Even simple mathematical models can be of practical use.
- The interdisciplinary process leading to the formulation of a model is as important as the resulting model itself.

In recent years, comprehensive whole-cell, whole-organ and whole-body models have emerged as visions for research in systems biology. Such models are built to analyse, simplify and reduce complex interactions, and to identify and quantify input–output relations as well as generic principles ('laws') that determine the functioning of the corresponding system. However, the value of models specifically tailored to answer particular research questions should not be overlooked. The use of cell-cycle models by the pharmaceutical industry, for instance, demonstrates that a whole-cell model is not essential for evaluating the effect of drugs.

Systems biology highlights the dynamical nature of the (mal)functioning of cells in the development and progression of diseases. Although disease progression can be slow (possibly years), it depends on cellular events, such as apoptosis, cell division, differentiation and neurons, which in their turn are regulated by subcellular processes taking place within a timescale of seconds or minutes. Hence, cells, organs and organisms rely on dynamic interactions between large numbers of components, the emergent behaviour being nonlinear in nature. The spatio-temporal dynamics of the system as a whole are of such complexity that their understanding challenges conventional, heuristic approaches and makes mathematical modelling a necessity.

Mathematical modelling provides *in silico* tools with which to carry out and iterate virtual experiments. One of the long-term goals is to support computational experiments that otherwise might be dismissed for being unethical, expensive, time consuming or simply impossible. In an era in which computer requirements are no longer a serious limitation, the growing field of systems biology is expected to blossom even further, leading to fundamental breakthroughs in both biology and medicine. However, to overcome existing hurdles in medical systems biology and to form a new generation of scientific investigators and decision makers that can sustain these exciting developments, new targeted funding initiatives for research and training are required.

Promising application areas

A number of medical areas where systems biology looks particularly promising have been selected for more in depth consideration. The following sections highlight some of the conclusions from these workshops.

Cancer diseases are systemic by nature, and reductionist approaches have failed to improve treatment and understanding substantially. Despite the variability in the nature of cancer, it is expected that systems biology can make essential contributions to

- Personalised medicine by building computer models of different stages of the disease.
- The identification of early markers for non-invasive prognosis by investigating tumour development.

• Improving treatment of later stages by comparing biochemical networks and gene expression levels in primary tumours and metastases.

In cancer research, it is important to promote the use of mathematical and computational methods to integrate data on gene expression, phosphoproteomics, epigenetics and metabolomics. Linking to existing, advanced models for fundamental processes (e.g. cell-cycle control, apoptosis and tumour growth) would also be beneficial. In addition, significant impacts are expected from recent developments in proteomics and low-cost sequencing.

In the first instance, large-scale systems biology efforts should concentrate on specific cancer types that have high medical relevance, well understood molecular pathology, high quality experimental models and a variety of targeted therapies available. Here it is expected that systems biology will provide new insight into why certain therapies fail, and thereby help in choosing the right therapy and treatment protocols in the near future.

As the population of the EU gradually grows older, the social and economic strain posed by **age-related diseases**, – such as cancer – is expected to become even greater. This prospect has created an urgent need for progress in the different areas of ageing research, with the ultimate goal of improving the quality of life of the elderly. Given that ageing is a complex multi-factorial process involving many biological/physiological phenomena, a multi-disciplinary approach has become essential to integrate the existing knowledge and, even more importantly, to generate new experimentally testable hypotheses. It is notable that, although cancer modelling and mathematical gerontology are both rapidly growing areas of research, little theoretical effort has been specifically aimed at enhancing our understanding of the interrelations between ageing and cancer. A basic requirement, namely discriminating between time-dependent and ageing-dependent events, constitutes a major challenge.

Inflammatory disorders encompass a large group of diseases, many of which are wide-spread, such as rheumatoid arthritis and asthma. A paradigm for the systemic nature of inflammation is its relation with cancer, where the inflammatory microenvironment may contribute to tumour progression, or an appropriate immune response may suppress tumours. Cytokine- and cell-based therapies for treating chronic inflammation and for anti-cancer immune therapies are being developed. Due to our lack of understanding of the complexity of the regulatory networks involved, the biological outcomes are not always as predicted, sometimes including dramatic failures, and can have considerable side effects. A systems biology approach on inflammation that can bridge from the molecular to the organism level is promising, as many tools for the necessary quantitative studies are available, including advanced techniques for monitoring molecular networks in primary immune cells, population dynamics of lymphocytes in animals and humans, and ready access to a great variety of mouse models. Recently, multi-photon microscopy has been applied to image immune system dynamics *in vivo*. Mathematical models are being developed for submodules at different levels, including signal transduction in lymphocytes and macrophages, lymphocyte differentiation and population dynamics.

Diabetes mellitus is rapidly becoming a worldwide epidemic, especially type 2 diabetes and the associated metabolic syndrome(s), driven by the increase in obesity. The common varieties of both type 1 and type 2 diabetes are multi-factorial polygenic diseases whose pathogenetic complexity has eluded conventional reductionist approaches. As a result, there are only a few efficacious drugs besides insulin, which is vital for people suffering from type 1 diabetes, and currently no cure nor prevention methods. It is generally believed that the genetic background of both type 1 and type 2 diabetes results from unfavourable combinations of multiple common gene alleles that determine beta cell function, susceptibility and survival, as well as metabolic homeostasis. Epigenetic factors probably also play a major role. Recent progress in genome-wide scans for association has started to unravel the genes that confer increased susceptibility to type 1 and type 2 diabetes, often unexpected from a candidate gene viewpoint. A complete elucidation of these genes will reveal critical nodes in pathways essential for normal beta cell function and survival as well as insulin sensitivity and metabolic stability.

The dynamics of proliferation, of metabolism, of brain activities and of immune response are strongly influenced by **the circadian clockwork**. Consequently, mathematical models of diseases and therapies have to take into account daily variations of gene expression, metabolism and behaviour.

Organisms have developed biological clocks that enable them to adapt to the 24 h period of the solar day. The so-called circadian clocks are autonomous oscillators that regulate the temporal organisation of physiology, metabolism and behaviour. The master clock in the hypothalamus is driven by transcriptional–translational feedback loops. DNA arrays reveal that hormone secretion, sympathetic innervation, body temperature, feeding time and activity rhythms regulate about 10% of all genes in peripheral tissues including cell-cycle regulators, cytokines and genes involved in detoxification. A first generation of mathematical models is already available to simulate the hierarchical organisation of circadian rhythms.

Chronotherapy (i.e. an optimisation of dose-time medication schedules) has been successfully applied for decades. The effects of chemotherapy, for instance, exhibit circadian rhythms since the proliferation of normal cells and cancer cells is gated by the circadian clock, cancer cells

doi: 10.1049/iet-syb.2009.0005

being less well synchronised. Moreover, the detoxification of cytostatic drugs depends on the time of their administration.

Neuropathologies affecting the basal ganglia, and in particular the dopaminergic system and its targets, form a major public health problem in western societies.

Neurodegenerative or neurodevelopmental diseases, such as Parkinson's disease, Huntington's chorea or schizophrenia are known to be on the rise. Many behavioural disorders that share the same biological substrate, such as drug addiction, depression, obsessive compulsive disorder (OCD), attention-deficit/hyperactivity disorder (AD/HD), and Tourette syndrome, also show an increasing trend.

Many of these neuropathologies are in fact multi-factorial diseases (e.g. Parkinson's disease or schizophrenia). However, the picture emerging from the conjunction of genome wide analyses and cellular assays shows that multiple causes can be linked through common signalling and metabolic pathways, urging for a systems biology approach. Although both the pathogenesis and symptomatology of these disorders seem very different, the technological bottlenecks faced when developing a systems biology approach are similar. Furthermore, they will all benefit from quantitative descriptions of the same biological systems.

On the basis of the societal impact, the possibilities of the current technology, the availability of existing animal models, the access to patients and the existing and foreseeable modelling effort, it is recognised that a systems biology approach would be mostly useful to study Parkinson's disease, schizophrenia and drug addiction.

The generation of kinetic information should be funded using classical distributed funding. This concerns chemical kinetics constants, diffusion constants, electrophysiological parameters of subcellular compartments and dynamics of neurotransmitter release.

Modelling infrastructure for systems biology

A core component of Medical Systems Biology is the ability to create dynamic models of biological processes. While European scientists have been well represented among those who have successfully demonstrated the importance of building theoretical models, several critical issues need to be addressed to fully exploit the potential applications of systems biology in the medical areas discussed in this report.

Several fundamental biological processes play a central role in more than one of the diseases discussed here. Among them are cell division, differentiation, programmed cell death (apoptosis) and signalling from the cell surface to the nucleus (signal transduction). The common interest in these phenomena makes understanding the networks responsible for their regulation a priority for systems biology research. Notably, although they involve molecular interactions taking place in a timescale ranging from nanoseconds to hours or days, these networks can affect disease processes that evolve over years or even decades. Formulating a multi-scale model for a disease thus implies the challenge of incorporating models describing molecular dynamics over short-time intervals into long-term macro-scale models. It is unfeasible to retain the fine level of detail at the subcellular level. Yet there are slow, macro-scale processes that can influence fast, subcellular events, and vice versa.

Analogous issues arise when dealing with morphological and spatial features. In some scenarios, spatial effects are negligible, whereas in others they influence the behaviour of a system substantially. Cells can function differently depending on their location within an organ or tumour and signal transduction pathways can be switched on or off depending on the cell's shape and size. Again, this creates a need to divide the system into parts that can be modelled separately and then decide how to combine the results. In systems biology, assembling various parts is not only an issue for modellers. An intrinsic fragmentation needs also to be overcome on the experimental side. Key cellular processes, for example, tend to be studied in relative isolation by independent experimental groups, using different sets of techniques. Yet it is certain that these processes are interlinked within the cell.

Recommendations

- The availability of integrated models of fundamental processes in normal, healthy cells are relevant for the diseases discussed in this essay. Advances in modelling the (dys)function of these processes for one particular disease can have significant benefits for research on another pathology. Therefore, systems biology approaches on the regulation of fundamental phenomena (e.g. gene expression, the cell cycle, apoptosis and cellular metabolism) are essential for progress.

- Techniques for coupling/embedding models of components built on disparate time and length scales, and often with different modelling techniques, into larger models spanning much longer scales are in their infancy and require further investigation. This also underlines the need to extend existing standards for model description to support this integration.

- Progress requires cooperative interdisciplinary research, not only between modellers and experimentalists, but also between experimentalists working in different topic areas. The use of different methods and technologies results in operational divisions between research groups focusing on different cellular processes. This leads to discontinuities in the type, quality and extent of the information available to modellers. Integrated projects are the best way to overcome these problems.

• Funding schemes for Medical Systems Biology should support the design of novel techniques for data-based system identification and analysis, including theoretical concepts for the design of experiments, hypothesis testing and effective algorithms to solve problems of computational complexity.

General recommendations

The limited availability of **high-quality quantitative data** still constitutes a major bottleneck for the application of mathematical modelling in biology and medicine. The generation of such data (e.g. quantitative proteomics) is more costly and time consuming than conventional experiments, making it unfeasible for small research teams. We therefore recommend the creation of interdisciplinary centres of advanced technologies, including high-throughput (HT) DNA sequencing, metabolomics, proteomics and phosphoproteomics (SILAC, HT mass spectrometry, capillary isoelectrofocusing), advanced antibody-based methods using array and FACS technology, HT microscopy/imaging, protein–protein interactions, and combined RNAi, in order to promote efficient standardisation, access and sharing of data.

The **integration of data and models** is required. This implies

• Comparative studies and integration of knowledge gained from different experimental model systems (cell lines, animal models, patient samples).

• Merging mathematical models of gene expression, regulation, signal transduction and metabolic networks (multi-level modelling).

• Combining different conceptual frameworks for mathematical modelling (e.g. deterministic/stochastic and discrete/continuum models).

• Coupling information from bioinformatics resources, data mining and pattern recognition with dynamic models.

• Integration of models at different temporal and spatial scales (multi-scale modelling). This involves the integration of functional models (e.g. signalling) and structural models (e.g. tumour growth).

Successful systems biology projects serve to illustrate the importance of the modelling process *per se* and to demonstrate the intrinsic value of interdisciplinary interactions that cannot be overemphasised. The path to a model is already a goal. The process by which a model is established should not (and cannot) be automated. However, if done correctly, there is nothing more practical than a good model/theory.

The culture and particularities of training in the medical sciences hamper the introduction of modelling courses in the curricula. As only a few European institutions have succeeded in overcoming this problem, there is a growing need for opportunities to enable experimentalists working in a medical environment to liaise with modellers. The key recommendations are to introduce students to systems biology at an early, undergraduate level and to encourage students who wish to pursue a career in this field to acquire both experimental and theoretical skills by switching between the two environments at the Masters, doctoral and/or postdoctoral levels.

Acknowledgments

The report was compiled on the basis of workshops organised by SysBioMed, an EU FP6-funded Specific Support Action (SSA), supported by the Directorate F: Health Research, Unit F.4: Genomics and Systems Biology. The project was initiated by Dr. Rosita Cottone (German Federal Ministry for Education and Research). The co-applicants for the proposal were Professor Olaf Wolkenhauer (University of Rostock, scientific coordinator), Dr. Carole Moquin-Pattey (European Science Foundation – European Medical Research Councils), Dr. Astrid Lunkes (European Science Foundation – LESC), Dr. Thomas Bruhn (European Science Foundation – ESF) and Dr. Karsten Schürrle (Dechema e.V., project management).

doi: 10.1049/iet-syb.2009.0005