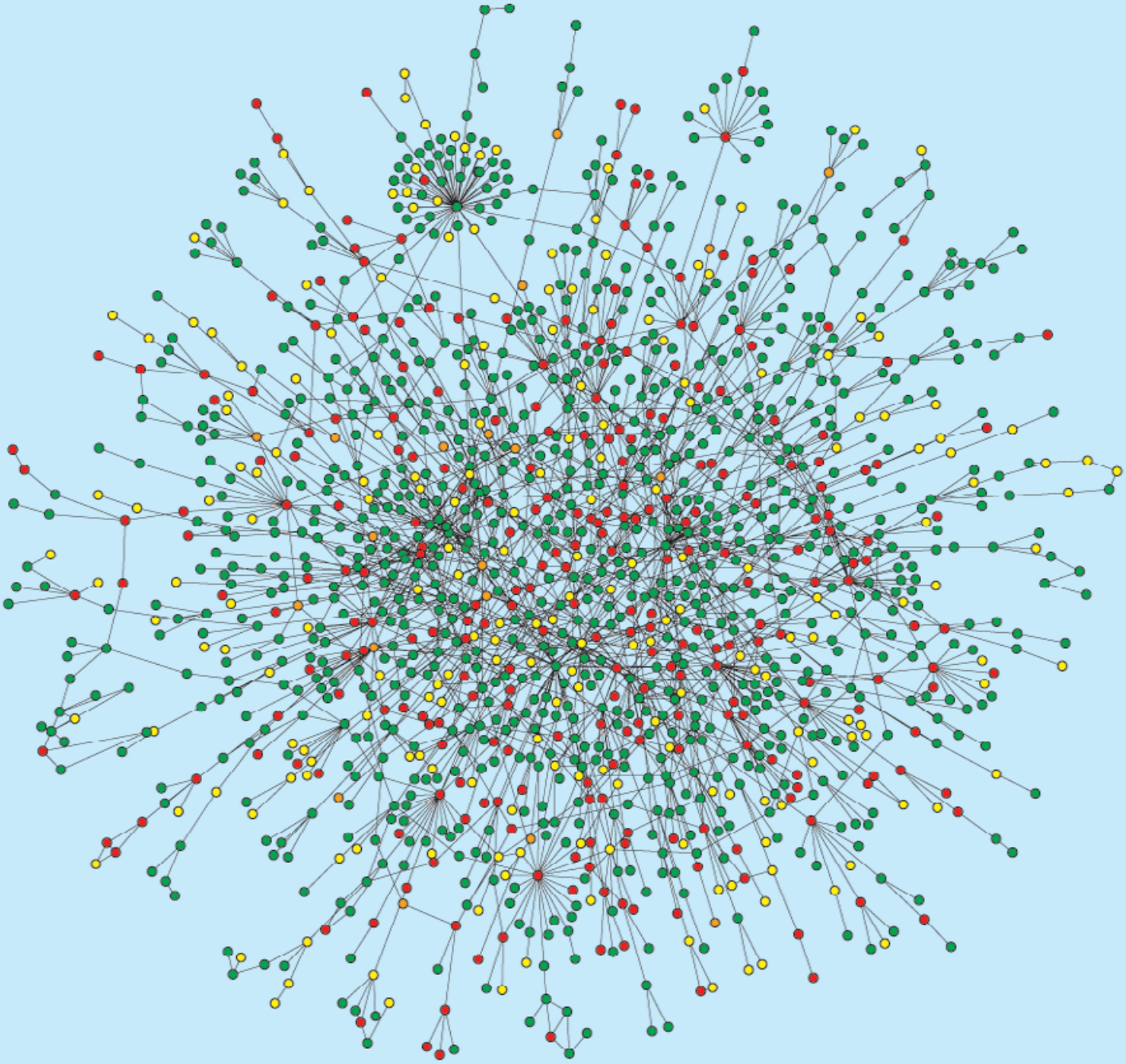
EUROPEAN SCIENCE FOUNDATION

SETTING SCIENCE AGENDAS FOR EUROPE

FORWARD LOOK

# Systems Biology: a Grand Challenge for Europe



www.esf.org

**The European Science Foundation (ESF) was established in 1974 to create a common European platform for cross-border cooperation in all aspects of scientific research.**
**With its emphasis on a multidisciplinary and pan-European approach, the Foundation provides the leadership necessary to open new frontiers in European science.**
**Its activities include providing science policy advice (Science Strategy); stimulating co-operation between researchers and organisations to explore new directions (Science Synergy); and the administration of externally funded programmes (Science Management). These take place in the following areas:**
**Physical and engineering sciences; Medical sciences; Life, earth and environmental sciences; Humanities; Social sciences; Polar; Marine; Space; Radio astronomy frequencies; Nuclear physics.**
**Headquartered in Strasbourg with offices in Brussels, the ESF's membership comprises 75 national funding agencies, research performing agencies and academies from 30 European countries.**
**The Foundation's independence allows the ESF to objectively represent the priorities of all these members.**

**Acknowledgements**

The European Science Foundation (ESF) is grateful to all the experts who contributed to the ESF Forward Look report on "Systems Biology" and especially to Prof. Roel van Driel and Prof. Hans Westerhoff, who agreed to chair this activity, for their commitment during the entire process. They have been supported in their task by the Chairs and Heads of Unit of the Scientific Standing Committees at ESF in charge of:

- Life, Earth and Environmental Sciences (LESC): Prof. Alexandre Quintanilha, Dr. Martina Hildebrandt (until July 2005) and Dr. Arja Kallio (from September 2006) who initiated the activity;
- Physics and Engineering Sciences (PESC): Prof. Michel Mareschal and Dr. Patrick Bressler;
- Medical Sciences (EMRC): Prof. Clemens Sorg (until April 2006), Prof. Liselotte Højgaard (from September 2006) and Dr. Carole Moquin-Pattey, who brought the activity to completion.

A special thank is addressed to the external experts who provided us with invaluable advice and to Dr. Thomas Bruhn, Science Officer at EMRC, who coordinated and oversaw the editing of the report.

**Editor:**
Prof. Roel van Driel

Cover:
A graph representing the interaction network of yeast (S. cerevisiae) proteins. Each circle (node) represents a protein and each connection between two circles a protein-protein interaction. This picture shows 1870 proteins and 2240 interactions. The colour of a node signifies the phenotypic effect of removing the corresponding protein (red, lethal; green, nonlethal; orange, slow growth; yellow, unknown). Among others the graph shows that some proteins interact with many partners, whereas others make contact with only a few. The figure also illustrates that it is difficult to depict complex biological networks. This one is from a paper published in 2001. Present protein-protein networks of yeast proteins comprise many thousands of nodes and can no longer be presented in an image as the one depicted here.
Reprinted by permission from Macmillan Publishers Ltd. Jeong H. and Mason S.P. (2001) Lethality and centrality in protein networks. *Nature* 411: 41

# Contents

# A. Foreword

Biological and biomedical research is undergoing revolutionary developments that are likely to have a lasting impact on society. These developments involve scientific disciplines including physics, chemistry, mathematics, engineering and computer science among others. They enable us to know and measure the properties and interplay of the molecules that constitute life. In principle we are now capable of unravelling complete sets of chemical reactions, interactions and dynamic structures through which molecules, cells and organs carry out specific functions of living organisms, including humans. Integrating the explosively growing amounts of data available on these components and generating understanding on how their maze of interactions in time and space govern life is termed *Systems Biology.*

Systems Biology evolved by recognising that biological systems are far too complex to be solved by classic biological approaches. Systems Biology tightly integrates expertise from physicists, mathematicians, engineers with biological knowledge. It gives a central role to predictive mathematical models that integrate all relevant data on the topic of investigation and exploits such models to decide which experiments are most effective. In this way, an effective and goal-oriented iterative cycle of model-driven experimentation and experiment-driven modelling is initiated.

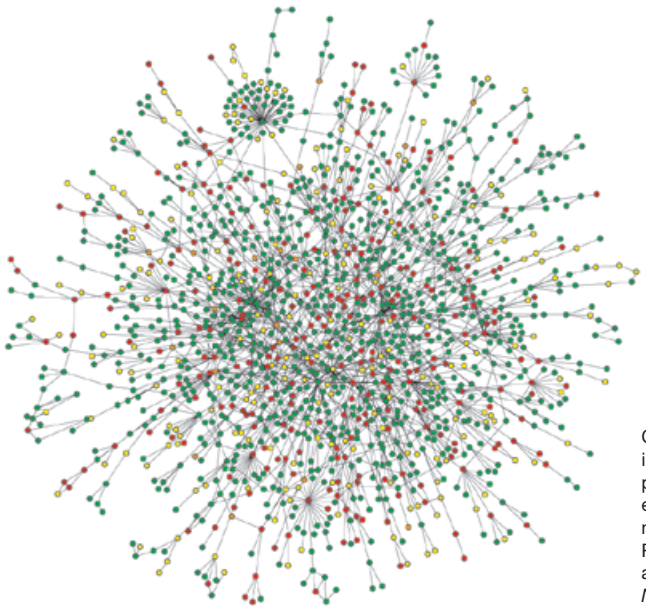
As Systems Biology progresses, multifactorial diseases, such as diabetes, arthritis, heart failure and cancer, may be understood in terms of failure of molecular components to cooperate properly. Consequently, complex diseases may be approached and treated in a much more rational and effective way. It should be Europe's ambition to be at the forefront of pinpointing the systemic causes of diseases, aiming at the rational design of targeted therapies and drugs.

This report is the outcome of the ESF Forward Look on Systems Biology that has been conducted in 2004 and 2005, involving high-level international experts from academia and industry, who agreed upon a set of specific recommendations dedicated to the needs and requirements of a European Systems Biology approach of key problems in health and biotechnology (see Section D of this report). The recommendations were first published in the ESF Science Policy Briefing No 25 in October 2005. Since then the implementation of recommendations has begun, in part through ERANets under FP6; further steps of implementation are expected through FP7 and other instruments.

To underpin these recommendations this report contains a number of essays written by experts in the field, covering a wide range of Systems Biology-related issues that are important in the European context. The report has been subjected to rigorous quality assurance through peer review by international experts in the field and submission to a high-level scientific advisory board.

John Marks
ESF Chief Executive

Bertil Andersson
Former ESF CEO



Characteristics of the yeast proteome. Map of protein–protein interactions. The largest cluster, which contains ~78% of all proteins, is shown. The colour of a node signifies the phenotypic effect of removing the corresponding protein (red, lethal; green, nonlethal; orange, slow growth; yellow, unknown).
Reprinted by permission from Macmillan Publishers Ltd. Jeong H. and Mason S.P. (2001) Lethality and centrality in protein networks. *Nature* 411: 41

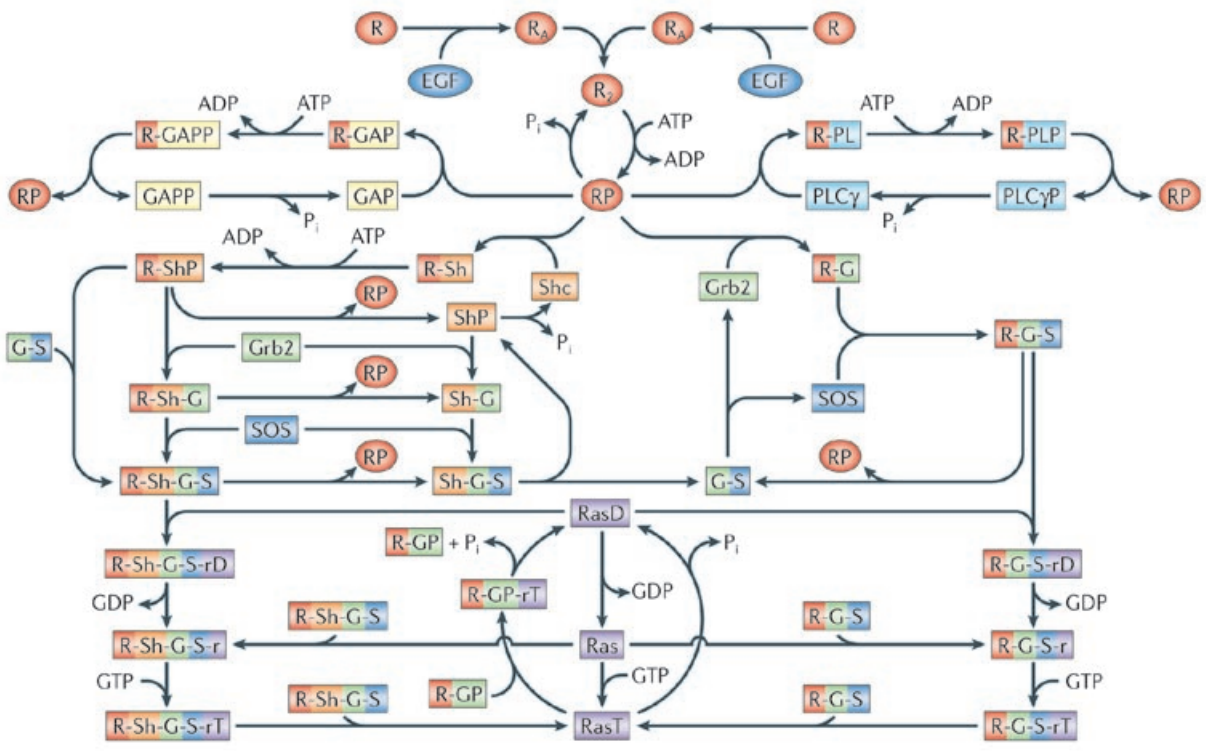
# B. Executive Summary

This report presents a vision of how the field of Systems Biology should develop in the European life sciences arena. This vision is based on extensive discussions during a number of focused workshops and meetings between scientists and policy makers from academia and industry. The result is a set of specific recommendations that aims at synergising Systems Biology efforts in Europe.

The idea for this Forward Look was born early 2003 when a number of scientists began to understand that Systems Biology opens new and promising opportunities for biomedical, pharmaceutical and biotechnological as well as for fundamental biological research. At the same time it became clear that biological research, at least in part, would change to 'big science', involving the joined effort of many investigators in a broad range of disciplines and institutions in many countries. It was easy to foresee that the lack of adequate instruments, language and funding mechanisms for such cross-discipline and large-scale biological research would constitute a major challenge. This vision has been the basis of a proposal for an ESF Forward Look on Systems Biology, which was approved by the ESF Executive Board.

Now, four years after initiating this Forward Look it is clear that our vision about the development of Systems Biology was correct. Large investments in Systems Biology are being made in various European countries, calling for tight cooperation of large numbers of scientists from different disciplines. An early and seminal Systems Biology initiative is the German HepatoSys programme[1], aiming at a comprehensive understanding of the human hepatocyte. Its development in the past three years illustrates both the potential as well as the challenges of Systems Biology. In parallel, an increasing number of new research institutes are being established in Europe, focusing on Systems Biology.

1. HepatoSys programme see www.systembiologie.de




Mechanistic models: keeping track of molecular processes - simplified scheme for the signalling routes that emenate from the epidermal growth factor receptor (EGFR). Reprinted by permission from Macmillan Publishers Ltd. Kholodenko B.N. (2006) Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology* 7: 167.

More than ever, it is clear that we need to develop new instruments in order to focus, synergise, manage and fund large national and international programmes that exploit the possibilities offered by Systems Biology in an efficient and cost-effective way.

Analysis of the Systems Biology field has been carried out by a Steering Committee that has organised eight meetings in the period autumn 2003 to spring 2005, in which different aspects were explored and resulted in the recommendations formulated in this report. The body of this report consists of 12 short essays addressing different aspects of Systems Biology. These include the opportunities in bioengineering and drug development, the need for standardisation of biological experimentation and for new mathematical tools, and the role of European science organisations and industries in developing Systems Biology at the international level. The overriding issue in many contributions is the extreme complexity of biological systems, requiring a true paradigm shift in biological, biomedical and biotechnological research. The starting point for this is the notion that any biological property is the result of the interaction in time and space of a large set of different molecules, cells, organs and/or organisms. Given the extremely complex behaviour of such multilevel networks of interactions, intuitive approaches are ineffective. It requires quantitative and predictive mathematical modelling that helps the biologist to decide what the most informative experiments are. The iterative cycle of model-driven experimentation with experimental data-driven modelling, in combination with novel systems analysis tools, constitutes the very heart of Systems Biology. In this sense, Systems Biology appears to be a rational and effective way towards understanding living organisms, including humans.

This report makes a number of specific recommendations to exploit Systems Biology in Europe.

- Develop a common European road map to invest in Systems Biology to achieve breakthroughs in biomedical, pharmaceutical and biotechnological research, building on European strengths in these fields.
- Create a network of prominent European Systems Biology institutes that take the lead in international standardisation of experimentation, modelling and data management, and making data and models available to the whole scientific community.
- De-fragment European biomedical and biotechnological research and initiate a small number of large-scale, man-on-the-moon type of research programmes that are necessary to achieve true breakthroughs. For this, define the top priorities in European health, such as diabetes and other, metabolic syndrome, ageing and multifactorial diseases. Develop the necessary tools to finance and manage such large programmes.
- Make the teaching programmes less mono-disciplinary. Develop teaching that tightly integrates biology, chemistry, physics, mathematics and engineering.
- Achieve cooperation and synergy between the different national and transnational initiatives in Systems Biology in Europe, starting from the notion that only at the European level we will be able to overcome the bottlenecks of complexity in biology.

Together, these recommendations constitute a solid basis for a European vision for a road map aiming at using Systems Biology to significantly improve the health and economy in Europe in a global context.

Roel van Driel
Hans Westerhoff

# C. Systems Biology: scientific views about what, why and how

## 1. Systems Biology: definitions and perspectives

Lilia Alberghina
*Department of Biotechnology & Biosciences, University of Milano-Bicocca, Milan, Italy*

Living organisms are modulated and selected by the environment in a never-ending process of natural selection, the integration of behavioural properties being the trait that confers selective fitness on each individual organism. Only very rarely are these functional properties dependent upon one or a few genes. More often they derive from the activity of molecular networks that require the action, coordinated in time and space, of a large number of gene products whose molecular function is often unknown. To understand how complex cellular or organismal behaviours are generated is one of the most challenging tasks in present-day biology. Genomic and post-genomic technologies are producing increasingly detailed molecular descriptions of physiological and pathological processes, but they are not able to provide insight into the causal chain of events that characteristically specify, for instance, cell transformation or the development of an egg into an organism. There is a growing awareness that this limitation of the high-throughput analysis is because biological entities are 'systems', i.e. 'a large collection of simple interacting parts that function as a unit'.

In the 20th century, physics and engineering have structured the theory of systems to promote the development of various sophisticated technological fields, and have shown the usefulness of mathematical models in describing the structure of systems and in predicting their behaviour under given sets of circumstances. Models are symbolic representations of a reality that foster understanding and support decision making. Mathematical models permit us to reach a quantitative appreciation of the dynamics of the system that is quite often counter-intuitive because of the presence of regulatory links in the architecture of the system. Quantitative modelling of excitable cells, red cells and of large metabolic networks has paved the way for the use of mathematical formalism and computation in the study of biological processes at the molecular level (see footnote 1 for a review). The term 'Systems Biology' started to be widely used around the year 2000 to indicate the integration of experimental and computational approaches to achieve comprehension and prediction of complex cellular functions[2]. A characteristic trait of Systems Biology is that it relies on an iterative process of model building, comparison with new sets of experimental data, improvement of the model to account for new features and so on[3]. Systems Biology can start by describing the structure of the system as a block diagram and proceed to blow up selected relevant modules of the system, in order to obtain finer resolution, down to the molecular level[4]. Models on different scales and with different levels of resolution have been reported. For example, the electrical behaviour of cardiac cells, sustained by potassium and calcium transport mechanisms, has been incorporated into anatomically detailed models of the ventricles in an attempt to reconstruct the behaviour of the whole organ in healthy and diseased states[5].

Both technological and biological systems are endowed with two features of great interest: (i) function as an emergent property, and (ii) robustness.

(i) A function derives as an emergent property when it is not present in the individual components of the system, but emerges when the various parts interact following an appropriate organisational design. Several important and previously unexplained aspects of biological processes are accounted for, by Systems Biology investigation, as emergent properties of their underlying networks; for instance, integration of signals across multiple time scales and the generation of distinct outputs depending on input strength and duration in signalling pathways[6]; the distribution of control in metabolic pathways and its summation law[7,8]; setting of the critical cell size to enter into S phase to coordinate cell growth with cell cycle progression[9].

(ii) Robustness is the ability to maintain stable functioning despite internal and external perturbations. The analysis of various biological systems indicates that biological robustness is based on aspects of architectural organisation, such as modularity which locally contains perturbations and damage, the presence of decoupling or buffering mechanisms which isolate genetic variations from protein functionality, for instance by chaperone-assisted folding of misfolded and/or mutated proteins, redundancy which allows the substitution of essential proteins with orthologs, and the presence of control circuits, such as feedbacks – and in particular negative feedbacks, which permit the adaptation to a wide range of stimuli[10]. Robustness and emergent properties are strictly linked to the evolvability of biological entities and many authors indicate that robustness may be a feature positively selected by evolution[11]. It is expected that computational genomics will shed light on the evolution of indispensable complex regulatory circuits such as those that control cell proliferation and differentiation. Robustness is not absolute and cells are, in general, robust in the face of frequently occurring perturbations but fragile when dealing with rare events. Moreover, robustness has a cost in terms of allocation of resources. The evolutive acquisition of robustness appears to be one main source of complexity for biological systems.

Computer replicas of chemical reactions and of macromolecular interactions that take place in cells and in organisms are the concrete aims of Systems Biology, which will take decades to complete. In the meantime, the modelling of specific functions, such as metabolism, cell signalling, cell cycle and apoptosis, are going on, using the findings of both post-genomic analysis[12] and of specific hypothesis-driven small-scale experiments. And it is assumed that their elucidation in conjunction with a more profound appreciation of robustness and emergent properties, both as mathematical theory and as new experimental approaches, will offer the paradigm shift that substantially improves drug discovery and development, and allows the establishment of predictive, preventive and personalised medicine[13,14].

References


1. Westerhoff, H.V., and B.O. Palsson. 2004. The evolution of molecular biology into Systems Biology. *Nat Biotechnol*. 22:1249-52.
2. Borodina, I., and J. Nielsen. 2005. From genomes to in silico cells via metabolic networks. *Curr Opin Biotechnol*. 16:350-5.
3. Kitano, H. 2002. Systems Biology: a brief overview. *Science*. 295:1662-4.
4. Alberghina, L., R. Rossi, D. Porro and M. Vanoni. 2005 In *Systems Biology, Definitions and Perspectives*, L. Alberghina and H.V. Westerhoff, eds., Springer Verlag.
5. Noble, D. 2004. Modelling the heart. *Physiology (Bethesda)*. 19:191-7.
6. Bhalla, U.S., and R. Iyengar. 1999. Emergent properties of networks of biological signaling pathways. *Science*. 283:381-7.
7. Heinrich, R., S.M. Rapoport, and T.A. Rapoport. 1977. Metabolic regulation and mathematical models. *Prog Biophys Mol Biol*. 32:1-82.
8. Groen, A.K., R.J. Wanders, H.V. Westerhoff, R. van der Meer, and J.M. Tager. 1982. Quantification of the contribution of various steps to the control of mitochondrial respiration. *J Biol Chem*. 257:2754-7.
9. personal communication
10. Kitano, H. 2004. Biological robustness. *Nat Rev Genet*. 5:826-37.
11. Kurata, H., H. El-Samad, R. Iwasaki, H. Ohtake, J.C. Doyle, I. Grigorova, C.A. Gross, and M. Khammash. 2006. Module-based analysis of robustness tradeoffs in the heat shock response system. *PLoS Comput Biol*. 2:e59.
12. Joyce, A.R., and B.O. Palsson. 2006. The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol*. 7:198-210.
13. Butcher, E.C. 2005. Can cell Systems Biology rescue drug discovery? *Nat Rev Drug Discov*. 4:461-7.
14. Clayton, T.A., J.C. Lindon, O. Cloarec, H. Antti, C. Charuel, G. Hanton, J.P. Provost, J.L. Le Net, D. Baker, R.J. Walley, J.R. Everett, and J.K. Nicholson. 2006. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature*. 440:1073-7.


## 2. Marrying diverse partners – a mixed complementary approach for integrating bottom-up and top-down methods in Systems Biology


Matthias Reuss

*Institute of Biochemical Engineering and Centre Systems Biology, University of Stuttgart, Germany*


In Systems Biology commonly a dichotomy is made between bottom-up and top-down approaches. The two metaphors have incidentally already become very popular in several other disciplines, e.g. cognitive science, computer science, economics, etc. Equally, there are differences and debates, sometimes confusion in the Systems Biology literature about the definitions and context of activities to which the two terms refer. As far as modelling is concerned, the two classes differ mainly with respect to the emphasis placed on the molecular details and the quantification of the dynamic behaviour of the interactions between the components of the system. The two approaches characterise much of modern Systems Biology and, unfortunately, in many cases represent a sort of contrary thinking, driven by pronounced differences in experimental and modelling tools. Thus, there is a strong faction in the Systems Biology community that is convinced that the recent availability of masses of high-throughput 'omics' data is responsible for the growth spurt of Systems Biology and that further development and improvement of the corresponding technological platforms is of basic importance to achieve the high expectations. In contrast, there is also a firm conviction that the emergence of Systems Biology is driven by the demand for integration of fragmentary information about the components and their interrelationships into a coherent whole. Much the same contrast is made by the terms 'hypothesis-driven' and 'data-driven' modelling. This chapter aims at sorting out some of the different opinions and their importance, as well as elaborating a mandate for integration instead of polarity.

### Bottom-up approach

The bottom-up approach is basically a reductionist method and strongly promoted by the concepts and technology of biochemistry and molecular biology. At the core of this approach is the idea of initially aggregating detailed biological knowledge about individual components and quantitative information about their molecular interactions into appropriate modules and then to interconnecting these into architectures suitable for holistic analysis of the system of interest. Typical for the modelling strategy is the modelling cycle explained in Figure 2.1. Depending on any framework of choice, say, deterministic, or stochastic, continuum or discrete modelling approaches, the first

step involves verbal level modelling, where necessary information about the system is collected. This is followed by the model set-up and subsequent solution of equations, performing parameter sensitivity analysis. This process yields sufficient information about new experimental designs which can then be used for the quantification of individual components and their dynamic behaviour. Parameter estimation can then be followed, which paves the way for the testing and validation of the model. The final result is cycled until a satisfactory result is obtained. This modelling cycle is the key to the success of bottom-up or reductionist model building.

Although the reductionist approach is powerful in building logically simple hypotheses and devising ways to test them, it is very difficult to reconstitute a model for a whole biological system by combining the pieces of information it generates. First, using a reductionist approach, the entire system model must be reconstituted by combining information about every molecular step in the system. Any missing pieces of information may block the reconstitution of the system. Therefore, the bottom-up approach requires essentially complete information, including the dynamic behaviour of each step, to build a system model. Study of each molecular step requires intensive research. It is not realistic to expect that nearly complete information will be obtained about any biological system soon. Second, reductionism by definition focuses on information essential to a simplified question and intentionally discards extra information. Usually the success of this approach depends upon the concept of modularity, to shape the problem into practicable and manageable subunits. The major difficulty in applying this strategy, however, is the definition of criteria for the demarcation of these modules to guarantee a certain level of autonomy. Albeit a multitude of methods for decomposition of networks has been suggested, the specification and proof of existence of these modules is still a great challenge for the future. For the time being these modules are most often defined from an empirical, text book-driven decomposition of the network into subsystems performing particular physiological functions. Because of the absence of a rigorous definition of these subunits the question remains whether the fundamental organisation of biological networks or multi-organ systems is modular at all or distributed, or whether it is probably best described as being a little bit of both. This problem is particularly glaring when addressing issues of so-called 'hubs' in molecular networks or entities in which the elements have overlapping roles in superimposed subsystems.

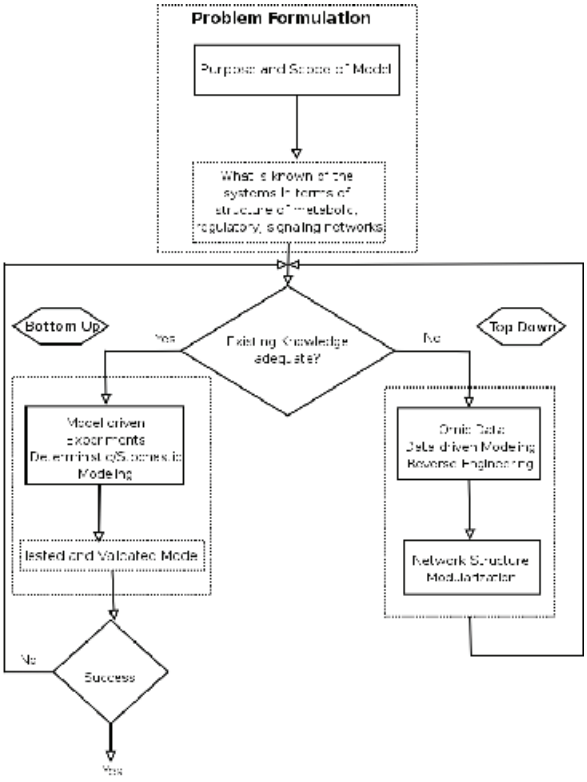



Figure 2.1: Modelling approaches in Systems Biology – bottom up and top down.

**Top-down approach**

The top-down approach is basically linked to a high-throughput reductionism (e.g. assigning biological functions to the genome of an organism). Another aspect, however, is characterised by exhaustive, simultaneous descriptions of biological systems, such as global profiling (transcriptome, proteome, metabolome, interactome, fluxome, etc.). Such broad and detailed information about a biological system provides us with a view different from reductionism – a view of how the system behaves as a whole.

A typical workflow in a top-down approach is shown in Figure 2.2. The primary focus here is the planning and execution of large-scale experiments to generate a lot of information about the genome, proteome, metabolome, etc. Hence the experimental design in this case is a crucial part that can make this strategy successful or otherwise. Perturbation experiments are planned accordingly to take care of environmental factors or genetic factors. This is followed by the design of further experiments, time series, stationary etc. Next step involves the large-scale data generation of 'omics' data. Data analysis follows, and this is one area where innovation keeps raising its head. Different re-

verse engineering methods have been developed and used successfully to reconstruct the network structure from 'omics' data. These include, deterministic methods, Boolean networks, other stochastic methods, neural networks, machine learning etc. The reconstructed network gives us an idea about the structure and the interactions between the players in the system and a general impression of its performance. There is also a possibility of studying the modularity in such reconstructed networks, by studying the interactions of sub-networks within the main network and pinning down their autonomous nature or their lack of it.

**The mixed complementary approach**

There should be no controversy about the need for a dual approach, but only about the relative importance in context with the existing knowledge related to a given problem (Figure 2.2). Once the problem has been formulated, the purpose and the scope of the model and the related known information about different aspects of the structure, and regulation of the system can be studied. If the knowns outweigh the unknowns, then the bottom-up approach can be taken with confidence. But in the case where there is a large number of unknowns, the top-down approach is the logical way to bridge the gap between the knowns and the unknowns. The workflow illustrated in Figure 2.2 additionally suggests an iterative strategy for linking the two approaches for the purpose of mathematical modelling. The characteristic of such a hybrid strategy is that the 'border' between the two approaches can be crossed multiple times in both directions. The top-down approach provides us with broad and detailed information about the biological system and enables us to start with a view of

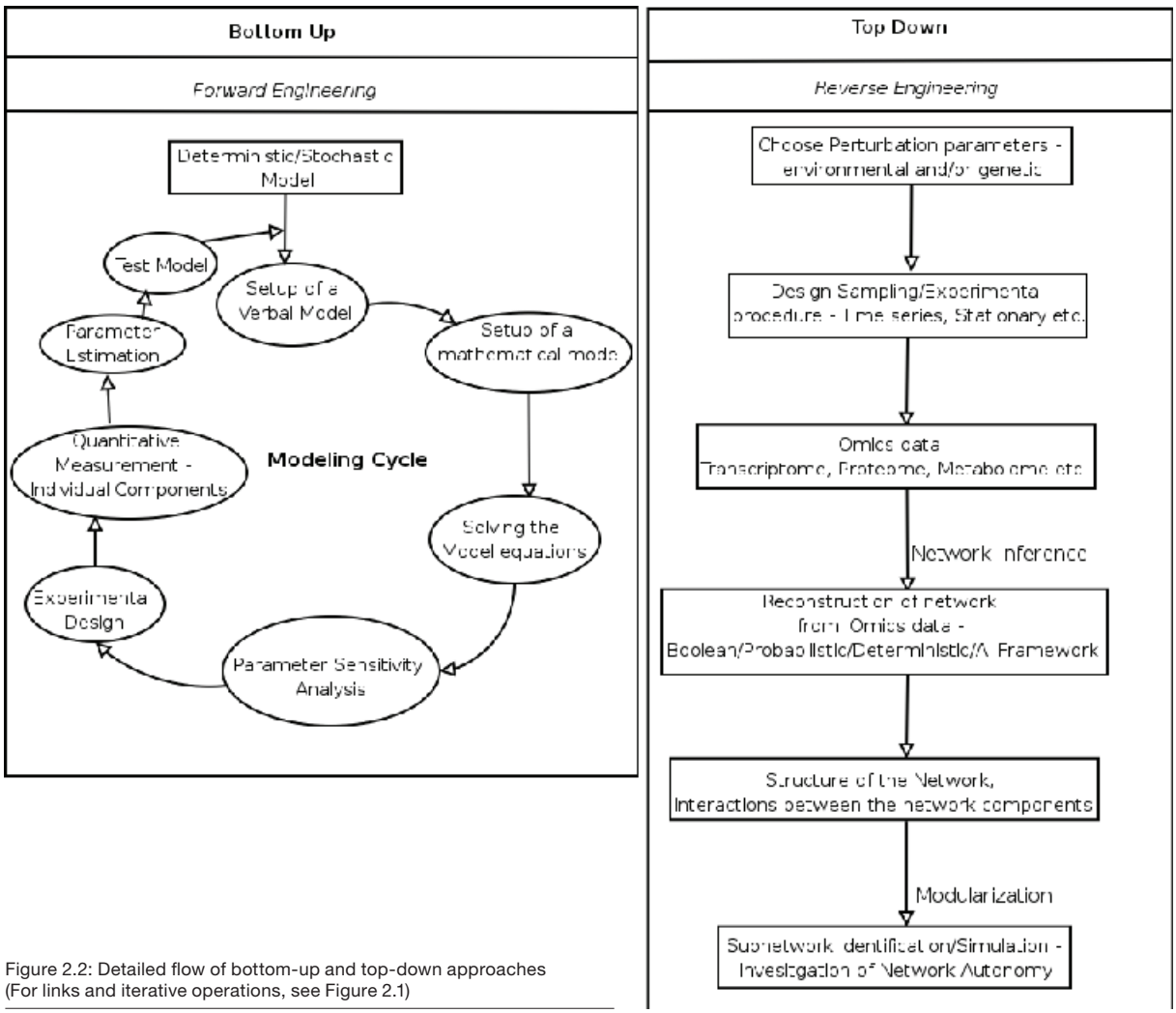



Figure 2.2: Detailed flow of bottom-up and top-down approaches (For links and iterative operations, see Figure 2.1)

how the entire system model should behave. However, if the ultimate goal is the mathematical modelling of the dynamic behaviour of the system – a sort of kernel of Systems Biology – the demand for quantitative data related to the components and their dynamic behaviour is extremely high and often requires targeted quantitative analysis focusing on single components such as quantitative metabolite measurements, protein and mRNA concentrations and quantitative data about the kinetics of protein–protein interactions. Large-scale, high-throughput 'omic' technology is usually not in the position to deliver the required reliability and accuracy of the data to identify the structure and parameters of the kinetic expressions of interest. Therefore, once the players in context with the formulated problem are identified and information about the modularity is provided from rigorous data analysis in the top-down area, the border should be crossed, and the modelling cycle of the bottom-up approach needs to be initiated. The ultimate goal of such a hybrid approach is that the characterisation of the behaviour of parts of the system should be consistent with the expected and/or observed behaviour of the system as a whole.

The top-down approach is to deconstruct the system into smaller parts. The bottom-up approach is to reconstitute elemental steps into larger parts. If the results of these approaches meet in the middle, and if they are consistent in terms of links between modules, multiple function of elements etc., we can be confident that we are on the right track. In other words, we can use information from the reductionist approach as constraints in large-scale model building and vice versa.

The proposed bridge for crosstalking between bottom-up and top-down approaches is driven by the pragmatic view of mathematical modelling of biological systems. This endeavour is possible only with strong coordination between experimental and modelling efforts. Rather than each playing a supplementary role to the other, it is highly important that both areas are tightly linked and function in tandem as one single effort. In this context, the experimental design within the top-down approach should be driven by the modelling strategies for reverse engineering. Therefore, it is the demand of superposition of the two approaches which eventually leads to a model-driven top-down/bottom-up approach instead of flooding computers with 'omic' data and dreaming about a data-driven miracle of Systems Biology modelling.

## 3. Systems Biology calls for coordinated large-scale international efforts to meet Grand Challenges in life sciences

Roel van Driel
*Netherlands Institute for Systems Biology and University of Amsterdam, Amsterdam, The Netherlands*

In this chapter I will argue that biological research is entering a phase in which we will be able to really understand the key processes of life. This offers exciting opportunities, for instance, to effectively tackle major diseases in the next 10 years. Nevertheless, progress is slow. The reasons are the extreme complexity of biological systems and the fragmented way biological (including biomedical) research is carried out. I will make the point that we have to initiate large-scale, highly focused, goal-oriented and systematic international efforts concentrating on a limited number of carefully selected key issues in health and biotechnology. Systems Biology offers basic tools to manage such concerted efforts, giving quantitative and predictive models a guiding role and by tightly combining experiment-based mathematical modelling with model-based experimentation.

**What do we want to know in life sciences?**

The answer to this question depends on the vision and the perspective of the person that answers it. Probably everyone will agree that, eventually, the aim is to understand the complex multilevel biological systems in which molecules, cells, organs, organisms and ecosystems interact. Specifically, results should help us to fight complex multifactorial diseases and detrimental effects of ageing, improve food quality, develop personalised medicines and therapies more efficiently and to better deal with our environment. Doubtless, real understanding of the functioning of multilevel biological systems can bring major societal and economic benefits. This answer to the above question seems almost trivial, because it may be far beyond our present possibilities. However, is this really true?

**It is difficult to take big steps in life sciences: information versus understanding**

In biology we are dealing with extremely complex systems. If I limit myself to the level of molecules and cells, we are dealing with very large numbers of components (molecules, cells) that constitute interaction networks that act in time and space in a nonlinear and in part a precisely controlled manner involving multitudes of complex control loops. Moreover, biological systems are often hierarchical and mix deterministic and stochastic processes. Our present understanding of underlying principles of such highly dynamic

biomolecular and cellular networks is very limited. The complexity explains our slow progress in, for instance, understanding complex diseases, such as cancer, diabetes and metabolic syndrome.

However, there is another side of this coin that is related to the characteristics of the research arena. Biological research, including the biomedical and biotechnological fields, is remarkably fragmented. The only main exceptions are the genome projects, which are highly concerted efforts to sequence and annotate complete genomes of an increasing number of organisms, involving the coordinated effort of many researchers. Generally, however, biological, biomedical and biotechnological issues are addressed by small teams (small compared to the complexity of the problem) in ways that do not allow a simple adding up of qualitative – let alone quantitative – results from different groups. Each group often has its own specific system and conditions. Moreover, biology lacks a well-defined and broadly accepted common language to address and compare complex systems. This fragmented approach did have and still has major advantages, allowing one to hunt for novel components and interactions in highly diverse biological systems. Such an approach has been extremely successful and will remain so in the future. It has resulted in valuable lists of biomolecular and cellular components and insight into their interactions. In short, our, in most cases fragmented, research efforts have resulted in large amounts of extremely important information about complex biological systems, but have not led to a real understanding of biological systems.

**What does 'understanding' of complex biological systems mean?**

Evidently, biological systems are too complex to approach and to describe in an intuitive way. Computers should help us to integrate the multiple interactions of the large numbers of components in time and space. To do so we need quantitative mathematical models that have predictive value, i.e. allow us to make explicit statements about the behaviour of the system that are experimentally verifiable. In this context the word 'understanding' obtains its meaning. The better the model predicts the behaviour of a system, the better we understand it. Moreover, such models help us to uncover unpredicted (emergent) behaviour of biological systems. An interesting corollary of this is that 'understanding' becomes quantifiable: the better the model predicts the behaviour of the system, the better we understand the system.

**Approaching complex biological systems requires large-scale systematic experiments**

There is no unique way to analyse complex biological systems. One approach is to treat the system as a network of interactions. Network components may be molecules, genes, cells and organisms that interact in time and space. The ensemble of possible interactions constitutes the network wiring and largely determines the behaviour of the system. Obviously, this set of interactions is far from constant in time and changes in response to internal and external cues, such as cell cycle control and hormonal signals. By their very nature, biological networks are made up of large numbers of often quite diverse components and interactions. Analysing them in a comprehensive way requires very large numbers of highly focused experiments. Our present predominantly small-scale type of research is not fit to do this job.

**Managing large-scale research efforts: combining experiment-based modelling with model-based experimentation**

How to manage and coordinate large-scale research efforts and how to keep it focused, synergistic and goal-oriented? Systems Biology offers a logical tool by making experiments centre around a computational model that combines all relevant information and that is able to make specific predictions that can be experimentally verified. Specific model-based predictions become the rational guiding principle for selecting the best, i.e. most informative, experiments. Based on the outcome of the experiments the model is updated and improved continuously. Such an iterative cycle is at the very heart of Systems Biology. So far, examples of systematic integration of modelling and biological experimentation in solving complex biological problems are sparse. However, for instance, in physics and engineering the exploitation of the iterative cycle of experiment-based modelling and model-based experimentation is quite common. Now that we have an extensive toolbox for carrying out biological experiments and we are beginning to obtain experience in computational modelling of complex biological systems, the above rational approach is also becoming feasible in biology.

A discussion that often comes up is that even our best models often are utterly incomplete and largely simply wrong. This does not invalidate the proposed approach. Predictive models at any time represent in an explicit and quantitative manner our best understanding of the system under study and therefore allow choosing the best experiments that we can define at a particular stage of a project. The key issue is that the predictive model is a tool that acts as a central guiding principle of large-scale research projects and that it is continuously updated and improved as knowledge increases. It is only eventually, as research progresses, that such models will begin to effectively describe the real world.

# C. Systems Biology: scientific views about what, why and how

**Problems that must be faced**

Successfully and cost-effectively addressing complex biological systems depends probably more on sociological issues than on scientific and technological problems. First of all, investigators must organise their research efforts in a highly focused and synergistic manner. At first sight this leaves less room for individual scientific creativity and may seem to reduce academic freedom. Whether this is true remains to be seen. Within the context of the system under study and of the model that acts as the guiding principle, scientific creativity and serendipity are still essential. However, success will often be the result of cooperative efforts of large numbers of investigators, rather than from individual research groups. We better get used to this.

An important issue is the standardisation of experimentation, making it possible to add up quantitative results obtained in different labs. Related to this is the standardisation of modelling, allowing the integration of different models at a certain stage. Dr Klingmüller addresses this in Chapter 8 of this report. A major challenge is the management of large biological programmes, i.e. to keep them focused, goal-oriented and cost-effective. First, it requires effective cooperation between different disciplines, including biological, biomedical and biotechnological investigators, those that develop the mathematical models (in practice often researchers with a background in physics) and engineers that bring in expertise in how to handle complex systems in general. This is far from trivial. Because of our monodisciplinary education system in universities and polytechnics, communication between these disciplines is difficult. Experience shows that it requires considerable time before barriers of language and culture between disciplines begin to disappear. Special teaching and training programmes will be required. In several places this process has already started. Second, starting from the guiding principle of predictive modelling, priorities must be set continuously, making research efforts as synergistic and cost-effective as possible. It will require vision, persuasiveness and other special skills to oversee such large collaborative research efforts. Third, data management requires special attention. Heterogeneous large data sets obtained by the participants must be checked, integrated and made accessible and at the same time must be integrated into the central model. Fourth, teaming up of academia and industry is a must for large-scale goal-oriented international programmes exploiting Systems Biology. This requires a re-evaluation of the issue of intellectual property. In large-scale international projects that are organised around computational models to which all partners contribute, it is by no means clear who owns what type of information or discovery. Finally, goal-oriented research programmes that concentrate on solving highly complex biological problems require a long-term vision and stamina. We may hope to really achieve breakthroughs by focusing on specific issues for a period of probably at least 10 years. Again, this will require vision and stamina that do not fit any of the funding systems we operate today. We have to find ways to make this possible.

**What complex biological systems should be studied?**

From a purely scientific point of view it makes not too much difference what the precise focus of large-scale research programmes is. Whatever the goal, it is very likely that the scientific challenge will be big and that many investigators will be interested in participating. Therefore, choices will be made based primarily on societal, economic or political grounds. It is likely that understanding the molecular and cellular bases of human diseases will be a priority. In particular, multifactorial diseases, metabolic syndrome, diabetes, specific types of cancer, infectious diseases and ageing all are keywords in the forthcoming discussion about priorities. Rational drug development, personalised medicine and health food are closely related issues.

**Conclusion**

The tools for the systematic and necessarily large-scale tackling of major biomedical and biotechnological issues are within reach. Systems Biology creates a framework to address and manage large-scale efforts by putting predictive computational modelling at the central stage. A number of major hurdles have to be taken, including major changes in how biological, biomedical and biotechnological research is organised and managed and how it is financed. Evidently, demonstration projects are needed to present proof-of-principle, showing that the proposed approach is feasible, productive and cost-effective. This ESF report should convince the European community that this is a timely thing to do and worth major investment.

## 4. Systems Biology and the bio-engineering of living organisms

Uwe Sauer
*Institute of Molecular Systems Biology, ETH, Zürich, Switzerland*

Different from scientifically accurate descriptions of biological systems and their molecular mechanisms, engineering of living organisms ideally requires a deeper understanding that entails the capacity to predict complex responses when entirely new functions or pathways are added and existing ones are removed. In the early 1990s, the concept of metabolic engineering was conceived as a systems approach to optimise biotechnologically desired properties of microbes and higher cells[1]. The key element was to supplant the gene- or molecule-centric concept that focuses on a particular gene or reaction with an emphasis on pathways in their entirety. Improvements of such more complex properties were envisaged through multiple gene manipulations by iterating a design–synthesis–analysis cycle, in which quantitative experimental analyses were to be combined with mathematical models for data integration and rational design[2,3].

While great strides have been made in the past 15 years, progress was largely empirical and confined to individual pathways or simple traits. Truly rational engineering of complex networks remained somewhat elusive. Besides the absence of sufficient biological knowledge and understanding, the lack of global analytical methods and of theoretical concepts in the face of an overwhelming complexity in even the simplest microbes were the main limitations. The post-genomic wave of cell-wide analytical methodologies spawned not only Systems Biology, but addressed also one of the key limitations of metabolic engineering. Consequently, both fields have cross-fertilised each other right from the beginning; for example by applying 'omics' methods in metabolic engineering[4] or, conversely, applying methods originally developed for metabolic engineering such as flux analysis[5] and genome-scale models of metabolism[6] in Systems Biology. Thus, metabolic and cellular engineering are basically applied Systems Biology[7].

The limitation of appropriate theoretical concepts, however, remained for some time. Experience from metabolic engineering already demonstrated that the isolated application of any global analytical technology alone does not suffice to significantly advance genetic manipulation of living cells to a rational engineering process. While countless DNA chip and somewhat fewer proteomics experiments have certainly increased our knowledge of involved genes and proteins (system components), they have not helped much in promoting our understanding of the multiple biochemical and regulatory interactions among genes, proteins and metabolites, and hence have not noticeably increased the success rate of engineering strategies. The underlying reason is, of course, the complexity of the biological systems that is based on the quantitative, often nonlinear and highly regulated interaction of these components with each other. It is precisely in this area where the strongest impact of Systems Biology is expected. How can the aim of Systems Biology, to understand the interaction between the components in biological systems and to quantitatively predict their behaviour as an assembled system, be achieved?

What has been limiting so far, in science and metabolic engineering alike, are methods for efficient data integration in computer models with appropriate mechanistic and molecular detail to enable in silico experiments of sufficient predictive capability. While there is clearly a need for more and quantitative data, at higher quality, with greater time and space resolution, at the single cell level and as integrated functional read-outs, the timing is more appropriate now then ever to develop and apply novel theoretical approaches. To this end, several promising Systems Biology approaches are currently being developed that fall generally into two categories:

(i) detailed molecular models of semi-independent biological modules that can eventually be assembled into more complete models, and
(ii) coarse-grained models that start from the top of network topology to make their way down to the molecular detail[8].

Although these approaches are far from accurate, significant progress has been made. Genome-scale models of microbial metabolic stoichiometry, for example already, achieve qualitative predictions of simple mutant phenotypes and provide an interaction framework for interpreting metabolite or protein data within their biological context[9].

Systems Biology is starting to generate, for the first time, a quantitative, molecular and cell-wide knowledge base that is also tremendously relevant for biotechnology. Clearly, individual Systems Biology-driven technological advances, experimental or computational, will have an immediate impact on cellular and metabolic engineering. The key element, however, is the unique combination of integrating experimental data with computational and theoretical methods that go beyond statistical correlations but are based on molecular reasoning and actual mechanisms. Progress, however, will not stop with the re-engineering of existing, highly interconnected cellular networks in a metabolic engineering framework, but will attempt to synthesise them de novo. Largely driven by Systems

Biology, such synthetic biology emerged more recently as an attempt to synthesise increasingly complex biological entities – regulatory switches, gene circuits and pathways for starters but eventually entire organisms – from standardised biological components in an engineering-inspired approach from scratch[10,11].

The increasing knowledge on molecular component interaction in microbes, plants and animals that is gathered from and integrated through Systems Biological-related research is already paving the way towards a more predictable and rational approach to cellular and metabolic engineering, as it was originally conceived by its founders in the early 1990s. This has great potential to fundamentally change the way green and white biotechnology is pursued by opening up entirely new options for the production of chemicals, food products and in plant breeding. By replacing the current gene-centric view with a systems-perspective on interacting elements, Systems Biology will also make major inroads into the development of a novel, personalised medicine that holds promise in combating and preventing complex diseases such as cancer and metabolic disorders that are beyond our current capabilities.

References

1. Bailey, J.E. 1991. Toward a science of metabolic engineering. *Science*. 252:1668-75.
2. Vemuri, G.N. and A.A. Aristidou. 2005. Metabolic engineering in the -omics era: elucidating and modulating regulatory networks. *Microbiol Mol Biol Rev*. 69:197-216.
3. Nielsen, J. 2001. Metabolic engineering. *Appl Microbiol Biotechnol*. 55:263-83.
4. Lee, S.Y., D.Y. Lee, and T.Y. Kim. 2005. Systems biotechnology for strain improvement. *Trends Biotechnol*. 23:349-58.
5. Sauer, U. 2006. Metabolic networks in motion: 13C-based flux analysis. *Mol Syst Biol*. 2:62.
6. Patil, K.R., M. Akesson, and J. Nielsen. 2004. Use of genome-scale microbial models for metabolic engineering. *Curr Opin Biotechnol*. 15:64-9.
7. Stephanopoulos, G., H. Alper, and J. Moxley. 2004. Exploiting biological complexity for strain improvement through Systems Biology. *Nat Biotechnol*. 22:1261-7.
8. Stelling, J. 2004. Mathematical models in microbial Systems Biology. *Curr Opin Microbiol*. 7:513-8.
9. Price, N.D., J.L. Reed, and B.O. Palsson. 2004. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol*. 2:886-97.
10. Heinemann, M., and S. Panke. 2006. Synthetic biology – putting engineering into biology. *Bioinformatics*. 22:2790-9.
11. Barrett, C.L., T.Y. Kim, H.U. Kim, B.O. Palsson, and S.Y. Lee. 2006. Systems Biology as a foundation for genome-scale synthetic biology. *Curr Opin Biotechnol*. 17:488-92.

## 5. Does modelling of complex biological systems require new types of mathematics?

Mats Gyllenberg
*Department of Mathematics and Statistics, University of Helsinki, Finland*

Life is highly complex. However, research in genetics and molecular biology during the past two decades has revealed detailed information about the basic constituents of life and the dynamical interactions between processes at different hierarchical levels. At the same time, both computer science and hardware technology have evolved rapidly providing us with machines with an enormous computing power. This development has made it possible to build realistic models of living cells and organs that can be run on a computer and simulate life in a quantitative fashion. The study of the aforementioned dynamical interactions has been termed 'Systems Biology' and in a relatively short time it has become one of the most computer-intensive disciplines. It is self-evident that successful research in Systems Biology relies on a collaboration between biologists and computer scientists. Involvement of physicists and chemists is frequently needed. But what is the role of the mathematician? Every computer implementation of a system, be it physical or biological, depends on an underlying mathematical model. However, mathematicians have, in most cases, been conspicuous by their absence in Systems Biological enterprises. Perhaps they are not needed?

It is a well-known fact that the mathematics of most physical theories has been developed on an intuitive, as opposed to a rigorous, basis. It has then been the task of mathematicians of later generations to fill in the gaps and make sense of the sloppy calculations of physicists. For example, physicists have used the Dirac 'function' and other 'generalised' functions long before Laurent Schwartz made these objects mathematically precise in the 1940s. One can argue that from the point of view of physics, the real insight was made when the new objects were introduced and understood intuitively by physicists and that the later mathematical justification did not bring anything new to our understanding of the physical world. Is something similar true regarding mathematics in the emerging Systems Biology?

A common misconception is that a mathematical model is a system of equations describing the behaviour of a real, natural system. If sufficient data is available, one can tune the parameters of the model and obtain a good fit of the model to the data. Once the parameters have been chosen in such a way, the model can subsequently be used to predict the behaviour of the system in the future. The abovedescribed view of mathematical modelling is naive, to say the least, and

simply wrong. A mathematical model is a mathematical structure (e.g. a system of differential equations or a Markov process) together with an interpretation rule. If the interpretation of the proposed equations leads to biological absurdities (such as negative values of variables that by their very nature can take on only positive values), then the whole structure must be disqualified as a model. A good mathematical model of a real world system should be able to:
(i) describe quantitatively,
(ii) predict, and
(iii) explain
the behaviour of the system. Moreover, it should be mechanistic because the ultimate goal of Systems Biology is to understand the mechanisms at a lower hierarchical level (e.g. the cell) that lead to functional behaviour at a higher level (e.g. an organ). In fact, in most cases a model must be mechanistic in order to fulfill condition (iii) of explaining power.

The first step in modelling is to clothe the biological problem in a mathematical guise. This step is already important. Our everyday language is not sufficiently precise to express and analyse scientific data and knowledge. This can be done only with mathematics. But the use of mathematics forces us to be precise about assumptions and implications and therefore this step can give significant insight. Modelling might seem easy, but in fact it requires expertise exactly as all other endeavours in science. Ideally, it should be carried out in collaboration with biologists (and/or physicists, chemists etc.) and mathematicians. The biologists contribute by their knowledge of the true system and its mechanisms while the mathematicians translate this into mathematics, build models and check their consistency. Often the modelling process gives rise to new questions that force the biologist back to the laboratory to perform new experiments. At its best, modelling thus leads to the discovery of new empirical results.

The following is an example of how one can build models of Systems Biology. We consider the growth of a solid tumour. The tumour consists of cells. Let us assume that we know how the individual cells behave, given the environmental conditions. Here the environment of a cell is built up by all the other cells in the tumour and is also affected by blood flow etc. By saying that 'we know how cells behave' we mean that we have submodels for how cells grow, progress through the cell cycle, go into quiescence, divide, etc. The tumour model then tells us how the behaviour of the individual cells is reflected in the behaviour of the tumour as a whole. But the story does not end here, because in reality we do not know in advance how the cells behave. In fact their behaviour depends on a complicated nonlinear feedback from the level of the tumour to the level of the cells. For instance the size of the tumour, the presence of a necrotic centre, the degree of vascularisation, etc. affect the behaviour of the cells.

The example of tumour growth mentioned above is an instance of modelling-structured populations. A tumour is of course nothing but a population of cells. Such models have been considered using partial differential equations since the 1960s[1,2] and more recently using abstract integral equations[3]. Many mathematical techniques have been developed to analyse such models but still a lot of questions remain unanswered. It should be noted that the tumour model described above has only two hierarchical levels: the cell and the tumour. In Systems Biology we want to dig deeper. The cell is not the basic entity of a tumour; it forms itself a system with many sublevels. Nothing is known about how to treat such multilevel systems mathematically.

When modelling biological systems the focus should always be on the biological situation. The mathematical models and methods must be chosen to catch the essential nature of the system. Unfortunately, this is not always the case. Scientists are equipped with a certain 'toolbox' of mathematical method that they master and they tend to chose their models so that they can apply the tools in their box. Differential equations have formed the main modelling tool ever since the days of Newton. But what is natural for modelling planetary motion is not necessarily the right thing for intra- and intercellular interactions. It is, of course, natural to model enzyme kinetics and different pathways using differential equations, but when all kinds of dynamical interactions at different hierarchical levels of a biological system are put together, one arrives at a huge system of differential equations. In the literature, I have seen systems with hundreds of unknowns and tens of thousands of parameters to be determined. It is obvious that with such a system one can produce any kind of behaviour one wants to. But it is more than unclear how such a model could give any mechanistic explanation of the real biological system. The well-known quote by John von Neumann[4] comes immediately to mind: 'With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.'

So called 'complex' systems have been the object of intensive mathematical research during the past three decades or so. The perhaps most intriguing fact is that very simple models such as a scalar difference equation or a differential equation in three dimensions can give rise to extremely complicated or 'chaotic' dynamics. Chaotic systems may seem to be indistinguishable from random time series, but in fact they exhibit some sort of regularity in their statistic properties. Much of the modern theory of dynamical systems deals with the uncovering of such regularities. In Systems Biology we meet another kind of complexity. Here, as we have seen above, the models have a complex structure. So

if simple models can produce complex dynamics, then complex models are expected to produce complex behaviour as well. The characterisation of this behaviour is a great challenge for mathematicians.

In the beginning of this chapter I mentioned that the mathematics needed for major breakthroughs in physics has often been developed rather sloppily with a rigorous mathematical justification left to later generations and I pondered whether the same might be true for the mathematics of Systems Biology. There is a difference between physics and biology. The basic laws of physics take on a comparatively simple form that often can be derived in an intuitive fashion without rigorous mathematics. The immense complexity of biological systems requires real mathematics already at the modelling stage. Mathematics should not be considered as a toolbox that already contains everything that one could need. Modelling of new biological phenomena will most likely require the development of entirely new mathematics.

## References

1. Tsuchiya, H.M., Fredickson, A.G, and Aris, P. 1966. Dynamics of microbial cell populations, *Advan. Chem. Eng.* 6:125-98.
2. Bell, G.I., and E.C. Anderson. 1967. Cell growth and division. I. A mathematical model with applications to cell volume distributions in mammalian suspension cultures. *Biophys J.* 7:329-51.
3. Diekmann, O., Gyllenberg, M. and Metz, J.A.J. 2007. Physiologically structured population models: Towards a general mathematical theory. In *Mathematics for Ecology and Environmental Sciences* (K. Sato, Ed.), Springer, Heidelberg, Springer, Berlin-Heidelberg (2007), pp. 5-20.
4. Dyson, F. 2004. A meeting with Enrico Fermi, *Nature* 427, 297.

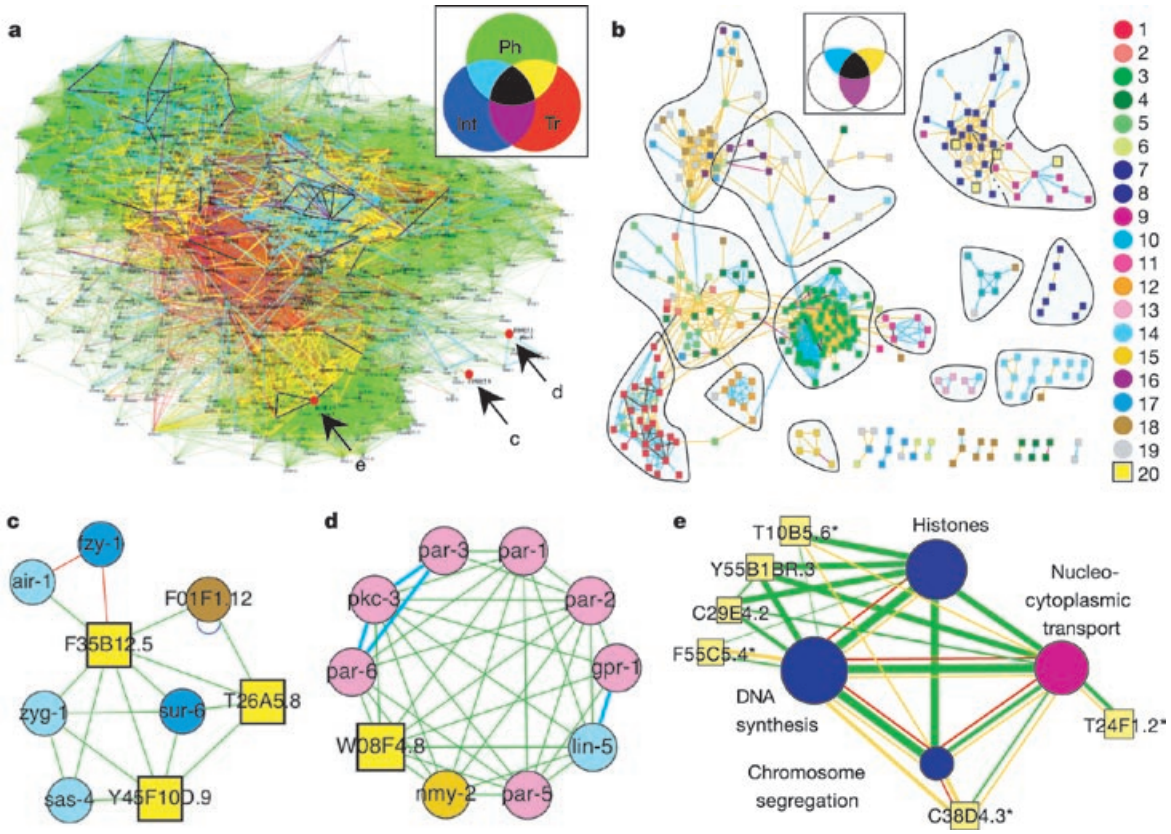



Figure 5.1: Integrated network analysis of a) entire early embryogenesis, b) multiple support network containing 305 nodes, c) Centrosome model; d) PAR cell polarity model; e) nuclear function model. Reprinted by permission from Macmillan Publishers Ltd. Gunsalus K.C., Ge, H., Schetter A.J, Goldberg D. Han J.-D.J., Hao T., Berriz G.F., Bertin N., Huang J., Chuang L.-S., Li N., Mani R., Hyman A.A., Sönnichsen B., Echeverri C.J., Roth F.R., Vidal M. and Piano F. (2005) Predictive models of molecular machines involved in Caenorhabditis elegans early embryogenesis. *Nature 436:861*

## 6. Engineering approaches: what can we learn from it in Systems Biology?

Olaf Wolkenhauer
*University of Rostock, Germany*

The predefined title of this short chapter is misleading: Systems Biology *is* the merger of systems theory (engineering approaches) and molecular/cell biology. While bioinformatics has been closely associated with the field of computer science, Systems Biology is particularly attractive to researchers from the engineering and physical sciences. What this indicates is that an 'engineering' or 'systems-theoretic' approach is different from the way cell biological systems have been studied up to now. The emergence of Systems Biology is, in part, a consequence of the limitations we have reached in genomics and bioinformatics. While those areas are a different approach to investigating cellular processes, they are clearly complementary: an engineering approach relies on information about which macromolecules do matter in any particular cellular process and what their physical characteristics are. What this hints at is the fact that a signal and systems-oriented engineering approach is quite demanding in terms of the data it requires. As I shall argue below, the engineering approach is a necessity, not a choice, if we are to understand the functioning of the cell. The most important question we should therefore ask is 'What is necessary to ensure systems-theoretic approaches can work?'

While genomics and bioinformatics have focused on an effort to identify and catalogue the components that make up the cell, including their molecular characterisation and study of associations, the signal and systems-oriented perspective of Systems Biology focuses on functional activity, that is the dynamics of intra- and intercellular processes that determine cell function. An engineering approach is a 'way of thinking'. What this means and what we can or cannot learn from the engineering approach is discussed below. An engineering or systems-theoretic approach is characterised by the use of mathematical models. The important role mathematical models play is a consequence of the complexity of cellular processes, specifically the large number of variables, nonlinear interactions and temporal processes. Mathematical models are the extended arm of common sense; the only means we have to deal with non-intuitive complexity – no more but also no less.

In an article in *Current Biology* (Vol.15, No.21, 2005), Ronald Plasterk criticised the engineering approach and argued that:

*'None of these modellers ever predicted that small microRNAs would play a role. One makes discoveries by watching, working, checking. They want to be Darwin, but do not want to waste years on the Beagle. They want sex but not love, icing but no cake. Scientific pornography'.*

While a mathematical model (or more precisely the mismatch between a model and experimental data) can indicate whether additional variables or others than those selected, should be included in the model, knowledge of the components, and to some extent information about their molecular characteristics, must be available before we can establish a model of a dynamic system. A modeller could never predict that microRNAs would play a role; instead, the purpose of the model is to elucidate what role components have in the functioning of the cell. A mathematical model is used to characterise the function a component may have in the regulation and control of a processes, say gene expression. A model and computer simulation helps to validate hypotheses about the dynamic properties of a system and mechanisms (feedback interactions) that give rise to the behaviour observed in experiments. System biologists are interested in the consequences of dynamic interactions and perturbations, that is, how spatio-temporal changes in molecular concentrations determine cell function, including differentiation, apoptosis, proliferation etc. Plasterk apparently did not understand the role of models and modellers:

*'One makes good models by watching gene expression, working on improved designs for experiments, checking hypotheses encoded by models. Modellers want quantitative data, but do not want to waste years in the lab (for which they are not trained). They want collaborations but not ignorance, support the experimentalist but not replace him. Interdisciplinary research'.*

A cell, organ, or organism, understood as a 'system', is a network of components whose relationships and properties are largely determined by their function in the whole. The functionality is observed as the 'behaviour' of the system. The first and probably most important lesson of systems theory is that we can understand the behaviour of a system only if we systematically perturb it and record its response. A systems approach is thus characterised by input/output descriptions and from this, the most important role of the modeller in Systems Biology is to support the design of stimulus/response experiments. The role of nonlinear systems and control theory is then to provide methodologies to encode interactions of genes/proteins in the structure of the mathematical equations that form a model. The terms of these equations will reflect such processes as (de)activation, dimerisation, (de)phosphorylation, while the signs of these terms can indicate synthesis, degradation, positive or negative feedback relations. Parameter values emphasise terms and relate to the particular experimental set-up, cell type or cell line.

# C. Systems Biology: scientific views about what, why and how

Taking cell differentiation as an example, the decision whether a cell differentiates or not will depend not only on the presence of a protein, whether a gene is 'on' or 'off', but on the history of various interacting proteins, a process referred to as a 'bistable system'. Even the simplest three-component model of such a process demonstrates that the observed behaviour can be understood only through experiments that vary not only initial conditions but also the duration and level (profile) of the stimulus. The system-theoretic concepts of 'identifiability', 'distinguishability' and 'observability' are important concepts in this context. The analysis of a model may reveal that there are multiple sets of parameter values that can all reproduce the same input-output behaviour. An improved design for the experiment may either remove this ambiguity or at least has the analysis alerting us to the uncertainty that can arise from such a situation. Closely related is the question of whether a given experiment would be capable of distinguishing between two hypothesised alternative mechanisms (model structures) that could generate the observed phenomena. What this discussion leads to is that experiments in Systems Biology tend to be more expensive and more time consuming. However, there is no alternative if we accept that in cells we are dealing with nonlinear dynamics. A consequence of this view is that research funding practices should appreciate the need for 'theoretical work', developing systems-theoretic methodologies, and that consumables budgets can increase if one generates quantitative time course data (including experiments to establish standards, normalise data and replicates to remove non-biological variability in measurements).

What the modeller describes as 'bistability', leads to switching-type behaviour; and an important task of Systems Biology is to identify functional units (subsystems) that realise such 'dynamic motifs', including for example 'oscillations', 'amplification', 'hysteresis' or

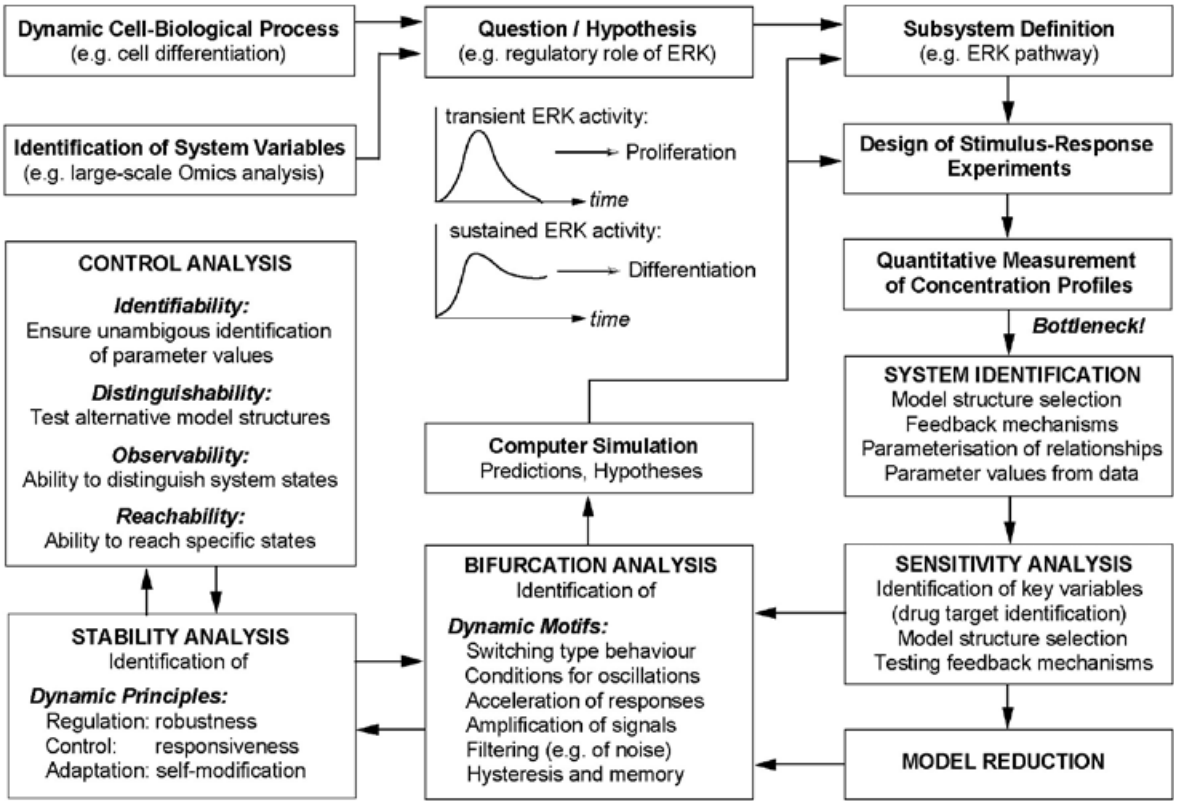



Figure 6.1: The (control) engineering approach as a Systems Biology workflow: merging cell biology with systems theory to study the functional organisation of cells, i.e., cell function understood as inter-and intracellular dynamic processes. The insert about the ERK signal transduction pathway is to provide an example in which the history of a signal (and not only the presence of a gene/protein) matters for the cellular process which decides upon the executed cell function. The role of an engineering approach is to elucidate the mechanisms (in particular feedback interactions) responsible for such observed phenomena: For nonlinear processes like the one above, these can be understood only with the help of mathematical modelling.

'homeostasis'. Stability and bifurcation analysis are important tools for this task. Under the heading of 'system identification', the control engineer gathers tasks for parameter estimation and model structure selection. Once a (preliminary) model is found, parameter sensitivity analysis allows us to identify the influence variables have on the overall behaviour of a network. This enables an investigation into the robustness and/or responsiveness of a system and is also a natural starting point to discuss potential drug targets. Many of the existing systems-theoretic techniques are not well suited for short time series, uncertainty in data and for systems involving many variables: there is a need for basic research to develop new methodologies. Systems Biology is not the application of existing engineering tools to cell biology but a merger of both fields; both fields should co-evolve.

The aim of Systems Biology is to understand the relations between things such as molecules or cells, not the things in themselves. Cell function arises from interactions between molecules and is not a property of any one molecule. The engineering perspective of Systems Biology is thus characterised by a shift towards an understanding of functional activity, away from the identification, molecular characterisation and cataloguing of the components that make up the cell. The complexity and limitations of Systems Biology are primarily a consequence of a large number of variables, interacting in space and time in a nonlinear fashion. Because of limited time frames for projects, funding constraints and also technological limitations that prevent us from quantifying large numbers of gene/proteins in time course experiments (at different levels of scale), a dynamic model of a pathway is necessarily 'wrong' – a phenomenological representation of a hypothesised principle that governs observed phenomena. Mathematical modelling is therefore the *art* of making appropriate assumptions; a process by which we represent one thing by another because understanding consists of reducing one type of reality to another. The purpose of modelling is therefore abstraction: the reduction of a complex reality to essential features. But even if inaccurate in this sense, a model can be useful by guiding the experimentalist in the design of his experiments, helping in the decision as to which variables to measure and how.

An important role of the modeller is therefore his/her involvement in the design of experiments. An advantage engineers and physicists have in this is that in addition to their analytical skills, they are not afraid of getting their hands dirty with experimental data. The sceptical wet-lab scientist may find that even if a mathematical model is a long way off, engineers and physicist can be helpful allies in understanding the physical properties (specifications) and limitations of measurement devices (e.g. its linear range, reproducibility, accuracy, etc.). Being able to quantify the accuracy and variability of instruments is an important step in interpreting experimental data. The real bottleneck for a success of engineering approaches in Systems Biology is advances in the generation of quantitative and sufficiently rich time series data sets. Progress in Systems Biology will depend on improved technologies that can quantify temporal changes in stimulus-response experiments. This can be done only in close collaboration with the engineering and physical sciences. What we can learn from engineering approaches is that measurement technologies to generate data and methodologies to analyse data cannot be separated.

## 7. The role of information technology for Systems Biology

Heikki Mannila
*Helsinki Institute for Information Technology (HIIT), Helsinki University of Technology and University of Helsinki, Finland*

Systems Biology aims at understanding biological entities at the systems level. To do this, we have to be able to observe many parts of the structure and dynamics of the entity, store and transform data, link it with many other types of observations, and model the data. Consider, for example, the task of modelling the behaviour of a single cell. We can obtain direct and indirect information about the genome of the cell, gene expression under experimental conditions, the metabolites and pathways in the cell, etc. Managing this data and using it to build a useful model of the cell will require huge advances also in information technology. Information technology is vital for Systems Biology: it is needed in measurement, in management and curation of the data, and in data analysis. Existing methods are not going to be enough, as Systems Biology poses unprecedented challenges to all these areas.

**Measurement, storage and retrieval, and analysis**

One key factor in the rise of Systems Biology is the rapid development of *measurement* technologies. We can measure many aspects of the operation of biological systems with high accuracy and in tremendous volume. The advances in high-throughput measurement techniques such as microarray methods have required many innovations from information technology.

*Data management and curation* are crucial for the accurate analysis of any larger mass of observational data. Especially in Systems Biology we have to understand well the conditions under which the data have been collected, otherwise the prediction of complex cellular functions cannot be achieved.

# C. Systems Biology: scientific views about what, why and how

Data management for Systems Biology requires that we can answer questions such as the following. How are large masses of data stored so that desired data can be retrieved efficiently? How can we keep track of the different versions of the data? How can we locate and correct errors in the data? What metadata, i.e., data about the data itself, should be stored, and how can it be used?

The genome projects are spectacular examples of the success of data management methods. Work done at NCBI and EBI are prime examples of how large volumes of data can be made available and usable. However, Systems Biology presents new, more difficult challenges to data management methods. Genome data is actually fairly simple compared with the multitude of information that the other 'omics' are yielding.

Biological data creates large demands for *data analysis*. The volume of data is large, and, more importantly, the dimensionality is often much higher than in traditional data analysis tasks. Moreover, the novel types of data pose new challenges to data analysis.

**The genome was easy**

The move from genome projects to Systems Biology will also require huge advances from information technology. One could even claim in hindsight that sequencing the human genome was an easy task for IT.

The genomic data is discrete and the error rates in sequencing are quite small and reasonably well understood. Moreover, the genome is a slowly changing object, with differences between individuals being mostly ignored in the initial sequence projects. Thus the goal in the genome projects was to find a *static* object. Partial descriptions of the genome are useful: if we know the first half of the genome, the work is half done; hence a piecemeal approach can be used.

Data analysis of genomic sequences is not easy, but the handling of sequence data has been one of the success stories of computer science in biology. Ingenious algorithms have been developed for tasks such as fast approximate string matching, sequence assembly, and for fitting hidden Markov models. These algorithms are in wide use in genome sequencing, search, and gene prediction[1,2]. Similarly, in gene mapping the data are (in principle) fairly simple: information about the phenotype, pedigrees, and the genetic markers. The task is to find good methods for predicting the locations of the possible genes influencing the phenotype. This goal turns out to be hard to achieve, but statistical genetics has been very successful in devising advanced methods that basically are able to extract all possible information from the available data. Systems Biology will require a similar concentrated effort in data analysis methods, but the multitude of data types will create novel difficulties.

**Complex new data: managing and using multiple data sources**

Compare sequence analysis or gene mapping with trying to understand how a single cell works (let alone a whole multicellular organism). A cell is a *dynamical* system, and in principle each component or metabolite can interact with all other components. We have to understand cell function resulting from interactions between molecules, with some reactions happening quickly, some slowly. The genome of the cell is a single object that we can (in principle) describe completely, but the dynamics of the cell are much more complex. Measurements give us only snapshot information about the state of the system. Partial information about the system can easily overlook some crucial data, and the selection of the correct temporal and spatial granularity is a hard task.

As an example, microarray data about gene expression is a snapshot of some aspects of the dynamical system. The errors in the measurement can be large, and it is hard to say whether the measurement conditions reflect the true biological environment. To compare different sets of gene expression measurements we have to be able to compare the measurement conditions. Thus metadata has a larger role for microarray data than for genome data[3]. The difference between static and dynamic data is even clearer when considering, say, data about signal transduction pathways.

Systems Biology will need many different types of static and dynamic data. To use such heterogeneous data sets we need to have ways of describing what the data is and how it can be used. Thus the development of different description languages for data is important; many interesting efforts are under way[4].

The multitude of data types and sources implies that many issues in data management analysis have to be solved. The basic problem of storing the data can be solved reasonably easily: while high-throughput biological analysis methods produce lots of information, storing it using database methods will probably be straightforward. However, storing spatial and temporal information poses new challenges, both in describing the semantics of the data and in implementing efficient retrieval of the data.

It is much harder to describe general ways of linking information together. How do we know in general that two sets of observations describe the same situation, and that they can be used in the same analysis? As we are studying a dynamical system by using snapshot information, how do we guarantee that the snapshots are actually from the same state? Data management techniques and tools will have to improve in order to address these issues.

**New data analysis methods**

The new types of data immediately lead to new data analysis tasks. How can we connect together different measurements of the function of the cell and use them to construct a predictive model? Given the plethora of data types, building models is going to be extremely hard.

Even simple cases of novel biological data can lead to difficult analysis problems. As a small example, in microarray data one can search for groups of genes and samples such that within the set of samples the expressions of the genes are strongly correlated. Such bi-clustering or co-clustering tasks turn out to have very interesting computational properties; a definite treatment is still lacking. The analysis is complex even though no dynamical behaviour is being modelled. Finding dynamical models and interaction models for the cell or organism will certainly require the solving of a multitude of hard computational problems.

One of the key challenges will be to develop the statistical and computational methods for keeping track of the different levels of uncertainty and errors in the data and in the analysis results. When we know that the original measurements have errors in a certain range, how trustworthy are the final results? Strong interaction between computer science, statistics and mathematics is going to be needed to answer such questions.

The high dimensionality of data obtained from high-throughput techniques poses challenges for determining the statistical significances of the discoveries. An SNP chip with 100000 or 500000 markers will clearly yield huge amounts of spurious correlations with any phenotype for the sample sizes that are economically feasible. Thus techniques for handling multiple testing are needed; as the distributional results are often hard to obtain, randomisation techniques must be used.

Randomising SNP data is not necessarily easy, but in Systems Biology the task will be much harder. Imagine a database with many different types of data, and a complex data analysis task that obtains an interesting looking result by combining these results in some way. How do we generate random instances that share the crucial properties of the original data set? The problem is difficult. Randomisation methods are computationally intensive, but typically they are relatively easy to parallelise. However, conceptual and algorithmic developments are still needed.

**Where are the bottlenecks?**

Systems Biology will need information technology. Where are the bottlenecks? Is computational power going to be the factor that limits the development of Systems Biology, or are the main problems in using IT elsewhere?

While many aspects of Systems Biology will certainly lead to heavy use of computing resources, many tasks can be handled by using fairly inexpensive clusters (throughput computing), as opposed to tightly coupled expensive multiprocessors (capability computing). It seems that while lots of raw computing power is going to be needed, even greater challenges for IT are elsewhere.

The main bottleneck in the role of information technology in Systems Biology is going to be in the development of data management and data analysis techniques and software. Even though database management systems are already quite powerful, we will need even better ways of managing the large masses of data with spatial and temporal components, and with possibilities for describing the uncertainty in the data. Equally important, coming up with novel algorithms for computing interesting quantities, for fitting models to data, and for assessing the significance of the discoveries will be crucial for Systems Biology.

References


1. Richard Durbin, Anders Krogh, Graeme Mitchison, Sean R. Eddy. 1997: *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids.* Cambridge University Press.
2. Dan Gusfield: 1997. *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*. Cambridge University Press.
3. Brazma A., Hingamp P., Quackenbush J., Sherlock G. et al. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics*. 29: 365 - 71.
4. Hucka M., Finney A., Sauro H. M., Bolouri H., Doyle J. C., Kitano H. et al. 2003. The Systems Biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*. 19, 4: 524-31.
5. www.sbml.org


## 8. The need for standardisation in Systems Biology


Ursula Klingmüller, Marcel Schilling, Sebastian Bohl and Andrea C. Pfeifer
*Group: Systems Biology of Signal Transduction, DKFZ, Heidelberg, Germany*


Cellular decisions are regulated by a complex interplay of signalling pathways activated in response to ligand binding to cell surface receptors. The components of several signalling pathways have been studied in great detail but it remains largely unknown as to how information is processed and how biological responses are coordinated. By combining mathematical modelling with empirical data, Systems Biology aims at understanding general systems properties and predicting the effect of perturbations. However, the majority of data currently available is of a qualitative nature

# C. Systems Biology: scientific views about what, why and how

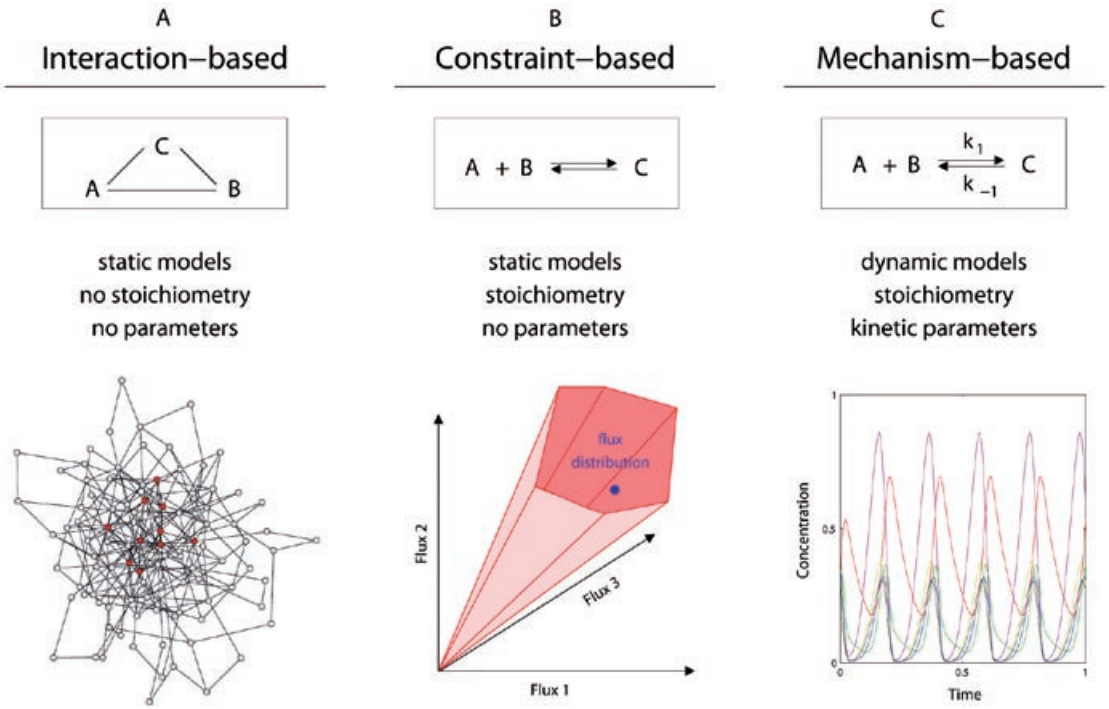



Figure 8.1: Approaches to the mathematical modelling of cellular networks. Mathematical models can start from network representations based on a) interactions alone, b) constraints including network topology, stoichiometries and reaction reversibilities, or c) detailed reaction mechanisms. Reprinted with permission from Elsevier Publishing Company. Stelling G. (2004) Mathematical models in microbial Systems Biology. *Current Opinion in Microbiology 7:513*

and has been generated by making use of cell lines. Conflicting results obtained underscore the importance of the cellular background and the need to work in a standardised cellular setting. Furthermore, current experimental techniques primarily aim at the identification of components and their interactions. These techniques need to be advanced and standardised to permit the generation of high quality quantitative data with sufficient temporal and spatial resolution for data-based dynamic pathway modelling.

**Standardisation of the cellular background**

Currently, several Systems Biology networks are being established or are already in operation such as HepatoSys, FORSYS, SysMO, YSBN, NucSys and Systems X. The success of these networks strongly depends on the transferability of the experimental data from the different participants. Therefore, standardisation of the biological material studied is essential. The requirements in standardisation are of varying complexity depending on the cellular context of interest. Systems Biology approaches in *E.coli*, yeast or eukaryotic cell lines should be based on identical strains or on a defined common source for a cell line. Furthermore, cultivation conditions and number of passages should be defined in detail. Even more complex are the requirements for the work with primary mammalian cells. Genetic heterogeneity is a major problem, thus inbred strains should be used, if possible, as a source for primary cells. Furthermore, age, sex and nutritional status should be defined. An example for successfully agreeing on an inbred mouse strain and a common standardised procedure for the preparation of primary cells is the consortium HepatoSys that studies complex cellular processes facilitating regeneration, detoxification and differentiation of hepatocytes using a Systems Biology approach[1,2].

**Standardised quantification of proteins and their modifications**

Signal transmission in a cell depends on both the stoichiometry of the communicating components and the dynamic changes in their transient modifications such as phosphorylation. The most widely applied method used to examine protein-protein interaction and transient modification is immunoblotting. This technique involves separation according to the molecular weight in a polyacrylamide gel and transfer to a membrane followed by antibody detection. Systematic evaluation of the sources for major errors in this technique

as well as the development of strategies for error correction and computer algorithms for automated data processing were established to reliably generate data for Systems Biology approaches by quantitative immunoblotting[2,3]. A simple yet useful strategy to improve data quality in immunoblotting involves randomisation of correlated samples reducing the error in quantitative immunoblotting by about 45%[2,3]. Further error reduction can be achieved by normalisation procedures. Normalisation involves calibrating measurements against a set of standards that are included with each run and then adjusting raw values to that of the standard[4]. As standards, 'housekeeping' proteins can be used to normalise immunobloting data. For immunoprecipitation-based approaches, calibrator proteins have been proven useful for normalisation – recombinant proteins that are recognised by the same antibody as the protein of interest. However, it is important that both calibrator or normaliser are of similar molecular weight as the protein of interest. Furthermore, a reliable and unbiased computer algorithm is necessary for automated data processing[3]. Using such computer algorithms, data of individual experiments can be merged and analysed together. Standardisation and automation of immunoblotting procedures will contribute to larger and more accurate data sets, fulfilling Systems Biology's demand for high quality quantitative data. Nevertheless, the number of samples that can be processed in parallel by quantitative immunoblotting remains limited. Thus, it is important to advance medium or high-throughput techniques for quantitative data generation. Considerable effort is being invested in the development of protein arrays and most promising are sandwich strategies using affinity-purification in combination with antibody-based detection in the microarray format[5]. The advantage of protein arrays is that small sample volumes can be analysed. However, standardised strategies are required to determine the specificity of the antibodies used and their extent of cross-reactivity. To generate absolute numbers quantitative proteomics holds great promise. To facilitate the quantitative analysis by mass spectrometry labelling with non-radioactive amino acids applying the Stable Isotope Labelling by Amino acids in Cell culture (SILAC) method[6] or iTRAQ based labeling[7] are being applied. Although recent developments show the power of mass spectrometric measurements of the phosphoproteom[8], standardised procedures for absolute quantification remain to be developed.

**Quantification of target gene induction**

Transcriptional regulation represents one of the most complex and important mechanisms in the processing of biological information. In the post-genomic era microarray and quantitative RT-PCR enable the generation of high throughput quantitative data of gene transcription. However, the data currently generated is affected by high background and noisy signals, as well as high biological variability. Thus, the transferability, reproducibility and validation of such data require improvement. Standardised protocols for the isolation of RNA and DNA as well as amplification and hybridisation are necessary and have been promoted by the consortia MAQC[9] and MIAME[10]. One of the most prominent advantages as well as challenges is the enormous amount of data generated by microrrays. In the last years several efforts were made and are still ongoing to improve and standardise the acquisition, storage, normalisation and analysis of microarray data (www.Bioconductor.org). Quantification of mRNA by quantitative Real Time (qRT)-PCR addresses a limited number of target genes compared with microarrays but enables a higher temporal resolution because of its lower costs. Although it is often described as a 'gold standard', it is far from being a standardised assay. The significant problems caused by variability of RNA templates, assay designs and protocols, as well as inappropriate data normalisation and inconsistent data analysis, are widely known but still require more effort[11].

**Standardisation of live cell imaging data**

Solving biological questions at a systems level not only requires high-quality temporal data, but also time-resolved data on the spatial relationships of proteins in single cells. While biochemical techniques usually average over the dynamic behaviour of large cell pools, live cell imaging can be applied to quantitatively follow the dynamics of protein localisation, concentration and interaction in single cells.

To generate high-quality microscopy data from living cells the choice of an appropriate cell system is essential and the appropriate behaviour of the pathway under investigation in this cell system should be validated by complimentary methods such as quantitative time-resolved biochemistry. Similar standards should apply for labelling the protein of interest for detection in live cell imaging. Green fluorescent proteins (GFP) and their colour variants have been widely used[12]. However, extreme care has to be taken when using fluorescent protein fusions. The functionality of the GFP chimera in all aspects important to the questions asked has to be carefully confirmed by biochemical methods. Furthermore, to avoid over-expression artefacts the fusion protein should be expressed at or close to endogenous levels. An important tool to achieve this is the use of an inducible system such as the Tet-system[13].

Imaging living cells deals with a very complex system that is disturbed by many factors, such as the

expression of a tagged protein, exposure to high-power light sources and suboptimal culture conditions. All these factors have to be controlled as much as possible to obtain meaningful quantitative data. During the experiment, the cells should be maintained at appropriate culture conditions. Ideally, the stage is surrounded by an incubation chamber that is capable of keeping the temperature, $CO_2$ level and humidity stable over many hours. For shorter experiments this could be replaced by a buffered medium and a heating device for the microscope stage, but the effects on the process have to be carefully monitored.

Unlike biochemical experiments that average signals of a large pool of cells, only one or a few cells are followed over time in a live cell imaging experiment. Therefore, it is necessary to perform a sufficient number of experiments to be able to draw a statistically significant conclusion from the data. Automated acquisition software with good algorithms to control x- and y-position of the stage as well as autofocus in combination with a motorised stage facilitate standardised data generation. While most modern microscope systems have their own acquisition and analysis software packages that fulfil the basic requirements for analysis, certain experimental setups require further analysis tools such as automated pattern-recognition software and particle-tracking algorithms. In the future, these software tools will become increasingly important to the generation of informative data for Systems Biology approaches.

**The future**

Currently, Systems Biology approaches suffer from a lack of high quality quantitative data available for mathematical modelling. To circumvent this, techniques for quantitative data generation have to be developed or have to find wider application. To ensure transferability of data and to avoid problems arising from heterogeneity of the biological background, standardisation of the biological system under investigation and the methods applied is absolutely critical. Therefore, standard operating procedures need to be developed. The establishment of a central data management structure will apply pressure to follow standard operating procedures and thereby will facilitate the transferability of the data generated. Thus, standardisation efforts will critically determine the success of Systems Biology initiatives.

References

1. Klingmüller, U., A. Bauer, S. Bohl, P.J. Nickel, K. Breitkopf, S. Dooley, S. Zellmer, C. Kern, I. Merfort, T. Sparna, J. Donauer, G. Walz, M. Geyer, C. Kreutz, M. Hermes, F. Gotschel, A. Hecht, D. Walter, L. Egger, K. Neubert, C. Borner, M. Brulport, W. Schormann, C. Sauer, F. Baumann, R. Preiss, S. MacNelly, P. Godoy, E. Wiercinska, L. Ciuclan, J. Edelmann, K. Zeilinger, M. Heinrich, U.M. Zanger, R. Gebhardt, T. Maiwald, R. Heinrich, J. Timmer, F. von Weizsacker, and J.G. Hengstler. 2006. Primary mouse hepatocytes for Systems Biology approaches: a standardized in vitro system for modelling of signal transduction pathways. *Syst Biol (Stevenage)*. 153:433-47.
2. Schilling, M., T. Maiwald, S. Bohl, M. Kollmann, C. Kreutz, J. Timmer, and U. Klingmüller. Quantitative data generation for Systems Biology – the impact of randomisation, calibrators and normalisers. *IEE Proc. Systems Biology* (2005), 152:193-200.
3. Schilling, M., T. Maiwald, S. Bohl, M. Kollmann, C. Kreutz, J. Timmer, and U. Klingmüller. 2005. Computational processing and error reduction strategies for standardized quantitative data in biological networks. *Febs J*. 272:6400-11.
4. Albeck, J.G., G. MacBeath, F.M. White, P.K. Sorger, D.A. Lauffenburger, and S. Gaudet. 2006. Collecting and organizing systematic sets of protein data. *Nat Rev Mol Cell Biol*. 7:803-12.
5. Nielsen, U.B., M.H. Cardone, A.J. Sinskey, G. MacBeath, and P.K. Sorger. 2003. Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc Natl Acad Sci USA*. 100:9330-5.
6. Ong, S.E., I. Kratchmarova, and M. Mann. 2003. Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res*. 2:173-81.
7. Schmidt, A., J. Kellermann, and F. Lottspeich. 2005. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics*. 5:4-15.
8. Olsen, J.V., B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. 127:635-48.
9. Shi, L., L.H. Reid, W.D. Jones, R. Shippy, J.A. Warrington, S.C. Baker, P.J. Collins, F. de Longueville, E.S. Kawasaki, K.Y. Lee, Y. Luo, Y.A. Sun, J.M. Willey, R.A. Setterquist, G.M. Fischer, W. Tong, Y.P. Dragan, D.J. Dix, F.W. Frueh, F.M. Goodsaid, D. Herman, R.V. Jensen, C.D. Johnson, E.K. Lobenhofer, R.K. Puri, U. Schrf, J. Thierry-Mieg, C. Wang, M. Wilson, P.K. Wolber, L. Zhang, and W. Slikker, Jr. 2006. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol*. 24:1151-61.
10. Brazma, A., P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, J. Aach, W. Ansorge, C.A. Ball, H.C. Causton, T. Gaasterland, P. Glenisson, F.C. Holstege, I.F. Kim, V. Markowitz, J.C. Matese, H. Parkinson, A. Robinson, U. Sarkans, S. Schulze-Kremer, J. Stewart, R. Taylor, J. Vilo, and M. Vingron. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet*. 29:365-71.
11. Nolan, T., R.E. Hands, and S.A. Bustin. 2006. Quantification of mRNA using real-time RT-PCR, *Nature Protocols* 1:1559-82.
12. Ward, T.H., and J. Lippincott-Schwartz. 2006. The uses of green fluorescent protein in mammalian cells. *Methods Biochem Anal*. 47:305-37.
13. Baron, U., and H. Bujard. 2000. Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol*. 327:4014-21.

## 9. Role of national and international funding organisations in Systems Biology

Alfred Game, *BBSRC, Swindon, UK*, Christina Kyriakopoulou, Fred Markus and Jacques Remacle, *European Commission, Directorate-General for Research, Directorate F (Health), Brussels, Belgium*

Dedicated funding streams for Systems Biology as such (apart from capacity building, e.g. the creation of centres) are unusual, even in the United States. This possibly reflects the latitude of definition and the cross-cutting nature of the field. There is, however a substantial and growing corpus of Systems Biology research supported by national and international funding agencies, albeit under a variety of badges and headings. In almost every country with active life sciences research some Systems Biology research is likely to be found, often falling under the themes of genomics, integrative biology or complexity. Unfortunately, the vagaries of definition and terminology are such that identifying or quantifying this investment is almost impossible.

Some projects in Systems Biology have been funded for many years (e.g. the electrophysiological systems modelling of heart function by Denis Noble in Oxford, which has been in progress for more than 20 years, funded by the UK Medical Research Council amongst others). However, the first significant funding initiative at the programme level in Europe was in 2001. As part of an initiative to stimulate medium-long term collaborative project in life sciences the German BMBF funded the Hepatocyte Alliance, a multi-sited coordinated €50 M programme involving 25 groups based around two technology platforms (cell biology and modelling) and two sub-projects (detoxification/de-differentiation and regeneration). Built heavily on previous strategic investments, e.g. in genomics and complexity studies, the Alliance began work in 2004. The Hepatocyte Alliance is of particular importance because the pioneering and very lengthy process of proposal development and evaluation set a benchmark for subsequent Systems Biology initiatives in Europe: the equal emphasis on modelling and experiment and on the integration of the two, the study of a common system, the use of common technology platforms etc. and, in particular, the long-term and ambitious vision of the whole.

The establishment of centres in Systems Biology – intended to integrate multi-disciplinary teams with sufficient stability and critical mass – has subsequently become a major theme in the last four years. Some of these have been established through the initiative of individual institutions in partnership with national funders. Others (e.g. in the UK, Germany and Ireland) have arisen from competitive calls from funders to establish one or more centres. Some have developed from existing genomics centres through the acquisition of additional skills. At least one – the Centre for Computational and Systems Biology at the University of Trento in Italy – is a private-public partnership initiative between the university, the regional administration and Microsoft Corp. A list of current and planned European Systems Biology research centres with national government funding – assembled by the ERASysBio initiative (see below) – is given in Table 9.1. However, a recent report from the UK Academy of Medical Sciences/Royal Academy of Engineering[1] listed 16 established and developing centres in the UK alone.

The majority of the established and developing centres have an emphasis on high-throughput genomic approaches to systems at the cellular level and below. The majority also have links with companies, often large multinationals in the pharmaceutical, agrochemical, food or personal care product sectors.

There have been relatively few research project funding initiatives at European national level aimed solely at Systems Biology. The Academy of Finland was the first, launching its SysBio initiative in 2004 with an initial fund of €9 M. The initiative has funded 17 projects through the Academy and an additional four through the technology transfer agency TEKES. The UK Biotechnology and Biological Sciences Research Council (BBSRC) currently has a call for project proposals under evaluation which is seeking to fund large Systems Biology projects in the €2-5 M range. This includes options for substantial industrial partnership. The BBSRC has also recently (2007) launched a €10 M bilateral call for Systems Biology research projects with the Agence National de Recherche (ANR), France.

Multinational European initiatives were effectively seeded by EUSysBio[2], a two-year FP6 'specific support action' with eight partners established in 2003. This focused mainly on surveying the state of Systems Biology in Europe and various networking activities, but from it sprang two major developments: SysMO and ERASysBio.

The Systems Biology of Micro-Organisms (SysMO) is a partnership of six countries (Austria, Germany, Netherlands, Norway, Spain, UK) who together launched and co-funded a call for collaborative research proposals focused on complex networks in microorganisms with relevance to biotechnological research and production, to start in 2007. Eleven outstanding transnational research projects have been selected for funding, with a total funding volume of approximately €28 M (plus contributions from universities etc). Groups from France, Switzerland and the Czech Republic have also participated with their own funds.

SysMO seems well placed to make pioneering developments in structures for data and model exchange and interoperability between its projects.

ERASysBio[3] is an association of European funding agencies in Systems Biology set up under the EU ERANet scheme to investigate how to work together to reduce barriers and enable joint research and training in Systems Biology. It includes all the SysMo partners, together with Begium, Finland, France, Israel and Slovenia and the Autonomous Province of Trento in Italy. The Russian Federation and Luxembourg are also affiliated. ERASysBio is currently exploring the scientific opportuntities for joint funding action, as well as considering topics such as transnational integrated projects (IP) issues and public communication about Systems Biology.

The European Commission itself has been a major actor in the development of Systems Biology in Europe, primarily through the support of research projects and through a number of key studies and workshop events which have helped to integrate this emerging community.

The EU supports research through its Framework programmes which are spread over several years and are essentially dedicated to transnational multidisciplinary research. Compared with FP5 (1998-2002), which mainly supported small and medium collaborative research projects, FP6 (2002-2006) has additionally offered more ambitious 'new funding instruments', namely integrated projects (IP) and networks of excellence (NoE).

The EU FP6 thematic activity on 'life sciences, genomics and biotechnology for health' has had a clear focus in the post-genomics era, strengthening the challenges following the sequencing of the human and other organisms' genomes. Fundamental genomics research received approximately €600 M support in FP6 for a large number of collaborative projects (small-medium (90) and large (40)). It is clear that the 'new instruments' of FP6 have enabled European scientists to achieve a real critical mass in very competitive areas of functional genomics and have really given European research a global profile.

The EU is emerging as a major world player in Systems Biology in FP6 and will increase its role in future years in FP7 (2007-2013). In FP6, a number of Systems Biology projects initiated in 2005 have already demonstrated that a systems approach can indeed work, both to provide a deeper understanding of biological processes and to provide predictive potential for applications. Funding of ongoing collaborative projects includes: (i) 14 STREP/CA/SSA* projects (approximately €25 M), covering basic signalling pathways (cell-cycle, modelling kinase signalling pathways, coordination of the yeast Systems Biology community, integrative genomics and chronic diseases phenotypes, coordination efforts of Systems Biology of cancer, modelling of post-translational modifications; (ii) several large funding schemes (IPs/NoEs) (approximately €60 M), covering areas such as bioinformatics tools for Systems Biology (ENFIN), Systems Biology for medical applications (BioSim) understanding the dynamic transcriptional regulation in bacteria and mammalian

Table 9.1: Leading Systems Biology centres in Europe

- The Oxford Centre for Integrative Systems Biology, UK
- Netherlands Institute for Systems Biology (NISB), Amsterdam
- German Cancer Research Centre, Heidelberg, Germany
- Freiburg Initiative for Systems Biology (FRISYS), Freiburg, Germany
- Multidisciplinary Centre for Integrative Biology (MyCIB), Nottingham, UK
- The Manchester Centre for Integrative Systems Biology (MCISB), UK
- Centre for Integrative Systems Biology of Ageing and Nutrition (CISBAN), Newcastle, UK
- Department of Plant Systems Biology, University of Gent, Belgium
- National Institute of Biology, Ljubljana, Slovenia
- European Bioinformatics Institute (EBI), Cambridge, UK
- Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany
- Centre for Systems Biology at Edinburgh (CSBE), UK
- Centre for Integrative Genetics (CIGENE), Ås, Norway
- Microsoft Research - University of Trento Centre for Computational and Systems Biology (CoSBi), Italy
- Warwick Systems Biology Centre, UK
- Centre Systems Biology (CSB), Stuttgart, Germany
- Potsdam-Golm BMBF-Forschungseinrichtung zur Systembiologie (GoFORSYS), Germany
- Centre for Integrative Systems Biology at Imperial College (CISBIC), London, UK
- Kluyver Centre, Netherlands
- Center for Medical Systems Biology (CMSB), Leiden/Amsterdam/Rotterdam, Netherlands
- Institute for Medical Genomics Research and Systems Biology (IMGuS), Austria
- Systems-X, ETH Zurich (Switzerland)
- Science Foundation Ireland Systems Biology Centre (proposal currently under evaluation)

models (BaSysBio, EUTRACC) and the developmental processes in plants (AGRONOMICS). These projects have major links to ongoing worldwide research programmes and also link and support European-wide (EMBL, EMBL-EBI), and national programmes. It is noteworthy to mention that the EU has also supported large-scale functional genomics initiatives in FP6 producing new knowledge on basic biological processes (mammalian cell cycle, tissue development and degeneration), human and animal stem cell differentiation processes, organelle function, endocytosis and post-translational modifications, that will pave the way for future Systems Biology approaches in Europe. To further our understanding of biological phenomena, there is a need for quantitative approaches and systematic modelling and analysis of the information gathered by high-throughput technologies. Among other things, Systems Biology requires modelling and simulation of the complex dynamic interactions between genes, transcripts, proteins, metabolites and cells using integrated systems-based approaches. Systems Biology is opening the way towards predictive and applied biology.

The EU FP7 programme will play a major role in this important and rapidly expanding research field by establishing the multidisciplinary networks in Europe that will catalyse the progress and the excellence in this field. This research will involve a wide variety of disciplines and critical mass that will require essentially collaborative efforts. The first FP7 call for proposals will support four large-scale integrating projects, in areas of Systems Biology approaches to unicellular organisms, T-cell activation, apoptosis and stem cell differentiation (approximately €48 M). In the second call, some tens of millions of Euros are available for several small and medium collaborative projects enabling Systems Biology approaches in basic biological processes relevant in health and disease.

By funding Systems Biology initiatives and by linking relevant programmes in Europe, the EU has provided a strong basis for a European Research Area (ERA) in Systems Biology in the coming years.

* STREP: specific targeted research projects;
CA: co-ordination actions;
SSA: specific support actions.

References


1. Academy of Medical Sciences/Royal Academy of Engineering (Feb. 2007) *Systems Biology: a vision for engineering and medicine* www.acmedsci.ac.uk
2. http://ec.europa.eu/research/health/genomics/newsletter/issue5/article05_en.htm
3. www.erasysbio.net


## 10. Opportunities and Challenges for Systems Biology in Drug Discovery & Development


Adriano Henney
*AstraZeneca, Global Discovery Enabling Capabilities & Sciences, Cheshire, UK*


**The opportunities**

The opportunities for Systems Biology in helping to bring innovative and effective new medicines to the market largely arise from the challenges that the pharmaceutical industry now faces. Over the last 10 years the cost of developing new drugs has escalated hugely, matched by a significant decline in the number of new medicines reaching the market. Further, compared with the 1960s, the time taken to develop a new drug has doubled to approximately 12 years. The available data from across the industry shows an increased productivity in the discovery phase, but this has not been matched by success in development, with a number of well-publicised failures recently of drugs in the later stages of the development pipeline. This suggests that whilst the *output* of projects from discovery into development has increased, the *quality* of the output has not. What is the reason for this disappointing return on investment?

There was optimism that the explosion of biomedical information following the mapping of the human genome would contribute not only to significant advances in understanding disease mechanisms at the molecular level, but more importantly perhaps, the expectation that this would translate rapidly into the delivery of novel therapies. However, the reality is that the gap between bench discovery and bedside application appears to have widened. As the 'low hanging fruit', i.e. adopting established approaches to treat complex disease, has been harvested, the need now is for increasingly innovative therapeutic solutions. 'Post-genome', the number of such novel targets to be considered in drug discovery has escalated hugely compared with what was available before. However, in many cases, novel targets present a major challenge. Current evidence shows that compounds directed at such novel targets have a much lower chance of success in development, with the major reasons for attrition being preclinical toxicity, followed by lack of clinical efficacy and inadequate clinical safety. In combination, these factors account for up to 60% of development project failures.

**How can Systems Biology help?**

Pharmaceutical R&D generally has been an empirically data-driven, qualitatively oriented, activity and modelling and simulation has played a relatively minor role. Whilst targets being studied in the discovery phase

may be placed within a network, it is not necessarily clear which of the many targets is the one worth pursuing therapeutically, because of the complexity of the biological system itself, compounded by variability between individuals. Typically, each drug and target combination tends to be considered in isolation, using target-driven, high-throughput, reductionist approaches that are removed from their physiological context. It has become clear that a reductionist approach focusing on individual entities in isolation can often be misleading, as the mechanisms contributing to the development of the diseases we need to treat are complex and not just the result of the contribution of a single gene or its protein product. The need, therefore, is to consider these 'novel' targets in the context of the functional networks that operate in the development of complex diseases. Specifically, we need to understand the pathophysiological context in which the selected targets are operating if new medicines are to be effective, and Systems Biological approaches are the tools that offer the hope that this can now be done. By combining mathematics, engineering, statistics and experimental science to create models that can be used to simulate complex physiological networks, the hope is that we will be able to predict how the system will translate a molecular stimulus into a physiological response, and thus avoid some of the unanticipated difficulties encountered in development. The potential applications of robust and successful Systems Biological strategies are mapped out in Figure 10.1.

**The challenges**

The relatively disappointing return on the investment made in genomics, in terms of delivering successful new medicines to the market has resulted in a pharmaceutical industry that is more cautious about the risks of embracing new technologies too readily. Paradoxically, the pressures the industry now faces logically demand the exploration of new ideas as alternatives to the status quo. All new scientific developments and technological innovations are met with a degree of scepticism and resistance to implementation, and Systems Biology is no different. However, before any vision for how Systems Biology might contribute to the development of novel therapies can become reality, there is a need to build confidence in its ability to deliver tangible value. In order to be recognised and accepted in this context, Systems Biology needs to show that it can deliver on its promise with strong examples of relevance and impact. This is its major challenge, which will need to be addressed through programmes of work designed to influence a change in attitudes and working practices, particularly in breaking down conventional disciplinary barriers. Immediate opportunities to demonstrate early success and impacts most likely lie in the application of Systems Biology approaches to existing projects, focusing on questions of toxicity and clinical efficacy. Even though this is a reactive approach, tackling the challenges faced now by projects, an ability to demonstrate success will help to build confidence in what Systems Biology can do, as well as identify what its

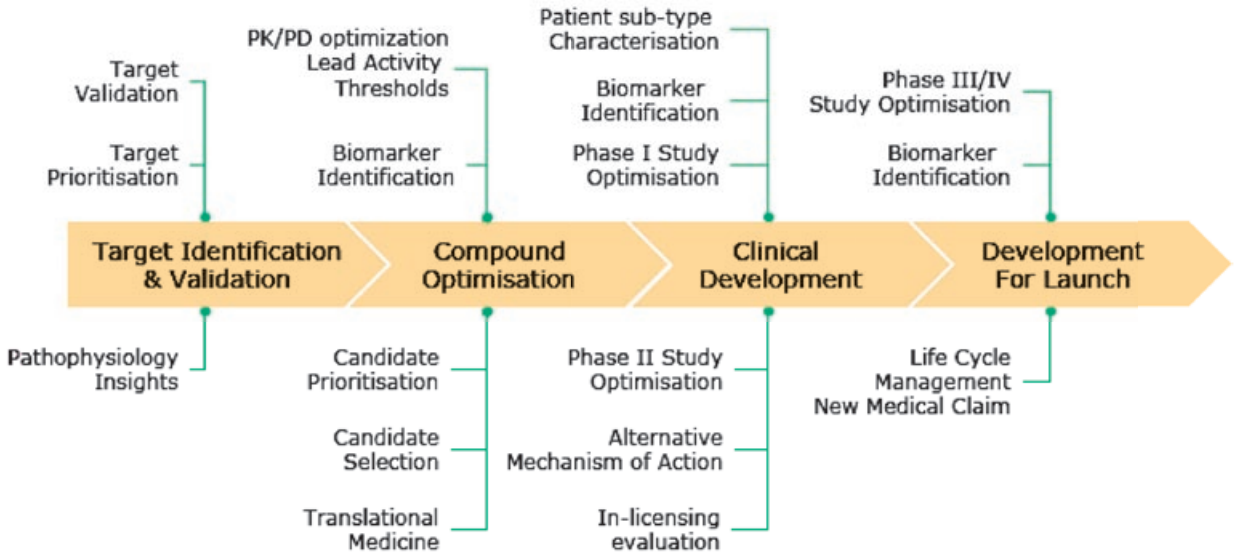



Figure 10.1: The Systems Biology: applications in drug discovery and development

limitations currently are. Strategic Programmes tackling broader complex problems in human biology and medicine, such as those suggested in this Forward Look will also help.

The key to this will be the delivery of the 'human capital' necessary to do the work; that is, the provision of scientists with the right skills. This represents a further challenge. Whereas many eminent scientists in the history of science were polymaths, modern scientific training and research is geared to the separation of the physical and biological sciences, with biologists tending to shy away from mathematics and vice versa. A successful implementation of systems approaches to understanding biology and medicine demands that researchers are trained to a level of numeracy and biological understanding that is currently hard to find. There is a need to build on the training programmes becoming established across Europe that will produce scientists that are accustomed to working routinely in an environment where mathematicians, engineers and biologists do not just work alongside each other on projects, but also to share the same language and understanding of the problems being addressed.

## 11. The importance of Systems Biology for the food industry – Systems Biology at Unilever

Janette Jones, *Port Sunlight Laboratory, UK*;
Guy Warner, *SEAC Colworth Park, UK*;
Bernd van Buuren, *Vlaardingen Laboratory, The Netherlands*; & Frans van der Ouderaa, *Corporate Research, Colworth Park, UK*

Unilever Europe is a leading fast-moving consumer goods company operating in over 150 countries world wide with approximately 200 000 employees and an annual turnover in Europe of €16 billion. Its field of operations is Foods and Home and Personal Care products. The Unilever R&D programme is carried out at many sites (development) in Europe and worldwide. The European sites for Research are at Vlaardingen (The Netherlands), and Port Sunlight and Colworth Park in the UK. The Safety and Environmental Assurance Centre (SEAC) also operates from Colworth Park. Unilever's mission is to deliver 'Vitality', i.e. functional products that make 'everyday people everywhere' look good and feel good and allow them to get more out of life. Examples of these products are skin products to prevent the signs of ageing, hygiene products for the home and for our bodies as well as more specifically, oral hygiene. Unilever's food products aim to provide energy, essential and functional nutrients, and, more specifically, products providing heart health. The branded products over the past 76 years have made significant contributions to human health, for instance by preventing malnutrition, improvements of personal hygiene, dental decay and cardiovascular disease both due to the products themselves as well as behavioural changes engendered by brand communication. Unilever's interest in Systems Biology is in the fields of:

- consumer safety,
- human biology applied to human nutrition and ageing, and
- microbiology.

### The role of Systems Biology in assuring consumer safety

It has been suggested that it is plausible for new risk-assessment paradigms to be developed to enable consumer safety decisions to be made by using data other than that generated in animal tests[1]. Consequently, SEAC is carrying out work to develop novel ways of assessing consumer safety. In particular, it is envisioned that Systems Biology approaches will contribute to the development of improved safety risk assessments through their application in:

- providing the biological context for integration and interpretation of new types of non-animal data (e.g. those generated by using 'omics' technologies),
- guiding the development of new biological in vitro models, and
- developing new computer-based (in silico) models of biological processes.

Currently, we are investigating the use of 'omics' approaches in making risk-based safety decisions. For example, protein microarrays are being used to attempt to gain an increased understanding of the role of intracellular signalling in the activation of human dendritic cells following treatment with skin-sensitising chemicals. To interpret the data and maximise the knowledge gained from these technologies, we are developing a systems-based analysis platform to integrate 'omics' data with biological network information. The objective is to identify the components that underlie biological mechanisms of relevance to consumer safety. We are also investigating the applicability of integrative modelling approaches. In a joint collaboration between Unilever and Entelos Inc., an in silico model of the induction of skin sensitisation has been developed. This approach combines data from different experiments and integrates them together mathematically in the context of a framework of biological understanding. The project has yielded new insights into the underlying immunology of skin sensitisation, and identified promising new avenues of exploration for developing in vitro assays more predictive of the in vivo biology.

**Human Systems Biology for nutrition and ageing**
Dietary compounds are being used in functional food products for health improvement. For instance, a weight of evidence supports the use of phytosterols for improved cardiovascular health. One of Unilever's key objectives is to understand and predict the health effects of dietary compounds. For example, the interactions between the gut microbes (microbiota) and the human host represent a significant source of, still untapped knowledge on the complexity of human health. Understanding this can come from a Systems Biology approach focused on the host-microbiota interactions. Three (sub)systems are involved in this, i.e. the dietary compounds (e.g., polyphenols), the microbiota that metabolise them prior to absorption, and the physiology of the human body. Knowledge is still lacking on how, for instance, polyphenols are metabolised by the microbiota and on their bioavailability. A mechanistic understanding of the interaction between these complex systems will give insights in potential nutritional intervention strategies to improve consumer health. This may be regarded as an initial milestone in demonstrating the opportunities of a Systems Biology approach. This and other routes will allow the food industry to generate insights to identify and develop novel functional foods with real and proven health benefits for the consumer.

Working together in a coherent relevant scheme is vital for achieving these scientific goals. A full Systems Biology approach requires the combined effort of several scientific disciplines which are currently spread across independent organisations over the globe. Open innovation between the foods industry, biotech companies and academia can lever Systems Biology from a promising scientific discipline to a fully functional and well-accepted research tool in both academia and industry. More specifically, Unilever's Healthy Ageing Corporate programme aims to pursue health prevention consumer options in the area of Human Ageing and Metabolism.

To that effect Unilever has a strongly internationally networked programme investigating aetiologies, potential consumer prevention routes as well as self-assessment tools, within the framework of improving people's ageing trajectories. Unilever strongly believes that in such a complex biological process as human ageing a Systems Biology approach is the only way to develop in-depth understanding comprehensively. We are working closely with CISBAN (Centre for Integrated Systems Biology of Ageing and Nutrition) at Newcastle, one of the recently formed BBSRC centres of excellence in Systems Biology. In addition we have, with our partners, incorporated 'systems' approaches in past and current EU Framework projects, such as ZincAge.

**Microbiology**
Microbial systems are complex and a Systems Biology approach may yield important insights in this field. To this end Unilever also has a collaboration with the Systems Biology modelling company Genomatica Inc. using the approach of constraint-based modelling to understand biochemical pathways within micro-organisms.

Reference

1. Fentem, J., Chamberlain M., and Sangster B. 2004 The feasibility of replacing animal testing for assessing consumer safety: a suggested future direction. *Altern Lab Anim*. 32:617-23.

## 12. Systems Biology for profit in SMEs

John Savin
*Wendover Technology, Wendover, UK*

Systems Biology is now established as a discipline, but is not yet mainstream in its major market of pharmaceutical development. This is partly because of organisational reasons in major companies and partly due to immaturity of the science and lack of methods to evaluate economic returns. SMEs are well placed to adopt systems approaches and a new development paradigm as they are unencumbered by historic organisational constraints. They have three potential strategies:

- *Service:* build and run simulations for traditional development companies (based on tacit, in-depth knowledge and proprietary and freeware models);,
- *Expedite development:* use to reduce the development risk of novel therapeutics;
- *Product revival and combination:* many projects fail in phase II clinical development with no recovery strategies available. By aiding understanding of the development issues, systems approaches can aid revival and lead to combination therapy development thus extending patent life.

**Introduction**
Despite its relative scientific maturity, in business terms Systems Biology is not widespread in major pharmaceutical companies. The level of small company application is still rare. Consequently, there is no defined business route for using Systems Biology for profit in small companies. In a study of the US pharmaceutical and biotech industries published in late 2006[1], Professor Pisano showed that biotech has never collectively made money until recently and then only because of Amgen profits. Its crude productivity is similar to that of major pharmaceutical companies, but with lower sales. Investors over the long run made

16% per year on capital, a long way below the 40% targeted. The remedy, according to Professor Pisano, is better organisational learning, as the science needs a long time to be learnt. Furthermore, as stock market investors and venture capital providers know, individual projects are very risky with only 10-15% of clinical-stage projects gaining approval. The problem is that new therapeutics (or diagnostics) are complex to design and then have to work in a system (the human body) that is still poorly understood as a molecular system. Hence, the majority of development time and cost is spent not on product design (discovery) but on exhaustive testing where any failure destroys a project. Improving clinical success rates by 10% would cut the cost of development by US$242 M (2004 prices) and, more importantly, would give higher productivity.

This chapter advocates Systems Biology as a fast way to systematic learning to reduce development risk and avoid inflexible formulaic development routes. Systems Biology approaches cannot substitute for well-designed and rigorous laboratory science and clinical development. Rather, systems provide a way to make these into efficient proof statements and not raw unpredictable experiments in their own right. This chapter will therefore look at the 2006 real market for services and Systems Biology products, the science behind the business and then draw some conclusions about how business can deploy this science for profit over the next five years.

**Technology applications**

Most life science technologies take 10-20 years to mature to the point where they become mainstream. The initial paradigm needs to be modified by experience. Systems Biology has been around for long enough for serious applications to start to emerge but not long enough for its use and economic benefits to be more than theoretical. In brief, the areas can be divided into the following domains.

***1. Bioinformatics***

Systems Biology is often classed as a subset of bioinformatics, but this is inaccurate. In bioinformatics, one searches the known databases. In systems, one replicates living systems. In terms of profits: as the databases are all free, and so is much software, there are few profit opportunities and these are basically operating outsourced services plus some specialist software. Genomics players, such as InCyte, failed to sustain their US$200 M on bioinformatics business as data became public; the same will happen to Systems Biology with data sets and models.

***2. High-throughput approaches***

Perhaps the best examples are Beyond Genomics (USA) or the Institute for Systems Biology. Here, gene chips, proteomics and metabolomics give massive data sets for data mining. The risk is that given enough noisy data, data mining often finds patterns that are not reproducible. But this approach fits easily with the industrialised R&D approach and has a lot of utility in toxicology. To do this, one needs lots of hardware (so heavy capital expenditure, not ideal for European SMEs) and pathway software.

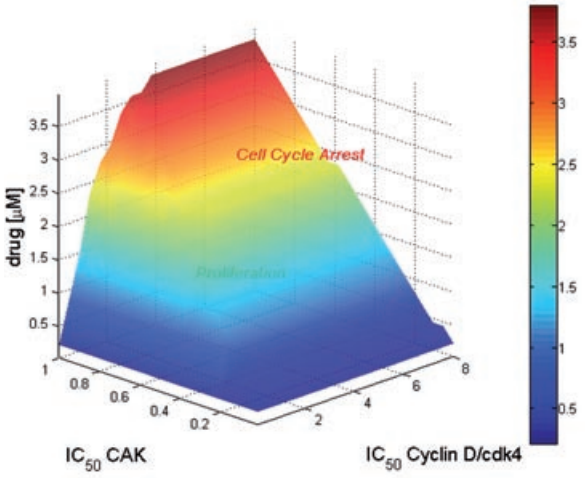



Figure 12.1: Dual kinase inhibition of mammalian cell cycle: $IC_{50}$s against drug concentration (Courtesy of Physiomics plc 2007).

***3. Pathway software***

Pathway software for mining and modelling is essential for high-throughput approaches, for example GeneGo. Genomatix, Ingenuity, Gene Network Sciences, These companies take large disparate data sets and assemble the most likely sets of interrelationships into pathways. This lets one find critical network points. These may be good drug targets or markers, or may not. The market is now crowded with many suppliers as barriers to entry are low and academic software and data is often free. Pathways are becoming integrated with replication approaches via software such as Cell Designer.

***4. Sophisticated process optimisation***

Process optimisation can have a fast payback by optimising expensive biological production systems; for example, for antibiotics and protein therapeutics. Examples of services suppliers are Genomatica (USA), Insilico Biotechnology (Germany) and Bayer Technology Services GmbH (Germany). This appears to have turned into a good market, but with probably room for only a few external specialist suppliers. Some use mathematical equations, others are based on stoichiometric metabolic networks.

***5. Replication***

The creation of mathematical models (two- or three-dimensional concepts assembled from the liter-

ature or experiment) to understand how a therapeutic works or a specific toxicity effect occurs. A simulation is when one adds the fourth dimension of time to a model allowing adaptation of the systems to change input variables; that needs a computer. This dynamic variation is a key differential from static bioinformatics and restricted time point high-throughput snapshots.

Replication can be developed as 'top-down', 'bottom-up' or 'middle-out', all being relative perspectives. Top-down is illustrated by Entelos' (USA) virtual disease-specific patients (PhysioLabs®). These simulations use decision theory and mathematics. They can encompass a wide range of clinical empirical observation and also biochemical detail. The investment in any such model is high; the technology needs to be built over many years and the best ones so far appear to be for diabetes and inflammatory disease. Entelos sells over US$30 M per year of services and could achieve US$5 M of profit in 2007. However, the capital investment (excluding sales) in Entelos is over US$70 M and such capital is difficult to access for service businesses, particularly in Europe. Bottom-up or middle-out is probably best illustrated by Physiomics plc (UK). This company develops world-leading models of cancer cells from a detailed understanding of the biochemical machinery of the cell cycle and signal transduction. The advantage of this approach is that pharmaceutical parameters, target identity, inhibitor $IC_{50}$ and administered dose can be input giving a predicted and verifiable drug dose-response profile. Figure 12.1, courtesy of Physiomics plc, shows how a dual kinase inhibitor efficacy varies depending on the $IC_{50}$ against different targets and on the intracellular drug concentration.

To get this data in an experiment would be tedious, to get it in a clinical trial impossible. One can see that the response profile of the CAK (cyclin activating kinase) is different and more sensitive to $IC_{50}$ than the cyclin D – Cdk4 complex. One can also see in Figure 12.1 that there is a cut-off below which cell cycle progression will happen, although it might be delayed. Bottom-up approaches are typically 'one cell at a time'-based but need to move to cell populations (such as virtual tumours) and organ level to make a major impact. SystemCell® from Physiomics plc may be one solution.

An essential replication technology is whole body pharmacokinetic simulation. Bayer Technology Services GmbH has developed sophisticated and powerful PK-Sim® software, Figure 12.2, to find the organ-specific dose of a drug over time. By linking drug effects on cellular systems with the pharmacokinetics, complex development situations can be explored, for example, how to optimise a combination of therapeutics. Alternative doses and scheduling can be explored when PK is linked to the biochemistry: most clinical developments test a pattern of only a few round numbers: doses of 25mg, 50mg, 100mg which is a human cultural pattern with no relation to the underlying compound and disease.

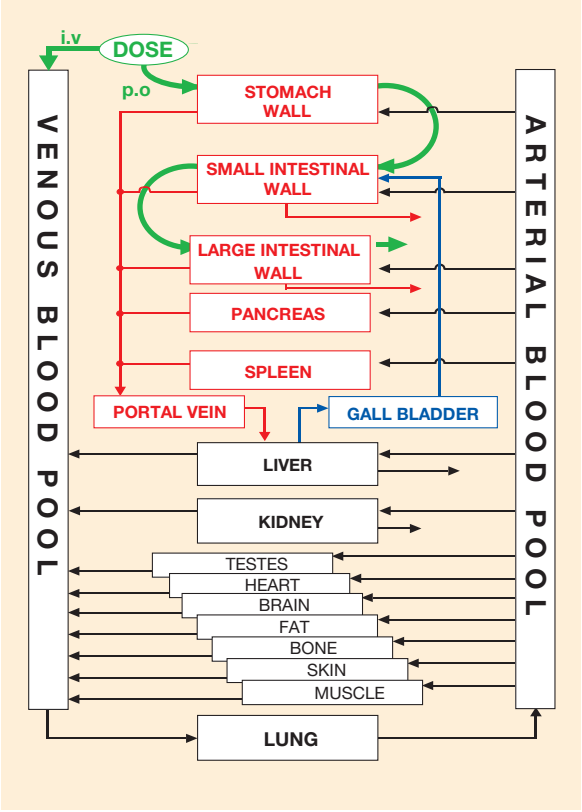



Figure 12.2: Structure of PK-Sim whole body pharmacokinetic simulation. (Courtesy of Bayer Technology Services GmbH).

**Market barriers to profit**

Most systems products suffer from limited markets and investment. One estimate from 2004 put the 2008 market at US$785 M. Systems Biology replication technology, which could have the biggest long-term impact, is not yet readily integrated into pharmaceutical development programmes as it works best on a multifunctional, integrated approach. Generally, large buyers of technology and services operate along narrow functional lines with huge cultural, communication and budgetary chasms between discovery, preclinical and clinical phases. Some companies (Astra Zeneca (UK) and GlaxoSmithKline (UK)) have very sophisticated and skilled in-house groups, but as yet it is not clear that they have significantly impacted on drug development. Roche (Switzerland) is stated to use systems

approaches, probably mostly PK simulation, in 50% of phase II and III trials. Lilly (USA) has invested in a high-throughput facility in Singapore. The concept of translational research is to bridge the preclinical to clinical gap but this has had limited impact in replication systems marketing to date. In addition, the commercial market for bottom-up approaches is limited by the availability of free academic models and various sites (www.siliconcell.net/sica/) and organisations such as the European Bioinformatics Institute (EBI) curate these and make them available (www.biomodels.net). Academic modellers have tended to cluster around well-understood pathways, for example, the EGF pathway has been reworked many times.

The top-down approach, because of its clinical complexity, has a higher barrier to entry. However, it suffers from rigidity whilst the bottom-up approach can offer custom-built models assembled from off the shelf modules. Modularity in bottom-up is in theory excellent, but there are no standards for coupling modules and running the enlarged simulation. To couple models at present, one needs to understand both and establish that parameters and assumptions are compatible. Even then, a previously stable simulation with a defined parameter set can become unstable when enlarged and need to be re-parameterised. Parameters for metabolic and replicate models come from the literature, whether top-down or bottom-up. As replication models aim to give quantitative results, accuracy is important. But literature values, if available, are highly variable and often found in pre-electronic literature. Government funding to provide standardised parameter sets would facilitate replicate adoption.

New targets often have little or no quantitative biochemistry or enzymology and some parameters are in any case experimentally intractable. Hence, all replication systems need to use mathematical fitting to find stable data sets. Biochemically, control analysis shows that cells tolerate a wide range of possible values before entering narrow unstable zones. To be effective, a pharmaceutical has to drive a cellular system into that critical zone and this frequently means cutting an enzyme activity to 5% or less of it normal value. Parameter fitting is best done by computationally intensive processes. Physiomics, for example, can link qualitative formal descriptions of cellular behaviour with quantitative differential equation models. Finally, there are 'academic' companies. Senior academics with good ideas can often do significant consulting work. This opens the way to maintain an SME's 'staff' on academic grants, going private only when capital or income allows. The problem is that while this can keep an enterprise going for a long time, it is difficult to combine the academic and commercial perspectives.

**Evolving business models**

There are three workable long-term strategies for independent companies.

- service,
- use in product development, and
- product revival and combinations.

**Service**

Service is a good strategy but will take time to implement. Entelos has just become profitable after nearly 10 years and is stock market listed (2006); Physiomics plc listed in 2004. The key to profit is a high level of tacit and proprietary knowledge in the company that is perceived as valuable by customers and is difficult to replicate by competitors (industrial and academic). Even so, service is generally a lower return business model and venture investors are wary. However, the current challenge for all systems service businesses are that potential customers do not see Systems Biology as necessary, do not know how to value it or what economic returns to expect and do not know how to extract value for their own organisation. The immediate application of process optimisation has a relatively fast payoff whereas it will take many years to have a full success story in therapeutic development.

**Product**

Product development is a good route for companies with novel therapeutic or diagnostic technologies and systems, which in theory, could help to reduce development risk. As an example, Cronos Therapeutics, part of the ValiRx plc group (UK) is planning to deploy replicate systems technology to understand the action of their novel gene silencing technology (GeneICE®). By doing so, Cronos will capture a huge amount of tacit knowledge about how its anti-cancer therapeutics work when targeted at various cell cycle targets. Issues such as dose, timing of therapy and the development of combination therapies can all be done by simulation and tested in focused experiments. Cronos, being a relatively new company with a virtual business model can adopt such strategies without organisational conflicts. A further example is Cyclacel Inc, (USA and UK), a more traditional small molecule cancer therapeutics company that has used Physiomics as a supplier and partner since 2002.

**Revive and combine**

Product revival and rescue is an application not previously discussed and is really only possible with systems approaches. It derives from the concept of re-profiling where molecules that failed in one therapy indication are found to be useful in an alternative indication. It is particularly popular with generics, if they can be repackaged. Revival involves taking a recent problem

project and using systems to overcome the problem. In a standard pharmaceutical development project, if an unexpected dose-response or toxicity issue arises, there is no way to fully understand what happened. Yet, particularly for phase II compounds (which typically suffer 60-70% attrition), there is a huge amount of data available with probably US$10-100 M sunk in investment. A systems approach using replicate models and simulating the pharmacokinetics could allow the available data and tacit knowledge of the project team to be integrated and alternative development strategies deployed. This depends, or course, on a good quality replicate model being available. However, such models can be constructed or adapted by the right team within 3-6 months. An example of this is Renaissance Pharmaceuticals Ltd. (UK), a new company in the area of coxibs and NSAID therapy. This market suffered a US$4.2 billion drop in sales with the withdrawal of Merck's Vioxx® and the uncertainties about cardiovascular risk attached to Pfizer's Celebrex® and the generic non-steroid anti-inflammatory drugs.

By using very powerful and intricate replicate technology, Renaissance has an excellent quantitative understanding of the underlying cardiovascular and gastrointestinal toxicity. By developing combination therapies, difficult and expensive to do by standard project protocols, new drug classes with extended patent life can potentially be created. This strategy could potentially be applied to many types of therapeutic and biomarker development and would be attractive to customers and investors.

**Don't try this**

Strategies unlikely to work are also crucial. Using systems approaches purely to find and validate targets is sensible only if off the shelf models are available but the approach offers no clues for the actual molecular shape and design of the therapeutics. Pure software approaches do not look to be sensible. At present, commercial software needs to be very powerful to compete against academic freeware. In addition, manufacturers of mathematics programs can develop simulation plug-in modules to add specific functions to their standard packages and thereby control the market at lower cost than any stand-alone systems operation. Simply selling replication model packages will also not work without a strong service element. Demand and pricing are too weak, development costs too high and the integrated project cannot be protected.

**Conclusions**

The terms 'profit' and 'Systems Biology' are not incompatible, as Entelos has now shown. But the use of Systems Biology by SMEs needs to be carefully applied to get the investment returns that the capital markets demand. Systems Biology will eventually become a mainstream tool but it is different in some regards from other biotech platforms (genomics, bioinformatics) in that it cannot just be plugged in as another R&D department or machine. Efficient use of systems demands organisational change. That might be the real impact of systems: a new type of flexible SME such as Cyclacel and Cronos deploying systems in a new way to manage development risk effectively or, like Renaissance, pioneering an emerging 'revival and combine' strategy. In that event, service providers such as Physiomics, Bayer Technology Services and Entelos will have strong profitable futures as their tacit knowledge and expertise is sought. But, as with all new technology, Systems Biology needs to prove that the economic benefits are real and the profits not a wishful simulation.

Reference

1. Pisano, G.P. 2006. *Science Business*. Harvard Business School Press, Boston, USA.

Further reading

- Karl Thiel. 2006. *Nature Biotechnology* 24(9): 1055-7.
- Tufts CSDD Impact Report (2002) Volume 4.
- Chassagnole, C., Jackson, R.C., Hussain, N., Bashir, L., Derow, C., Savin. J, and Fell, D.A. 2006. Using a mammalian cell cycle simulation to interpret differential kinase inhibition in anti-tumour pharmaceutical development. *Biosystems* 83: 91-7
- The Future of Systems Biology: Emerging technologies and their impact on drug discovery, development and diagnostics. *Business Insights*. 2005.

# D. Recommendations and Priority Areas

## 1. Recommendations to meet the Grand Challenge

Systems Biology offers exciting opportunities in the biomedical and biotechnological domains. If issues are tackled in a goal-oriented manner at the European level, results can have a major impact on society and the economy. The recommendations below may act as a guide to meet this Grand Challenge for Europe.

**1. A European road map**
A task force should be created in which the major stakeholders are represented, including top scientists, industry, European science organisations and funding agencies, and representatives of the European Commission. This group should undertake the following:

- define road maps for the Grand Challenge for the next 10 and 20 years;
- establish a financial plan;
- propose an adequate research management structure that allows cost-effective, goal-oriented, large-scale research efforts;
- define new concepts for effective technology transfer and commercialisation of results; and
- make a cost-benefit analysis.

**2. European Reference Laboratories (ERLs)**
A cost-effective coordination of a European Systems Biology programme, as proposed here, requires a consortium of European Reference Laboratories (ERLs). ERLs are research institutes that combine all relevant scientific disciplines and the know-how to provide outstanding expertise for core aspects of Systems Biology. They should be accessible to investigators from all over Europe so they can obtain the necessary introduction and training in this field. ERLs should develop the scientific basis for standardisation and quality control of experimentation and modelling. In current biological and biomedical research, standardisation and quality control systems play only a limited role. As a result, data obtained by different research groups cannot easily be combined; this is particularly true for quantitative studies. Since progress towards a true understanding of complex biological systems crucially depends on properly handling such data, it is imperative that academia and industry define adequate standards and that these standards are implemented into databases. At the same time such standards should remain adaptable to changes as knowledge increases.

**3. Cooperation between industry, academia and funding agencies**
The Europe-wide approach proposed here will require an integrative and larger-scale level of funding than provided by grant systems now. We propose that a new financial model is developed based on cooperation between academies, industry or EC-related organisations. All parties should contribute to and benefit from the programme, but with a clear understanding of what belongs to the public domain and what is rightfully the proprietary information of private companies. It will be a major challenge to align this heterogeneous set of parties so as to produce a synergistic cooperative programme.

**4. Public acceptance**
Large-scale European efforts will be viable only if the general public accepts and endorses the underlying ideas and goals of the Grand Challenge for European Systems Biology programmes. To achieve this, careful communication and explanation will be necessary from the outset, with initiatives in social sciences and ethics accompanying the Systems Biology research. A broad and open debate will enable the public to be aware of the socio-economic and health benefits, but also of any potential risks of the Grand Challenge, allowing a balanced and objective monitoring of the further development of this field from the very beginning.

**5. Training and education**
In contrast to the present practice of educating scientists in the classical disciplines, the Systems Biology approach requires new thinking across classic scientific borders. Currently, progress in biology is hampered by the largely monodisciplinary teaching systems in Europe. New interdisciplinary BSc, MSc and PhD teaching and training programmes should be implemented with high priority given to addressing this issue. In this respect, an open debate with responsible European stakeholders and higher education institutions (e.g. universities) should be initiated.

**6. Cooperation between different European Systems Biology initiatives**
As this Forward Look was initiated in 2004, Systems Biology was still in its infancy in Europe. In the mean time, however, a wide range of initiatives and funding programmes has started, both at the national and the European level. Examples are the ERANet programme ERASysBio, the SysMO programme and a variety of large-scale investments in new Systems Biology research institutes and research programmes in Germany, UK and Ireland and recently, with ANR in France. Similarly, Systems Biology-oriented programmes are developed in the context of the EU FP7 and ESFRI. Any European-level initiative in the Systems Biology field should align with and learn from these ongoing activities.

**Specific actions to implement recommendations:**

1. *Establish* a task force to define a European road map. It will be necessary to concentrate on a limited number of specific targets that are of major importance for human health and have a substantial impact on economic growth. These targets should create proof-of-principle and demonstration projects and fully exploit European strengths and expertise in this field. Defined milestones should be achievable within five and 10 years
2. *Organise* a high-level strategic workshop for national and European funding agencies, industry and charities to develop an appropriate financial plan for the Grand Challenge. In addition, this high-level group should work out a strategy plan aimed at the integration of ongoing activities in the field.
3. *Set-up* a network of European Reference Laboratories (ERLs).
4. *Involve* specific target groups, such as patient organisations, health insurances, regulatory bodies and publishing groups. It is important that they contribute to the development of the Grand Challenge. *Perform* business case studies to explore the costs and benefits of Systems Biology in general and of the large-scale pan-European Grand Challenge approach in particular. A review of the socio-economic value of such a large-scale approach will provide reliable data to funding institutions and policy makers on the expected outcome of the programme.
5. *Establish* a European training programme in Systems Biology to provide young investigators with the appropriate background. The programme could include regular summer schools, courses, graduate schools, BSc and MSc programmes.
6. *Stimulate* transnational collaboration and networking between ongoing Systems Biology programmes throughout the world to act in a synergistic manner and profit from their experiences.

## 2. Priority areas in Systems Biology for the next five and 10 years

It has been argued in this report that enforcing breakthroughs in biomedical, pharmaceutical and biotechnological fields requires the concerted, well-tuned and coordinated effort of large numbers of investigators from different disciplines, including biology, medicine, physics, chemistry, mathematics and engineering. This calls for a pan-European endeavour that will benefit society by boosting economy and improving health. Given that large-scale programmes will be necessary, we have to make choices. How should this be done?

**Priority areas should be scientifically challenging**

There are different ways of approaching this question. Since this report is written primarily by scientists, there is first of all the question 'what are scientifically challenging and at the same time feasible topics?' Several choices have already been made, for instance 'understanding the human hepatocyte' in the German HepatoSys programme (www.systembiologie.de/de/index.html) that started in 2004. This is an important and pioneering Systems Biology programme from which we can learn. Recently the transnational SysMO programme (www.sysmo.net) has been initiated, concentrating on the biomolecular networks that define the behaviour of microorganisms. A third example is the Yeast Systems Biology Network (YSBN) (http://www.ysbn.org) that exploits the yeast *Saccharomyces cerevisiae* as a model system to develop the field and to advance our understanding of the rules and principles of the dynamic operation of cellular systems. Finally, there is the Systems Biology programme *avant la lettre* guided by Dr Denis Noble (Oxford, UK) aiming at understanding the human heart by integrating molecular, cellular and tissue levels (http://noble.physiol.ox.ac.uk/People/DNoble/). Each of these large-scale programmes has its own scientific challenges. At the same time many scientific and managerial hurdles are remarkably similar, including standardisation, data integration, keeping large groups of scientists on track, etc. Importantly, in all cases economic issues played an important role in making choices as to what cell type and organism should be studied and what priorities should be set. The hepatocyte has a central role in drug metabolism, microorganisms studied in the SysMO programme are highly relevant for biotechnology and the same holds for yeast. This illustrates that from the point of view of scientific challenges many topics may be highly interesting to work on.

**Priority areas should boost European industry**

Large-scale European projects exploiting Systems Biology approaches are very likely to require teaming up of academia and industry. This will, as illustrated above, lead to choices that are important for biomedical, pharmaceutical, biotechnological and food industry. Chapters in this report underscore that industry is discovering Systems Biology as a key development in biological research at all levels of biological complexity. The hurdle in implementing Systems Biology in research carried out by industry is that to be successful it requires long-tem vision and large investments, which for most companies are difficult. This report strongly pleads for large-scale and goal-oriented investments by governments and the European Commission to overcome this hurdle by teaming up academia and industry in large-scale pro-

grammes. If successful, this investment will create a wide range of opportunities for application in the commercial domain.

**Priority areas should improve human health**

Progress in our understanding and rational tackling of major Western disorders, including diabetes, metabolic syndrome, Alzheimers, cancer and in several ways ageing, is disturbingly slow. As made clear in this report, this paucity is largely because of the extreme complexity of biological systems. Systems Biology is a systematic tool to come to grips with these problems. Considering the large number of people threatened by these disorders, it is likely that large-scale European projects will concentrate on some of these issues. Such focus no doubt also satisfies the interest of industry and academia. How to make choices in the health domain?

Here we must accept that Systems Biology is still in its infancy. Therefore, a thorough analysis should be made of what can be achieved at what cost and on what time scale. This report therefore recommends carrying out medium-term (five years) and long-term (10 years) cost-benefit analyses by a team of scientists that are proficient in large-scale Systems Biology, experts in the relevant biomedical or biotechnological domain and specialists that can look at the issue from the economic perspective. A rational approach, however, is needed as politics will play an important role and political support is needed.

**Setting the right priorities**

Summarising, the large-scale European programmes that Systems Biology calls for are likely to focus on health and biotechnology. The following steps should be taken in order to make the right choices.

- Careful cost-benefit analyses to decide what is feasible, most beneficial for society at what cost in the next five and 10 years.
- Team up all key stakeholders in Europe.
- Draw a road map for each potential focus, incorporating critical milestones and the budgetary consequences.
- Seek funding that will require novel ways of combining efforts of diverse European funding agencies[1].

1. The EuroBioFund initiative of the EC and the ESF is an interesting step in this direction.

# E. Appendix I

## 1. Activities

1. **First Grand Challenge meeting**
   Heidelberg, EMBO office, 2-10-2003
2. **Second Grand Challenge meeting**
   Brussels Airport, 9-1-2004
3. **Systems Biology Preparatory meeting**
   The Hague, 29-1-2004
4. **First Steering Committee meeting**
   Basle Airport, 28-6-2004
5. **Second Steering Committee meeting**
   Heidelberg, DKFZ, 15-10-2004
6. **Third Grand Challenge meeting**
   Amsterdam Airport, 24-11-2004
7. **Interim report FL-SB**
   January 2005
8. **Meeting SB workshop**
   Amsterdam Airport, 21/22-2-2005
9. **Final meeting**
   Gosau, 17/18-3-2005

## 2. Committees and Scientific Advisory Board

### Members of the Steering Group

**Dr Roland Eils**
German Cancer Research Centre (DKFZ)
Im Neuenheimer Feld 280
69120 Heidelberg • Germany
r.eils@dkfz.de

**Prof. Frank Gannon**
European Molecular Biology Organisation
Meyerhofstrasse 1
Postfach 1022-40
69012 Heidelberg • Germany
frank.gannon@embo.org

**Dr Igor Goryanin**
Glaxo-Smithkline
Gunnels Wood Road
Stevenage SG1 2NY • United Kingdom
goryanin@hotmail.com

**Dr Stefan Hohmann**
Göteborg University
Dept of Cell and Molecular Biology/ Microbiology
Box 462
40530 Göteborg • Sweden
hohmann@gmm.gu.se

**Prof. Karl Kuchler**
University and BioCenter of Vienna
Institute of Medical Biochemistry
Campus Vienna Biocenter
Dr Bohr-Gasse 9/2
1030 Vienna • Austria
karl.kuchler@meduniwien.ac.at

**Dr Bela Novak**
Budapest University of Technology and Economics
Faculty of Chemical Engineering
Dept of Agricultural Chemical Technology
1521 Budapest • Hungary
bnovak@mail.bme.hu

**Prof. Dieter Oesterhelt**
Max Planck Institute of Biochemistry
Abteilung Membranbiochemie
Am Klopferspitz 18a
82152 Martinsried • Germany
oesterhe@biochem.mpg.de

**Dr Uwe Sauer**
Institute for Molecular Systems Biology
ETH Hönggerberg
HPT C72
8093 Zürich • Switzerland
sauer@imsb.biol.ethz.ch

**Dr Barry Utting**
Lightwater
107, Ivy Drive
GU185YT Surrey • United Kingdom
butting@netcomuk.co.uk

**Prof. Roel van Driel *(Co-Chair)***
Netherlands Institute for Systems Biology and University of Amsterdam
Kruislaan 318
1098 SM Amsterdam
Netherlands
van.driel@science.uva.nl

**Prof. Hans V. Westerhoff *(Co-Chair)***
School of Chemical Engineering and Analytical Science
The University of Manchester
PO Box 88, Sackville St
Manchester, M60 1QD • United Kingdom
hans.westerhoff@manchester.ac.uk

### Members of the Preparatory Group

**Prof. Lilia Alberghina**
Deptartment of Biotechnology & Biosciences
University of Milano-Bicocca
Piazza della Scienza 2
20126 Milan • Italy
lilia.alberghina@unimib.it

**Prof. Gianni Cesareni**
Dipartimento di Biologia
Università di Roma Tor Vergata
Via della Ricerca Scientifica
00133 Rome • Italy
giovanni.cesareni@uniroma2.it

**Prof. Mats Gyllenberg**
University of Helsinki
Department of Mathematics and Statistics
Rolf Nevanlinna Institute
PO Box 68
00014 Helsinki • Finland
mats.gyllenberg@helsinki.fi

**Dr Gerhard Mulder**
Biolateral- Kennistking
Utrechtseweg 38
1381 GP WEESP • Netherlands
mulder285@planet.nl

**Dr Sergio Nasi**
Laboratorio Acidi Nucleici
Dipartimento Genetica e Biologia Molecolare
Instituto di Biologia e Patologia Molecolari
Università la Sapienza
CNR
P. le A. Moro 5
00185 Rome • Italy
sergio.nasi@uniroma1.it

**Dr Magali Roux-Rouquié**
Unité de Biosystémique, Modélisation, Ingénierie
Institut Pasteur
25-28, rue du Dr Roux
75015 Paris • France
Magali.Roux@lip6.fr

**Prof. Jaroslav Stark**
Imperial College London
Centre for Nonlinear Dynamics and its Application
Department of Mathematics
Exhibition Road
London SW7 2AZ • United Kingdom
jaroslav.stark@ic.ac.uk

# E. Appendix I

## 3. Workshop Participants

### Systems Biology Workshop, Amsterdam Schiphol Airport 21-22 February 2005

**Dr Mark Chamberlain**
Unilever Research Colworth
Safety and Environmental Assurance Centre
Colworth House
Sharnbrook
Bedford
Bedfordshire MK44 1LQ
United Kingdom
mark.chamberlain@unilever.com

**Dr Roland Eils**
German Cancer Research Centre (DKFZ)
Im Neuenheimer Feld 280
69120 Heidelberg • Germany
r.eils@dkfz.de

**Dr Alf Game**
BBSRC
Polaris House
North Star Avenue
Swindon SN2 1UH • United Kingdom
alf.game@bbsrc.ac.uk

**Prof. Willi Jaeger**
University of Heidelberg
Im Neuenheimer Feld 294
D-69120 Heidelberg • Germany
jaeger@iwr.uni-heidelberg.de

**Prof. Douglas Kell**
Department of Chemistry
PO Box 88 •
University of Manchester
Faraday Building
Sackville Street
Manchester M60 1QD
United Kingdom

**Prof. Karl Kuchler**
Medical University of Vienna
Campus Vienna Biocenter
Dr Bohr-Gasse 9/2
1030 Vienna • Austria
karl.kuchler@meduniwien.ac.at

**Dr Erik Mosekilde**
Department of Physics
Danmarks Tekniske Universitet
Bygning 307-309
2800 Kgs. Lyngby • Denmark
erik.mosekilde@fysik.dtu.dk

**Dr Clemens Sorg**
Faculty of Medicine
Medical University of Innsbruck
Innrain 52
6020 Innsbruck • Germany
clemens.sorg@i-med.ac.at

**Dr Frank Tobin**
Dir. Scientific Computational Modelling
Glaxo Smith Kline
980 Great West Road
Brentford, TW8 9GS • United Kingdom
Frank.Tobin@gsk.com

**Dr Barry Utting**
107, Ivy Drive
Lightwater GU185YT
United Kingdom
butting@netcomuk.co.uk

**Prof. Roel van Driel**
Netherlands Institute for Systems Biology and University of Amsterdam
Kruislaan 318
1098 SM Amsterdam
Netherlands
van.driel@science.uva.nl

**Prof. Hans V. Westerhoff**
School of Chemical Engineering and Analytical Science
The University of Manchester
PO Box 88, Sackville St
Manchester, M60 1QD • United Kingdom
hans.westerhoff@manchester.ac.uk

**Dr Gerhard Mulder**
Biolateral- Kennistking
Utrechtseweg 38
1381 GP WEESP • Netherlands
mulder285@planet.nl

**Dr Carole Moquin-Pattey**
Head of Unit Medical Sciences (EMRC)
cmoquin-pattey@esf.org

## 4. Participants of Gosau Conference
### 17-18 March 2005, Gosau, Austria

### Academic Experts

**Prof. Lilia Alberghina**
Department of Biotechnology & Biosciences
University of Milano-Bicocca
Piazza della Scienza 2
20126 Milan • Italy
lilia.alberghina@unimib.it

**Dr Marta Cascante**
Departament de Bioquímica i Biologia Molecular
Facultat de Química
Universitat de Barcelona
Martí i Franquès, 1
08028 Barcelona • Spain
martacascante@ub.edu

**Prof. Gianni Cesareni**
Dipartimento di Biologia
Università di Roma Tor Vergata
Via della Ricerca Scientifica
00133 Rome • Italy
giovanni.cesareni@uniroma2.it

**Prof. Pierre de Meyts**
Hagedorn Research Institute
Niels Steensens Vej 6
2820 Gentofte • Denmark
pdm@hagedorn.dk

**Dr Igor Goryanin**
Centre for Intelligent Systems and their Applications
School of Informatics
University of Edinburgh
Appleton Tower
11 Crichton Street
Edinburgh EH8 9LE • United Kingdom
goryanin@inf.ed.ac.uk

**Prof. Leslie A. Grivell**
European Molecular Biology Organisation (EMBO)
Postfach 102240
Meyerhofstrasse
69117 Heidelberg • Germany
les.grivell@embo.org

**Prof. Mats Gyllenberg**
Department of Mathematics and Statistics
Rolf Nevanlinna Institute
University of Helsinki
PO Box 68 • 00014 Helsinki • Finland
mats.gyllenberg@helsinki.fi

**Prof. Stefan Hohmann**
Department of Cell and Molecular Biology/ Microbiology
Göteborg University
Box 462
40530 Göteborg • Sweden
hohmann@gmm.gu.se

**Dr Boris Kholodenko**
Department of Pathology, Anatomy
and Cell Biology
Thomas Jefferson University
1020 Locust St.
Philadelphia PA 19107 • United States
boris.kholodenko@jefferson.edu

**Dr Hiroaki Kitano**
The Systems Biology Institute (SBI)
M-31 Suite 6A, 6-31-15, Jingumae
Shibuya-ku
150-0001 Tokyo • Japan
kitano@symbio.jst.go.jp

**Prof. Karl Kuchler**
Max F. Perutz Laboratories
Department of Medical Biochemistry
Medical University of Vienna
Campus Vienne Biocenter
Dr Bohr-Gasse 9/2
1030 Vienna • Austria
karl.kuchler@meduniwien.ac.at

**Dr Iain William Mattaj**
EMBL
Meyerhofstrasse 1
69117 Heidelberg • Germany
mattaj@embl-heidelberg.de

**Dr Sergio Nasi**
Laboratorio Acidi Nucleici
Dipartimento Genetica e Biologia
Molecolare
Instituto di Biologia e Patologia Molecolari
Università la Sapienza
CNR
P. le A. Moro 5
00185 Rome • Italy
sergio.nasi@uniroma1.it

**Prof. Dieter Oesterhelt**
Abteilung Membranbiochemie
Max Planck Institute of Biochemistry
Am Klopferspitz 18a
82152 Martinsried • Germany
oesterhe@biochem.mpg.de

**Dr Uwe Sauer**
Institute for Molecular Systems Biology
ETH Hönggerberg
HPT C72
8093 Zürich • Switzerland
sauer@imsb.biol.ethz.ch

**Dr Frans J. G. van der Ouderaa**
Colworth Laboratory
Unilever Research
Sharnbrook
Bedford MK44 1LQ • United Kingdom
frans.van-der-ouderaa@unilever.com

**Prof. Roel van Driel**
Netherlands Institute for Systems Biology
and University of Amsterdam
Kruislaan 318
1098 SM Amsterdam
Netherlands
van.driel@science.uva.nl

**Prof. Hans V. Westerhoff**
School of Chemical Engineering and
Analytical Science
The University of Manchester
PO Box 88, Sackville St
Manchester, M60 1QD
United Kingdom
hans.westerhoff@manchester.ac.uk

## Industrial Experts

**Dr Ivan Arisi**
PKO-Bioinformatics
Lay Line Genomics
Via di Castel Romano 100
00128 Rome • Italy
i.arisi@laylinegenomics.com

**Dr Thomas Flores**
Global Community Partnerships
GlaxoSmithKline
980 Great West Road
Brentford, TW8 9GS • United Kingdom
thomas.p.flores@gsk.com

**Dr Adriano M Henney**
Director
AstraZeneca
Global Discovery Enabling Capabilities
& Sciences
Mereside
Alderley Park
Macclesfield
Cheshire SK10 4TG
United Kingdom

Adriano.Henney@astrazeneca.com

**Dr Gerhard Mulder**
Biolateral- Kennistking
Utrechtseweg 38
1381 GP WEESP • Netherlands
mulder285@planet.nl

**Dr John Savin**
Wendover Technology
8 Walnut Drive,
Wendover HP22 6RT
United Kingdom
js180757@yahoo.co.uk

**Prof. Dominic John Tildesley**
Unilever CSO, Home & Personal Care
Quarry Road East
Port Sunlight
Bebington Wirral CH63 3JW
United Kingdom
dominic.j.tildesley@unilever.com

**Dr Barry Utting**
107, Ivy Drive
Lightwater GU18 5YT Surrey
United Kingdom
butting@netcomuk.co.uk

## ESF Member Organisations and European Commission Representatives

**Dr Ioannis Economidis**
Biotechnology Unit (SMDE 9/13)
Directorate General XII
European Commission
Rue de la Loi 200
1040 Brussels • Belgium
ioannis/economidis@dg12.cec.be

**Dr Edvard Beem**
Deputy Director MRC/NWO
The Netherlands Organisation for Health
Research and Development (ZonMw)
PO Box 93245
2509 AE The Hague • Netherlands
beem@zonmw.nl

## ESF/EMRC and LESC and European Commission Representatives

**European Science Foundation**
1 quai Lezay-Marnésia
BP 90015
67080 Strasbourg cedex • France

**Dr John Marks**
Director of Science Strategy
(until March 2007)
jmarks@esf.org

**Dr Martina Hilger-Hildebrandt**
Head of Unit Life, Earth
and Environmental Sciences (LESC)
(until July 2005)
Hildebrandt@bremen.be

**Dr Carole Moquin-Pattey**
Head of Unit Medical Sciences (EMRC)
cmoquin-pattey@esf.org

**Dr Rosita Cottone**
Guest of EMRC
as Deputy Head of Unit at
Bundesministerium für Bildung
und Forschung (BMBF)
Referat 113 EU Politik, Eureka
Heinemannstrasse 2
53175 Bonn • Germany
rosita.cottone@bmbf.bund.de

# F. Appendix II

## Systems Biology: a Grand Challenge for Europe

EUROPEAN SCIENCE FOUNDATION

25
October 2005

# Systems Biology: a Grand Challenge for Europe

**European Science Foundation Policy Briefing**

## Contents



## Foreword

Thanks to spectacular advances in the "-omics"[1] disciplines and in information and communication technologies, the biosciences are heading for another revolution: Systems Biology, targeting pathways, cells, organs and complete organisms by integrating experimental data with computational and theoretical approaches. Systems Biology combines concepts from different scientific disciplines to obtain an integral understanding of complex biological systems in terms of their components and their interactions.

A better understanding of the underlying mechanisms of life should open perspectives for a deeper insight into human diseases and the development of new therapies. Thus, Systems Biology promises to have a profound impact on medical research, including drug development and biotechnology.

The integrative character of Systems Biology might be one reason for the fascination of younger scientists for this subject, encouraging them to overcome borders between disciplines and to foster collaboration far beyond national borders. European policy makers should seize this unique opportunity in order to counter the impending lack of young investigators. In addition, the expertise from Eastern European countries should be integrated into the framework of a European Research Area (ERA) in Systems Biology.

The success of European Systems Biology in a global arena will essentially depend on a better coordination of national and European efforts, a rapid adaptation of training schemes and long-term investment in cutting-edge research, which requires immediate and determined action.

This Science Policy Briefing represents an ambitious endeavour moderated by the European Science Foundation (ESF) to set up a European action plan towards a Grand Challenge for European Systems Biology.

**Bertil Andersson**
ESF Chief Executive

1. The term "-omics" describes the genome-wide study of entities, in this case the DNA, RNA, protein, or other molecular components of cells, tissues, or organisms.

## Introduction

Biological and biomedical research is undergoing revolutionary developments that have an immense and lasting impact on society. These developments involve other sciences, including physics, chemistry, mathematics and informatics. They enable us to know and measure the properties of the molecules that constitute life. We are now capable of revealing the complete sets of chemical reactions, interactions and dynamic structures through which molecules, cells and organs determine the functioning of living organisms, including humankind. Integrating the rapidly growing amounts of data available on these components and their interactions and generating understanding on how they govern life is termed *Systems Biology* or *Integrative Biology*.

As Systems Biology progresses, multifactorial diseases (such as diabetes, arthritis, heart failure and cancer) might be understood in terms of failure of molecular components to cooperate properly. Consequently, complex diseases may be approached and treated in a much more rational and effective way.

*It should be Europe's ambition to be at the forefront of pinpointing the molecular and systemic causes of diseases, aiming at the rational design of targeted therapies and drugs.*

This brief report is the outcome of the ESF Forward Look on Systems Biology in 2004-2005, involving an international high-level expert group, whose members agreed upon specific recommendations dedicated to the recent needs and requirements of European Systems Biology. The recommendations are summarised at the end of this report. Input from representatives of industry was taken into account. A full Forward Look report will provide more detailed insight into the considerations of the experts and their conclusions.

This briefing is intended to trigger targeted efforts of relevant stakeholders, including the ESF and its Member Organisations, governments, the European Commission, European industry and European academia.

## Impact of Systems Biology

Systems Biology involves the goal-oriented and systematic gathering of knowledge at all levels, from molecules to entire living organisms, and the subsequent integration into comprehensive and quantitative computer models. These will enable the accurate simulation of the processes of life[2]. Not only will this provide key insights into the functioning of living organisms, more importantly it will uncover underlying principles of how life itself operates. The societal and economic potential and value of such knowledge is immense.

### Health

Models based on Systems Biology may well facilitate the more accurate prediction of the properties and behaviour of living organisms in its physiological, pathological and technological context. Integrative Biology is expected to have a major impact on the paradigm shift in medical research towards dynamic multidrug treatments and personalised medicine[3]. It will allow a more cost-effective development of drugs and therapies for major diseases that plague our society today.

### Biotechnology

Molecular knowledge of microorganisms, plants and animals gathered from and integrated by Systems Biology-related research will enable a more predictable and rational approach to their genetic and metabolic engineering. This has the potential to change green and white biotechnology in fundamental ways. The rational development of new food products, the production of special chemicals and novel approaches to plant breeding come within reach, as the huge amounts of information about genes, proteins and metabolic pathways can now be integrated using mathematical models and integrative bioinformatics. Eventually, Systems Biology approaches may largely replace animal testing. Systems Biology is a key to innovation in the area of biotechnology. As the food industry is the largest sector in EU manufacturing, it may be among the Systems Biology outlets with the greatest economic potential. This may help to increase the average education base of this important branch of industry and consolidate one of Europe's most important export flows to the USA.

### Socio-economic potential

Systems Biology contributes to Europe's endeavour to take a leading position in the generation of new products such as drugs, therapies and biotech-based goods and knowledge. By increasing the safety and efficiency of biotechnological production processes, Systems Biology has the potential to strengthen Europe's economic competitiveness and to improve the quality of life for EU citizens.

*Europe should decide whether it chooses to lead in these areas or to become a net importer of Systems Biology-based knowledge and products.*

## European Strengths

Opportunities for Europe should be seen in their global and historical perspectives. Systems Biology has developed into two major branches[4]. One originates from functional genomics with its genome-wide sequencing and microarray analyses. This has developed rapidly in Japan and the USA. That approach provides insights into transcriptional patterns and related processes. However, it gives little information about how the dynamic interplay between the molecular components of living organisms results in their function or dysfunction.

The second branch is the bottom-up and hypothesis-driven approach of Systems Biology that aims at discovering how the properties of living organisms arise from the interactions between their molecular components. Historically, this type of approach has a solid basis in Europe (including the Eastern European countries) in theoretical biology and physiology. This is the branch of Systems Biology where Europe is traditionally strong.

## European Systems Biology: Current Status and International Context

Recently, national and transnational funding programmes on Systems Biology have been launched in Finland, France, Germany, the Netherlands, Switzerland and the UK and in some of the new EU member states[5]. The European Commission is funding collaborative research projects and coordination activities related to Systems Biology (e.g. ERASySBio) under the 6th Framework Research Programme (FP6)[6] and will implement the topic in FP7[7]. Further transnational activities have been set up in the EUREKA framework (InSysBio), the International *E. coli* Alliance (IECA), the SysMO initiative and the Yeast Systems Biology Network (YSBN). These projects represent important steps towards coordination, networking and sharing of resources in Systems Biology in Europe and clearly show that its importance is broadly acknowledged.

Major investments have recently been made in the USA in the field of Systems Biology. The Alliance for Cellular Signaling[8] initiated by the Nobel prize winner Alfred Gilman and the Genomes to Life Initiative[9] of the US Department of Energy represent prominent examples of large-scale activities.

---

2. Proof of this concept is the model of the human heart that integrates the molecular, cellular and organ levels in a quantitative and predictive manner (see Noble *Modelling the heart: from genes to cells to the whole organ.* Science 295:1678-1682 [2002]).
3. The ultimate goal of this new field is to tailor drugs and therapies precisely to the needs of individual patients. Personalised medicine has the important potential to improve the safety and efficacy of medical treatment.
4. See Westerhoff and Palsson *The evolution of molecular biology into Systems Biology.* Nat Biotech. 22:1249-52 (2004).
5. EUSYSBIO: Survey on the current status of Systems Biology-related research in New Member States & Associated Candidate Countries, the Russian Federation & Newly Independent States, Western Balkan Countries and the People's Republic of China
6. Examples of Systems Biology-oriented projects funded under FP6 are: EUSYSBIO, ERASysBio, BIOSIM, QUASI, BIOSAPIENS, COMBIO, EMI-CD, COSBICS, DIAMONDS.
7. COM (2005) 119 final 'Building the Europe of knowledge'
8. Alliance for Cellular Signaling: http://www.signaling-gateway.org/
9. DOE's Genomes to Life program: http://doegenomestolife.org/

Furthermore, the Institute of Systems Biology in Seattle[10] (co-founded by Leroy Hood) and the Molecular Sciences Institute in Berkeley[11], which was set up in 1996 by the Nobel Laureate Sidney Brenner, have achieved an international reputation as pioneering research institutes dedicated to Systems Biology-related research.

In 1996 Japan launched the E-Cell Project[12], followed two years later by the Kitano Symbiotic Systems Project[13], which includes research projects with the California Institute of Technology and various other research labs outside Japan. Other Asian countries, including Singapore, China and Taiwan have also entered the international competition by investing large sums in this promising domain.

*Despite the large number of ongoing European initiatives in the field of Systems Biology, it is evident that their scope is too limited to achieve major breakthroughs in health and biotechnology in a reasonable time frame and to compete successfully with the USA and the Far East.*

## The Grand Challenge: Giving Europe the Lead in Life Sciences

Despite the major investments in life sciences in Europe and elsewhere, existing initiatives lack a strategy that results in efficient integration of and synergy between different programmes and that can overcome the fragmentation of the R&D landscape in life sciences. The Grand Challenge for Europe is to strive for a leadership position in life sciences by creating breakthroughs in health and biotechnology. This is beyond the capabilities of individual European countries and industries. It requires a goal-oriented, pan-European effort and new creative ways of organising, coordinating and funding biomedical and biotechnological research in academia and industry. Furthermore, it needs a change in programme management and scientific culture by encouraging more collective efforts and cooperation.

In spite of attempts to do otherwise, most European granting systems for biomedical and biological research result in incremental advances by (small groups of) individual scientists, rather than an individual's contribution to larger efforts that may have much greater impact[14]. Meeting the Grand Challenge requires an approach that resembles the Human Genome Project, but has more dimensions and is of greater complexity. It calls for cross-border cooperation on a scale new to biological and biomedical research. Such a programme is highly ambitious and demands an objective-driven approach similar to that taken by the USA in the last century in landing a man on the Moon and returning him to Earth.

Meeting the Grand Challenge for European Systems Biology requires a robust multisource financing model. It will need novel ways of cooperation between industry and non-profit organisations, including governments, the European Commission and charities.

10. Institute for Systems Biology: http://www.systemsbiology.org/]
11. The Molecular Sciences Institute: http://www.molsci.org
12. E-Cell Project: http://www.e-cell.org/
13. The Kitano Symbiotic Systems Project: http://www.symbio.jst.go.jp/symbio/index.html
14. See Liu *Cell* 121:505 (2005).

## Making Choices

To achieve significant results in a reasonable time, it is essential that Europe initially concentrates on a limited number of targets. Choosing targets should be done carefully. They should create enthusiasm and at the same time expectations that can realistically be met in the context of science, industry and politics, and society at large. Linking choices to key issues in human health and in biotechnology is important.

The following considerations should be taken into account.

- Build on expertise that is available in European industry and academia.
- Concentrate on issues that contribute significantly to improving health of European citizens and economic competitiveness.
- Set realistic goals for a 5- to 10-year time horizon and broader objectives for a period of 20 years.
- Implement an approach that leads to the rapid development of generic knowledge, of tools in data acquisition and integration and of complex quantitative modelling.
- As the Grand Challenge progresses, take advantage of the generation and integration of new knowledge to pick up speed and address a broader range of issues, as was done in the Human Genome Project as it moved from a curiosity-driven to a technology-driven approach.

A first survey by the experts involved in this Forward Look identified the following *examples* of Europe-wide targets.

- Full molecular understanding of the bacterium *Lactococcus lactis*, the single cellular eukaryotic yeast *Saccharomyces cerevisiae* and the *human hepatocyte* (liver cell), which are biotechnologically and medically highly relevant. At the same time they represent robust systems for developing generic tools for integrating multilevel biological information. A concerted and goal-oriented European effort focusing on these issues could significantly enhance the health of EU citizens and industrial competitiveness.
- An ambitious and outstandingly relevant health issue is the *metabolic syndrome*, which includes obesity and type-II diabetes. A concerted European Systems Biology effort promises to make a substantial contribution to effective treatment and prevention of this complex disease in the next five to ten years, thereby addressing an immensely important biomedical, economic and societal issue.

## A European Action Plan

Since no single country or industry is able to manage such a large-scale initiative by itself, it is imperative to launch a European concerted and objective-driven effort. Such a project will differ decisively from all endeavours that currently define the European biomedical and biotechnological research landscape. Below, we address some key aspects of how a Europe-wide programme could meet the major requirements of the Grand Challenge. An action plan may consist of the following steps:

### 1. A task force to develop a European road map

A small task force should be created in which the major stakeholders are represented, including top scientists, industry, European science organisations and funding agencies, and representatives of the European Commission. This group of people should undertake the following:

- define road maps for the Grand Challenge for the next 10 and 20 years;
- draft a financial plan;
- propose an adequate research management structure that allows cost-effective, goal-oriented, large-scale research efforts;
- define new concepts for effective technology transfer and commercialisation of results;
- make a cost-benefit analysis.

The task force may be based on senior officers from European international funding and science organisations as well as from industries with major Systems Biology-related activities in Europe.

### 2. European Reference Laboratories (ERLs)

A cost-effective coordination of the European Systems Biology programme, as proposed here, requires the support of a consortium of European Reference Laboratories (ERLs). ERLs are research institutes that combine all relevant scientific disciplines and the know-how to provide outstanding expertise for core aspects of Systems Biology. They should be accessible to investigators from all over Europe to obtain the necessary introduction and training in this field.

ERLs should develop the scientific basis for standardisation and quality control of experimentation, which will then be implemented by the ESBO. In current biological and biomedical research, standardisation and quality control systems play only a limited role. As a result, data obtained by different research groups cannot easily be combined. This is particularly true for quantitative studies. Since heading for a complete understanding of complex biological systems crucially depends on properly handling such data, it is imperative that academia and industry define adequate standards and that these standards are implemented into databases. At the same time such standards should remain adaptable to changes as knowledge increases.

ERLs should be responsible for the following aspects of the Grand Challenge:

- development of generic tools for data integration and data storage that can be implemented by the ESBO;
- development of tools for the validation of data generated by the participating research groups;
- development of relevant training programmes in collaboration with the ESBO;
- providing excellent experimental facilities and expert advice on Systems Biology and its methodologies.

### 3. Cooperation between industry, academia and charities

A European large-scale effort as outlined in this paper requires a re-thinking of the present practices of cooperation between industry, academia and charities. All parties should contribute to and benefit from the programme but with a clear understanding of what belongs to the public domain and what is rightfully the proprietary information of private companies. An appropriate and sustainable public-private partnership model has to be developed. Re-evaluation of the notion of intellectual property will be investigated, distinguishing accurately between efforts leading to generic pre-competitive tools and those addressing specific commercial biomedical and biotechnological objectives.

### 4. Public acceptance

A major European effort as addressed in this report will be viable only if the general public accepts and endorses the underlying ideas and goals of the Grand Challenge for European Systems Biology. To achieve this, careful communication and explanation will be necessary from the onset, with initiatives in social sciences and ethics accompanying the Systems Biology research. A broad and open debate will enable the public to be aware of the socio-economic benefits, but also of any potential risks of the Grand Challenge, allowing a balanced and objective monitoring of the further development of this field from the very beginning.

### 5. Training and education

In contrast to the present practice of educating scientists in the classical disciplines, the Systems Biology approach requires new thinking across scientific borders. New interdisciplinary BSc, MSc and PhD teaching and training programmes should be implemented as a matter of urgency to address this issue. In this respect, an open debate with responsible European stakeholders and higher education institutions (e.g. universities) should be initiated.

>>>

### 6. Financing the Grand Challenge

The Europe-wide approach proposed here will require a higher level of funding than that provided by vehicles available now. We propose that a new financial model is developed based on cooperation between national, international, industrial, charities and EC-related organisations. It will be a major challenge to align this heterogeneous set of parties so as to produce a synergistic cooperative programme.

### 7. A European Systems Biology Office (ESBO)

Systems Biology depends on a close cooperation of a broad range of scientific disciplines, including medicine, pharmaceutics, biology, physics, chemistry, mathematics, nanosciences and ICT. Furthermore, a European effort of the scale foreseen here will involve numerous research groups from all over Europe, in connection with non-European initiatives.

This cross-border cooperation should be supported by an independent coordinating body that has an overview of ongoing activities, a timetable of planned research and an awareness of areas where information is missing and how these gaps should be filled. Furthermore, this body, named the European Systems Biology Office (ESBO), should foster international networking and be responsible for the management of public relations affairs of the Grand Challenge. The Office should mediate between the scientific, industrial and political stakeholders and provide them with expert advice.

The ESBO should raise awareness of ethical, socio-economic and intellectual property right (IPR) issues relevant to a European Systems Biology programme. Other important tasks may include questions related to training of undergraduate and postgraduate students and other scientists, as this is seen as a major bottleneck in Systems Biology. Furthermore, the ESBO should be charged with data storage and the implementation of a quality control system and standardisation in experimentation.

## Recommendations

- Establish a task force to define a European road map. This could be mediated by the ESF.
- Organise a high-level strategic workshop for national and European funding agencies, industry and charities to develop an appropriate financial model for the Grand Challenge. In addition, this high-level group should work out a strategy plan aimed at the integration of ongoing activities in the field.
- Involve specific target groups, such as patient organisations, health insurances, regulatory bodies and publishing groups. It is important that they contribute to the development of the Grand Challenge.
- Concentrate on a small number of specific targets that are of major importance for human health and have a substantial impact on economic growth. The chosen targets should allow the exploitation of European strengths and expertise. Defined milestones should be achievable within 10 years.
- Perform business case studies to explore the costs and benefits of Systems Biology in general and of the Grand Challenge in particular. A review of the socio-economic value of such a large-scale approach would provide reliable data to funding institutions and policy makers on the expected outcome of the programme.
- Set up a European Systems Biology Office (ESBO) and identify research institutes that could serve as European Reference Laboratories (ERLs), providing appropriate expertise, resources and facilities for the Grand Challenge.
- Stimulate transnational collaboration and networking with other ongoing Systems Biology initiatives throughout the world to act in a synergistic manner and profit from their experiences.
- Establish a European training programme in Systems Biology to provide young investigators with the appropriate background. The programme could include regular summer schools, courses, graduate schools, BSc and MSc programmes.

**EUROPEAN SCIENCE FOUNDATION**

**European Science Foundation**
1 quai Lezay-Marnésia, BP 90015
67080 Strasbourg cedex, France
Tel: +33 (0)3 88 76 71 00
Fax: +33 (0)3 88 37 05 32
www.esf.org

# ESF Forward Look on Systems Biology

## Members of the high-level expert group

**Roel van Driel** *(Co-Chair)*, University of Amsterdam, Amsterdam, Netherlands
**Hans Westerhoff** *(Co-Chair)*, University of Manchester, Manchester, United Kingdom, Vrije Universiteit Amsterdam, and University of Amsterdam, Amsterdam, Netherlands
**Lilia Alberghina,** University Milano-Bicocca, Italy
**Gianni Cesareni**, Università di Roma, Italy
**Roland Eils**, DKFZ, Heidelberg, Germany
**Frank Gannon**, European Molecular Biology Organisation, Heidelberg, Germany
**Mats Gyllenberg**, University of Helsinki, Helsinki, Finland
**Stefan Hohmann**, Göteborg University, Göteborg, Sweden
**Karl Kuchler**, University and BioCenter of Vienna, Vienna, Austria
**Sergio Nasi**, Università la Sapienza, Roma, Italy
**Bela Novak**, Budapest University of Technology and Economics, Budapest, Hungary
**Dieter Oesterhelt**, Max Planck Institute of Biochemistry, Martinsried, Germany
**Uwe Sauer**, ETH, Zurich, Switzerland
**Barry Utting**, consultant, Surrey, United Kingdom

## Advisers from industry

**Adriano Henney**, AstraZeneca
**Frank Tobin**, GlaxoSmithKline
**Pierre de Meyts**, Novo Nordisk
**John Savin**, Physiomics
**Dominic Tildesley**, Unilever
**Frans van der Ouderaa**, Unilever
**Wanda Andreoni**, IBM

## ESF office in Strasbourg

**Martina Hilger-Hildebrandt**, Head of the Life, Earth and Environmental Sciences Unit (LESC)
**Carole Moquin-Pattey**, Head of the Medical Sciences Unit (EMRC)
**Rosita Cottone**, Deputy Head of Unit at Federal Ministry of Education and Research (BMBF), guest at EMRC Unit

European Science Foundation Policy Briefings are published by the European Science Foundation (ESF). They address selected science policy issues of key concern to the Foundation's Member Organisations and the wider scientific community. By drawing on the advice and expertise of the ESF's membership, the briefings aim both to provide information and to promote discussion.

Further information on the ESF's scientific and science policy activities is available from the Communication and Information Unit,
European Science Foundation
1 quai Lezay-Marnésia, BP 90015
67080 Strasbourg cedex, France
Tel: +33 (0)3 88 76 71 25
Fax: +33 (0)3 88 37 05 32
E-mail: communications@esf.org
www.esf.org

ISRN ESF-SPB-05-25-FR+ENG


Dans les villes : November 2005 – Print run: 2000

# G. Appendix III

*Please note that this is a necessarily incomplete and in part personal selection of the editor.*

## 1. Selected reviews on Systems Biology published in 2006


Albeck, J.G., G. MacBeath, F.M. White, P.K. Sorger, D.A. Lauffenburger and S. Gaudet. 2006.
Collecting and organizing systematic sets of protein data. *Nat Rev Mol Cell Biol*. 7:803-12.

Aldridge, B.B., J.M. Burke, D.A. Lauffenburger and P.K. Sorger. 2006.
Physicochemical modelling of cell signalling pathways. *Nat Cell Biol.* 8:1195-203.

Aloy, P. and R.B. Russell. 2006.
Structural Systems Biology: modelling protein interactions. *Nat Rev Mol Cell Biol*. 7:188-97.

Alves, R., F. Antunes and A. Salvador. 2006.
Tools for kinetic modeling of biochemical networks. *Nat Biotechnol*. 24:667-72.

D. Baker, M., P.M. Wolanin and J.B. Stock. 2006.
Systems Biology of bacterial chemotaxis. *Curr Opin Microbiol*. 9:187-192.

Brazma, A., M. Krestyaninova and U. Sarkans. 2006.
Standards for Systems Biology. *Nat Rev Genet*. 7:593-605.

Brent, R. and J. Bruck. 2006.
2020 computing: can computers help to explain biology? *Nature*. 440:416-7.

Cho, C.R., M. Labow, M. Reinhardt, J. van Oostrum and M.C. Peitsch. 2006.
The application of Systems Biology to drug discovery. *Curr Opin Chem Biol*. 10:294-302.

Cornish Bowden, A. 2006.
Putting the systems back into Systems Biology. *Perspect Biol Med.* 49:475-89.

Fitzgerald, J.B., B. Schoeberl, U.B. Nielsen and P.K. Sorger. 2006.
Systems Biology and combination therapy in the quest for clinical efficacy. *Nat Chem Biol.* 2:458-66.

Forst, C.V. 2006.
Host-pathogen Systems Biology. *Drug Discov Today.* 11:220-7.

Griffith, L.G. and M.A. Swartz. 2006.
Capturing complex 3D tissue physiology in vitro. *Nat Rev Mol Cell Biol*. 7:211-24.

Janes, K.A. and M.B. Yaffe. 2006. Data-driven modelling of signal-transduction networks. *Nat Rev Mol Cell Biol*. 7:820-8.

Jaqaman, K. and G. Danuser. 2006.
Linking data to models: data regression. *Nat Rev Mol Cell Biol*. 7:813-19.

Joyce, A.R. and B.O. Palsson. 2006.
The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol*. 7:198-210.

Karsenti, E., F. Nedelec, and T. Surrey. 2006.
Modelling microtubule patterns. *Nat Cell Biol*. 8:1204-11.

Kell, D.B. 2006.
Metabolomics, modelling and machine learning in Systems Biology - towards an understanding of the languages of cells. *Febs J.* 273:873-94.

Kersey, P. and R. Apweiler. 2006.
Linking publication, gene and protein data. *Nat Cell Biol*. 8:1183-9.

Kholodenko, B.N. 2006.
Cell-signalling dynamics in time and space. Nat *Rev Mol Cell Biol*. 7:165-76.

Kitano, H. 2006.
Computational cellular dynamics: a network-physics integral. *Nat Rev Mol Cell Biol.* 7:163.

Michelson, S. 2006.
The impact of Systems Biology and biosimulation on drug discovery and development. *Mol Biosyst*. 2:288-91.

Moffat, J. and D.M. Sabatini. 2006.
Building mammalian signalling pathways with RNAi screens. *Nat Rev Mol Cell Biol*. 7:177-87.

Mustacchi, R., S. Hohmann and J. Nielsen. 2006.
Yeast Systems Biology to unravel the network of life. *Yeast*. 23:227-38.

Nickell, S., C. Kofler, A.P. Leis and W. Baumeister. 2006.
A visual approach to proteomics. *Nat Rev Mol Cell Biol*. 7:225-30.

Pepperkok, R. and J. Ellenberg. 2006.
Innovation - High-throughput fluorescence microscopy for Systems Biology. *Nat Rev Mol Cell Biol*. 7:690-96.

Swedlow, J.R., S.E. Lewis and I.G. Goldberg. 2006.
Modelling data across labs, genomes, space and time. *Nat Cell Biol*. 8:1190-4.

Trewavas, A. 2006.
A brief history of Systems Biology –'Every object that biology studies is a system of systems' Francois Jacob (1974). *Plant Cell.* 18:2420-30.

Wiley, H.S. 2006.
Systems Biology – beyond the buzz. *The Scientist*. 20:52-7.

Wingreen, N., and D. Botstein. 2006.
Back to the future: education for systems-level biologists. *Nat Rev Mol Cell Biol*. 7:829-32.

Winzeler, E.A. 2006.
Applied Systems Biology and malaria. *Nat Rev Microbiol*. 4:145-51.


## 2. Selected books


An introduction to Systems Biology: Design principles of biological circuits by Uri Alon, June 2006, Chapman&Hall/CRC, Taylor and Francis Group, ISBN: 1584886420

Stochastic Modelling for Systems Biology by Darren James Wilkinson, April 2006, Chapman & Hall/CRC Press, ISBN: 1-58488-540-8

Cancer Bioinformatics: From Therapy Design to Treatment by Sylvia Nagl (Ed.), March 2006, John Wiley and Sons Ltd, ISBN: 0-470-86304-8

Systems Modeling in Cellular Biology by Zoltan Szallasi, Joerg Stelling, Vipul Periwal, March 2006, MIT Press, ISBN: 0-262-19548-8

Systems Biology: Properties of Reconstructed Networks by Bernard Palsson, January 2006, published by Cambridge Univ. Press, ISBN: 0521859034

Computational Systems Biology by Andres Kriete and Roland Eils (eds.), Elsevier 2005, ISBN: 0-12-088786-X

Systems Biology: Definitions and Perspectives Alberghina, L. and Westerhoff, H., Springer, 2005, ISBN: 354022968X

Metabolome Analyses: Strategies for Systems Biology Vaidyanathan, S. et al (eds.), Springer-Verlag, 2005, ISBN 0387252398

Systems Biology in Practice: Concepts, Implementation and Application Klipp, E et al., John Wiley & Sons Inc. 2005, ISBN 3527310789

Foundations of Systems Biology Kitano, H. (ed.), The MIT Press 2001-10-15, ISBN: 0262112663

Systems Modeling in Cellular Biology Szallasi, Z. et al., MIT Press, 2006, ISBN: 0-262-19548-8

# G. Appendix III

## 3. Congresses and workshops

### • 2006

CSBi Symposium 2006
Systems Biology of the Stem Cell
January 12, 2006 Boston, USA.
MIT, Kresge Auditorium (W16)
http://csbi.mit.edu/events/annualsymposium

The First ETH Symposium on Synthetic Biology
February, 24, 2006,
Zurich, Switzerland
Conference Web site

Keystone Symposia on Systems Biology: Integrating Biology, Technology and Computation
March 5- 0, 2006 Taos, New Mexico, US.
http://www.keystonesymposia.org/Meetings/ViewMeetings.cfm?MeetingID=789

Cold Spring Harbour Symposium on Systems Biology: Global Regulation of Gene Expression
March 23-26, 2006 Cold Spring Harbour, NY, USA.
http://meetings.cshl.edu/meetings/systems06.shtml

International Specialised Symposium on Yeasts (ISSY25)
Systems Biology of Yeasts – From Models to Applications
June 18-21, 2006 Helsinki, Finland
http://issy25.vtt.fi/

The First Conference on Systems Biology of Mammalian Cells (SBMC 2006)
July 12-14, 2006 Heidelberg, Germany
http://www.sbmc06.de/

The Fifth International Conferences on Bioinformatics of Genome Regulation and Structure (BGRS 2006)
July 16-22, 2006 Novosibirsk, Russia
http://www.bionet.nsc.ru/meeting/bgrs2006/

The First International Conference on Computational Systems Biology
July 20-23, 2006 Shanghai, China
http://life.fudan.edu.cn/ICCSB

Cold Spring Harbor/Wellcome Trust Conference
INTERACTOME NETWORKS
Mapping macromolecular interactions in the cell
August 30-September 3, 2006
Wellcome Trust Genome Campus, Hinxton, UK
http://meetings.cshl.edu/meetings/interuk06.shtml

The Third Integrative Bioinformatics Workshop
September 4-6, 2006 Rothamsted Research, Harpenden, Hertfordshire, United Kingdom
http://www.rothamsted.bbsrc.ac.uk/bab/conf/ibiof/

2006 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB 2006)
September 28-29, 2006
Renaissance Hotel Downtown, Toronto, Ontario, Canada
http://www.cibcb.org

International Conference on Systems Biology 2006 (ICSB-2006)
Oct 8-13, Yokohama, Japan
http://www.icsb-2006.org

International Conference on Computational Methods in Systems Biology
18-19 October 2006,
The Microsoft Research-University of Trento,
Centre for Computational and Systems Biology,
Trento, Italy
CMSB-06 Web site

The 3rd International *E. Coli* Alliance Conference on Systems Biology (IECA 2006)
Oct 31-Nov 3 2006, Jeju Island, Korea
http://www.ieca2006.org/

The second annual Systems Biology Symposium entitled 'Progress in Systems Biology'
Nov 9-10 2006, Ottawa Institute of Systems Biology
http://mededu.med.uottawa.ca/oisb/sympindex.htm

'Systems Biology in drug discovery'
- A part of the 2nd Modern Drug Discovery & Development Summit
Dec 4-6 2006, Pennsylvania Convention Center, Philadelphia, PA
http://gtcbio.com/confpage.asp?cid=8

### • 2007

BioSysBio 2007 – Systems Biology, Bioinformatics, Synthetic Biology
Jan 11-13, 2007 Manchester, UK
http://www.biosysbio.com

Winter School on Systems Biology for Medical Applications
February 27-March 2, 2007
Puerto de la Cruz, Tenerife, Spain

Computational Cell Biology
March 6-9, 2007
Cold Spring Harbor Laboratory, NY, USA

FEBS-SysBio 2007
March 10-16, 2007
Sport & Erlebnis Hotel Gosau, Austria
http://www.univie.ac.at/sysbio2007/

The Second Annual Conference of the Association for General and Applied Microbiology
April 1-4, 2007 Osnabrück, Germany

5th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics
April 11-13, 2007, Valencia, Spain
http://evonet.lri.fr/TikiWiki/tiki-index.php?page=CFP+EvoBIO+2007

Workshop 'Storage and Annotation of Reaction Kinetics Data'
May 21-23, 2007 Heidelberg, Germany
http://projects.eml.org/sdbv/events/workshop2007/index_html

2nd International Course in Yeast Systems Biology
June 4-21, 2007 Gothenburg, Sweden
http://www.icysb.org

2nd Conference Foundations of Systems Biology in Engineering (FOSBE 2007)
September 9-12, 2007 Stuttgart, Germany
http://www.fosbe.org

International Conference on Systems Biology (ICSB-2007)
October 1-6, 2007 Long Beach, California, USA.
http://icsb-2007.org

### • 2008

Systems Biology
ESF-UB Conference in Biomedicine
12-17 April 2008
Hotel Eden Roc, Sant Feliu de Guixols (Costa Brava), Spain
Chairs: L. Serrano (Heidelberg), R. van Driel (Amsterdam) & R. Aebersold (Zürich)

## 4. Examples of National, European and International Initiatives

### National

#### Finland

The Academy of Finland and Finland's National Technology Agency, Tekes, have a joint **Systems Biology Initiative**. To date, €9 Million has been committed.
http://www.aka.fi/index.asp?id=031ed13451eb43f69d8f747a232c3311

#### Germany

**HepatoSys**, funded by the German Ministry for Research and Education (BMBF) with €13 Million for the first 3 years and €24 Million promised for the following three, is a nationwide German systems biology project. The long-term goal is the creation of a virtual liver cell. The initiative comprises 30 partners from academia and industry. http://www.systembiologie.de/en/index.html

**BIOMS**, Heidelberg's Center for Modelling and Simulation in the Biosciences is a collaboration between the European Media Laboratory, the European Molecular Biology Laboratory, the Max Planck Institute for Medical Research, and the University of Heidelberg, in Germany.
http://www.bioms.de/

**Systems Biology**, based at the University of Stuttgart, Germany, is investigating prokaryotic and eukaryotic processes. The Max Planck Institute for Dynamic of Complex Technical Systems in Magdeburg is a partner in this initiative.
http://www.sysbio.de/

#### Switzerland

**System X** effort of the University of Basel, the University of Zurich, and the ETH Zurich (Swiss Federal Institute of Technology in Zurich). Systems X funded projects will be hosted throughout Switzerland but will be concentrated in two new institutes: the Center for Biosystems Science and Engineering in Basel, and the Institute for Molecular Systems Biology at the ETH Zurich. The pharmaceutical firm Roche is an industrial partner.
http://www.systemsx.ch/

#### The Netherlands

**Silicon Cell Initiative** is a Netherlands-based initiative that aims to develop precise computer models of living cells. Such models will be stored centrally so they are accessible for in silico experimentation.
http://www.siliconcell.net/sica/

### European

#### STREP/CA/SSA

**COMBIO** is an integrative approach to cellular signalling and control process.
http://www.pdg.cnb.uam.es/COMBIO

**COSBICS** - Computational Systems Biology in Cellular Signalling- E.U initiative focused on better understanding cell signalling in the context of tumor development. Partners are based in Germany, Bulgaria, Scotland, and Spain.
http://www.sbi.uni-rostock.de/cosbics/

**DIAMONDS** is a E.U. project that aims to demonstrate the power of a Systems Biology approach to study the regulatory network structure of the most fundamental biological process in eukaryotes: the cell cycle.
http://www.cordis.lu/lifescihealth/genomics/home.htm

**EMI-CD** is a European modeling initiative combating complex diseases.
http://pybios.molgen.mpg.de/EMICD/

**ESBIC-D** - European Systems Biology Initiative combating complex diseases using cancer as a prototypical problem.
http://pybios.molgen.mpg.de/ESBIC-D

**YSBN** stand for the Yeast Systems Biology Network and aims to bring researchers working on yeast systems biology together. http://www.gmm.gu.se/YSBN/

**BIOBRIDGE** - is on Integrative Genomics and Chronic Disease Phenotypes
http://tbd

**QUASI & AMPKIN** are respectively i) a EU-funded project that joins experimental and computational approaches to better understand signalling processes in yeast. Partners are from Sweden; Spain, Switzerland, Austria, and Germany; ii) A EU-funded prject with the overall objective to generate mathematical models of the AMPK pathway to be used for drug target identification and drug screening.
http://www.idp.mdh.se/quasi/
http://www.gmm.gu.se/AMPKIN/

**STREPTomics** stands for "Systems biology strategies and metabolome engineering for the enhanced production of recombinant proteins in Streptomyces", is a FP6 research program initiated in January 2007.
http://www.streptomics.org/

**SYSBIOMED** - is a EU-funded Strategic Support Action within the FP6. Its core objective is to explore the potential of systems biology for medical research, therapy and drug development. The other main goal is the formation of a network of young scientists who define the framework for future research programmes in 'Medical Systems Biology' (MSB)
http://www.sysbiomed.de/

**VALAPODYN** is aimed at generating validated predictive dynamic models of complex intracellular pathways related to cell death and survival.
http://www.recherche.ulg.ac.be/ard_eu/6fp_ulg.html

#### NOE/IP

**ENFIN** (Experimental Network for Functional Integration) a €9 Million E.U. initiative that started in 2005, is funded for 5 years and consists of 20 partner groups from 13 countries with a mix of theorists and experimentalists.
http://www.enfin.org

**BaSysBio** (Bacillus Systems Biology) has started on 1st November 2006 and involves 15 European research organisations and an Australian university for the purpose of developing "systems biology" techniques. It aims to study the global regulation of gene transcription in a model bacterium: Bacillus subtilis.
http://www.basysbio.eu/

**BioSim** is a E.U-funded network of researchers investigating biomedical questions-particularly pharmacological ones- using a systems biology approach. Forty partners across Europe are involved, including 10 industrial and four regulatory bodies. The effort is coordinated by the Technical University of Denmark.
http://biosim.fysik.dtu.dk:8080/biosim/showorganisation0.jsp?menu=showOrganisation

**EUTRACC** is a project of systems biology for transcription in mice
http://www.eutracc.eu/

**AGRON-omics** stands for "Arabidopsis GROwth Network integrating OMICS technologies" and has been launched in November 2006.
http://www.agron-omics.eu/

#### Others:

**NISIS** (Nature Inspired Smart Information Systems) is a E.U-funded network with a substantial systems biology component. The initiative includes members from all over Europe.
http://www.nisis.risk-technologies.com/

**ERASysBio** is a transnational funding initiative to support the convergence of life sciences with information technology & systems science.
http://www.erasysbio.net

**SysMo**-Systems Biology of Micro-organisms- is a collaborative effort between Austria, Finland, Germany, the Netherlands, Spain, Norway, and the U.K. A host of micro-organisms are under investigation. There is a sizable industrial component.
http://www.sysmo.net/

# G. Appendix III

### Marie Curie Actions

**NucSys** (Systems biology of nuclear receptors: A nutrigenomic approach to aging-related diseases) is a E.U-funded Marie Curie Research Training Network. This group is using a systems biology approach to explore how nuclear receptor transcription factors orchestrate responses to environmental changes that cells experience. The consortium includes partners from Finland, the UK, the Netherlands, Belgium, Spain, Germany, Austria, and Italy.
http://www.uku.fi/nucsys/

**SYSTEM** (Systems biology of stem cell function in Arabidopsis thaliana) is a Marie Curie Early Stage Research Training Network funded by the E.U where students spend part of their time in a dry and the other part in a wet lab. Partners are Belgium, France, Germany, Poland, Switzerland, UK.
http://www.sy-stem.ethz.ch/

## International

**IRGT** (International Research Training Group) This Genomics and Systems Biology of Molecular Networks is formed by groups from Humboldt University Berlin, Free University Berlin, MPI for Molecular Genetics, Boston University and Kyoto University and it focuses on the education of doctoral students. The German part is funded by the DFG.
http://www2.hu-berlin.de/biologie/irtg/

**Genes to Cognition** is a neuroscience systems biology initiative. The Welcome Trust's Sanger Institute in Hinxton, U.K, and the University of Edinburgh are the main partners, but the project includes collaborators worldwide.
http://www.genes2cognition.org/

# 5. Selected Systems Biology Institutes

### Canada

The **Ottawa Institute of Systems Biology** is located at the Faculty of Medicine of the University of Ottawa, Canada. The aim of the institute is to develop and apply systems biology to biological studies relevant to human diseases.
http://intermed.med.uottawa.ca/Associations/OISB

### Ireland

The **Centre for Systems Biology** is located at Trinity College in Dublin.
http://www.systemsbiologyireland.org

### Japan

The **Institute of Advanced Biosciences** established in the spring of 2001, is a pioneering research institute part of Keio University.
http://www.iab.keio.ac.jp/

The **Systems Biology Institute** in Tokyo aims to apply systems biology to medicine and engineering.
http://sbi.jp/

### The Netherlands

The **Netherlands Institute for Systems Biology** in Amsterdam concentrates on the integration of biomolecular networks in metabolism, signal transduction, gene expression and force and cellular shape generation.
http://www/sysbio.nl

### Portugal

The Systems Biology Unit is located in a biotechnology research park. The Unit is devoted to pursuing fundamental and applied research in the Life Sciences from a systems biology standpoint.
http://www.biocant.pt

### Spain

The **Center for Genomic Regulation** (CRG) is located at the Barcelona Biomedical Research Park (PRBB) building
http://www.crg.es

### Switzerland

The **Institute for Molecular Systems Biology** is located at the ETH in Zürich.
http://www.cbb.ethz.ch/research/units/biology/imsb

### United Kingdom

The **Centre for Integrative Systems Biology** brings together researchers from several faculties within the Imperial College in London. One of the main topics will be host-pathogen interaction. The project received €3.5 Million in funding from the UK's Biotechnology and Biological Sciences Research Council (BBSRC) and €0.7 Million from the Engineering and Physical Sciences Research Council (EPSRC). http://www3.imperial.ac.uk/cisbic/about
http://www.doc.ic.ac.uk/bioinformatics/CISB/

The **Centres for Integrative Systems** are located at the Universities of Manchester and Newcastle. €9.5 Million funding from BBSRC and EPSRC to concentrate on yeast and the ageing cell.
http://www.mcisb.org/

### United States

The **Institute for Systems Biology (ISB)** in Seattle, Washington, has 170 scientists from many different disciplines working on a wide range of scientific problems. ISB was the first major institution to focus exclusively on systems biology.
http://www.systemsbiology.org/

The **Department of Systems Biology** at Harvard Medical School in Boston has 20 faculty members (including departmental faculty, affiliated faculty, visiting faculty, Instructors, and lecturers) working on biomedical research projects.
http://sysbio.med.harvard.edu/

The **Systems Biology Group** at the **Pacific Northwest National Laboratory** in Richland, Washington, employs 90-plus staff scientists working on proteomics, microbial-cell dynamics, cell and molecular imaging and spectroscopy, computational biology, and bioinformatics. Particular foci include biomolecular systems, pathogen biology, computational science, and environmental biomarkers.
http://www.sysbio.org/index.stm

The **Massachusetts Institute of Technology's Computational and Systems Biology** Initiative applies large-scale numerical methods to the study of molecular, cellular, and structural biology. Particular areas of interest include gene finding and analysis, protein design, network-based signal analysis, and image informatics.
http://csbi.mit.edu/

The **Molecular Sciences Institute in Berkeley**, California, combines genomic experimentation and computer modeling to predict the behaviour of cells and organisms in response to genetic and environmental changes.
http://www.molsci.org/

# 6. Systems Biology Organisation

A portal site for Systems Biology:
http://www.systems-biology.org/

# 7. World Technology Evaluation Center (WTEC) Inventory

http://www.wtec.org/sysbio/report/

Published by the European Science Foundation.
August 2007.
Printing: IREG, Strasbourg.
ISBN: 2-912049-65-2

EUROPEAN SCIENCE FOUNDATION
SETTING SCIENCE AGENDAS FOR EUROPE

1 quai Lezay-Marnésia | BP 90015
67080 Strasbourg cedex | France
Tel: +33 (0)3 88 76 71 00 | Fax: +33 (0)3 88 37 05 32
**www.esf.org**

Print run: 2 500 – August 2007