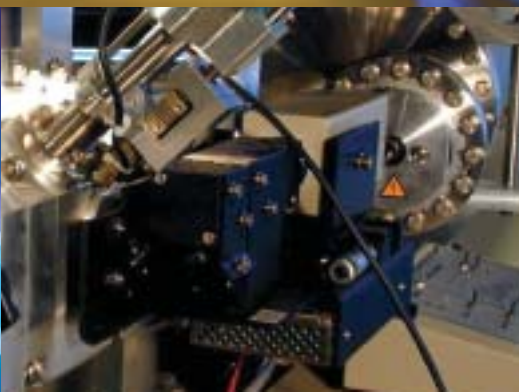
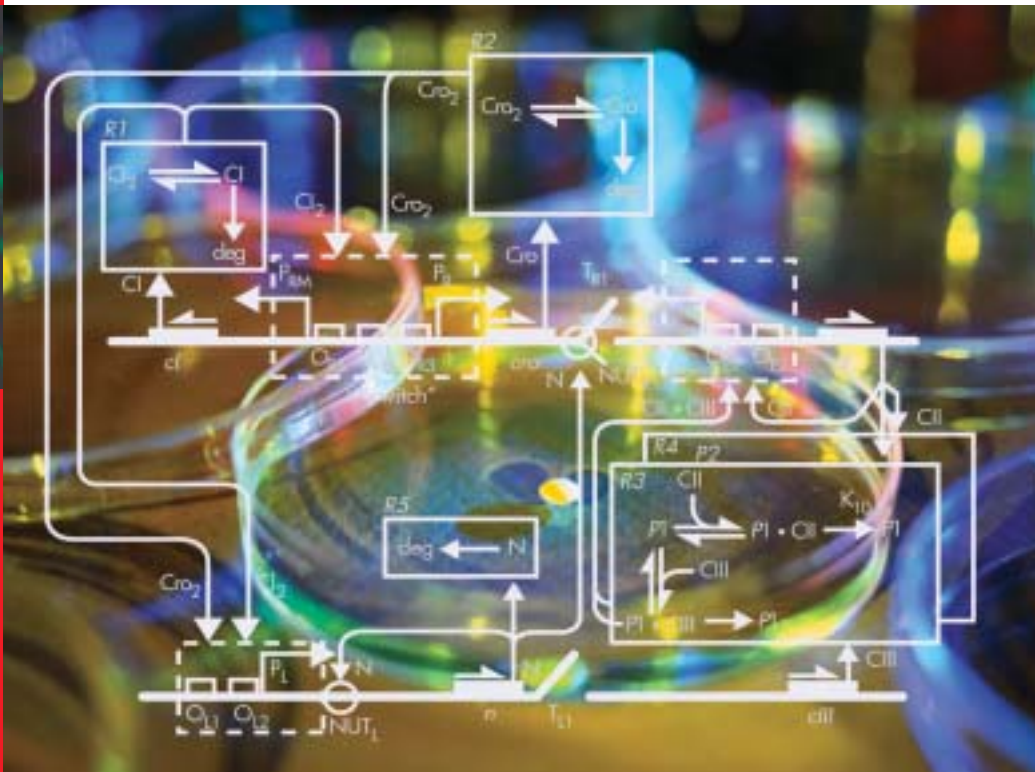




Systems of Life

Systems Biology

BMBF PUBLIK



Published by

Federal Ministry of Education
and Research (BMBF)
Public Relations Division
53170 Bonn

Orders

by mail to
BMBF
Postfach 30 02 35
53182 Bonn

by phone: 01805-BMBF02
or 01805-262302
by fax: 01805-BMBF03
or 01805-262303
0,12 Euro/min.

e-mail: books@bmbf.bund.de
Internet: www.bmbf.de

Edited by

Dr. Karsten Schürle
DECHEMA e. V.
Informationssekretariat Biotechnologie ISB

Author

Dr. Thomas Reiß
Fraunhofer-Institut für Systemtechnik
und Innovationsforschung

Layout

Christian Beck

Printed by

Druckerei Hörning, 69123 Heidelberg

September 2002

Printed on recycled paper

Copyrights

Bayer	6
Birgit Schoeberl	15
BMBF	27, 29, 30
DECHEMA e.V.	30
Boeing	9
DECODON AG;	18
Dietrich Keppler, Anne Nies/DKFZ	26
Digital Stock	4, 8, 18, 19
DOE	25
DOE/ AdamArkin	12
Eric Davidson	13
GEN	17
ISB Seattle	23
Karsten Schürle	5, 15, 16
Jörg Kopka	14
K. Maier, W. Tuszynski/ Universität Oldenburg	17
Monsanto	19
NASA	20
Nature	6
Pat Browne	16
Rechenzentrum Garching	10
Roman Bezjak	24
Rosita Cottone	28
Sigma-Aldrich	7



Federal Ministry
of Education
and Research

Systems of Life

Systems Biology



Introduction	3
1. Where we started: The biosciences at the end of the 20th century	4
High-throughput technology in genome research	6
2. Systems biology - attempting to understand the complexity of life	
Data capture	11
Modelling	12
Computer models	12
The research approach in systems biology	15
The basis for systems biology	16
Benefits of systems biology	18
3. International activities in systems biology	20
Systems biology in other European countries	21
Systems biology in Japan	22
Systems biology in Canada and the USA	22
4. The Systems of Life - Systems Biology research and funding programme	26
A new research priority emerges	27
Systems of Life - Systems Biology	28
Complementing other BMBF research activities	30
Systems biology links	32
Glossary	33
Literature	34

Only a few years ago large electronic appliances such as colour TV sets were made up of thousands of separate components: transformers, transistors, condensers, resistors, coils, diodes, switches, buttons, the picture tube and many other things. Let's assume an electronics specialist wants to reconstruct such an appliance, based on a complete parts list containing every single component with its precisely specified characteristics. Would he be able to join all the parts together to form a properly functioning system by just using the information available?

Certainly not! For without knowing the underlying circuit diagram that shows how the various components have to be wired and connected with each other the expert will lack the critical data enabling him to attain his goal.

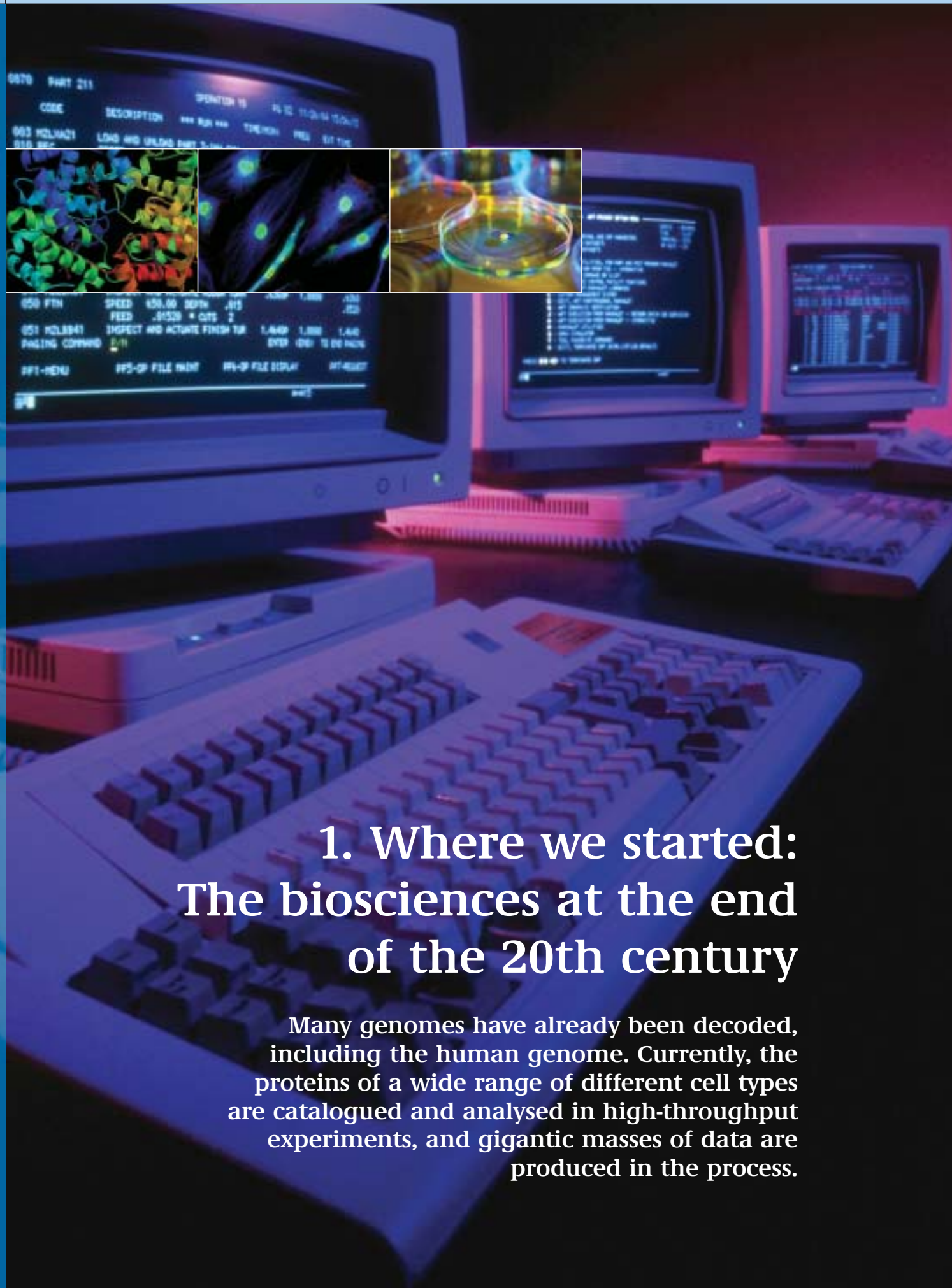
This simple analogy attempts to illustrate the situation currently prevailing in molecular biology research – with the complexity of cellular building plans and the number of components being much, much higher than in the intellectual experiment outlined above. Another weakness of this analogy is the fact that at present the "parts lists" of most living systems are still relatively incomplete. Even though the genome sequencing projects produce an immense wealth of data, information on the individual components of the genome, i.e. genes, is still fragmentary. Neither their number nor their functions are precisely known or understood. The active carriers of cellular functions, the hundreds of thousands of proteins, have hardly been mapped yet, either. Proteomics projects are currently cataloguing these proteins, and this exercise helps gradually to shed light on their interaction in cellular processes – a crucial prerequisite for 'reconstructing' biological building plans. There is no doubt that, thanks to modern high-throughput processes, the required amount of high-quality biological data will be available in the future.

But as the above example indicated, these data alone – as indispensable as they are – will not necessarily lead to an understanding of the function of a biological system, say, a cell. Many questions remain mostly unanswered, such as: How do various biological sub-systems, e.g. signal conduction pathways, interact with each other? How do cellular systems respond to external perturbations? How are complex signals recognised and interpreted?

To tackle these questions the expertise of many different research disciplines needs to be pooled, concepts borrowed from biology, computer science and systems studies need to be combined to form the discipline of 'systems biology'. In view of the improving data situation and the tremendous increase in computer performance a vision is now becoming a distinct possibility that only a few years ago would have been dismissed as an illusion: The simulation of entire living systems on a computer!

Such models could make a major contribution to understanding living systems. Modelling and simulating biological processes is of great economic interest as well. The *in silico* design of experiments with virtual cells promises, among other things, to reduce the time and costs needed to develop medical drugs, and perhaps also the number of animal experiments. The obvious potential of systems biology has already spawned several research initiatives at the international level, and also in Germany.

This brochure is intended to familiarise the reader a little with the ideas, the fantastic prospects and the realistic opportunities of this challenge.

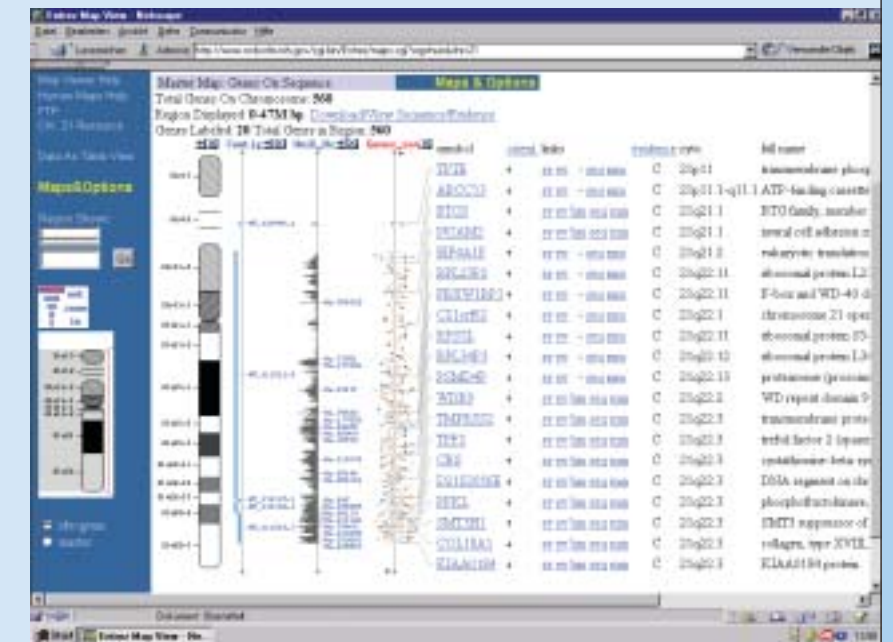


1. Where we started: The biosciences at the end of the 20th century

Many genomes have already been decoded, including the human genome. Currently, the proteins of a wide range of different cell types are catalogued and analysed in high-throughput experiments, and gigantic masses of data are produced in the process.

Over the past three decades the biosciences have been extremely successful in identifying and analysing the individual components of biological systems. As a result, a host of information on organs, tissues, cells and cell compartments down to the level of molecules was produced. This development culminated, for the present, in the decoding of the human genome and the publication of the results on 15 February 2001 (Venter et al. 2001), (International Human Genome Sequencing Consortium, 2001). A surprising result of the human genome sequencing exercise was the low number of genes which is estimated to total 30,000 to 40,000. This means that only a very small proportion (1.1 per cent to 1.4 per cent) of the entire genome sequence codes for proteins. Compared with other model organisms such as yeast (6,000 genes), fruit fly (13,000 genes), nematode (18,000 genes) and arabidopsis (26,000 genes), the number of human genes appears to be extraordinarily small, especially in view of the high complexity of the human organism.

The human genome was sequenced under two parallel large-scale projects: Celera, a private-sector US company, and the publicly funded international Human Genome Project (HGP), consisting of twenty centres in six countries, simultaneously succeeded in presenting the public with the first draft of the genome sequence. Groups of researchers from Germany also made a crucial contribution towards the success of this mammoth biology



Gene map of human chromosome 21

project. Supported by groups from France, Switzerland and the USA, German and Japanese scientists were able to decode the DNA sequence of human chromosome 21 (Hattori et al. 2000).

Internationally, the term 'genomics' was established to denote the analysis of the entire genome of an organism. The next step to follow the decoding of genetic information (DNA) is the analysis of the gene products, the ribonucleic acids (RNAs) and proteins. Accordingly, these activities are subsumed under the term 'functional genomics' which may be subdivided into three major areas: The analysis of genes that are active in a biological system (transcriptomics), the study of proteins and their interactions (proteomics), and the decoding of the three-dimensional structure of proteins which are coded by a genome (structural genomics).

Given the high level of complexity of the human organism, the number of human genes appears to be extraordinarily small

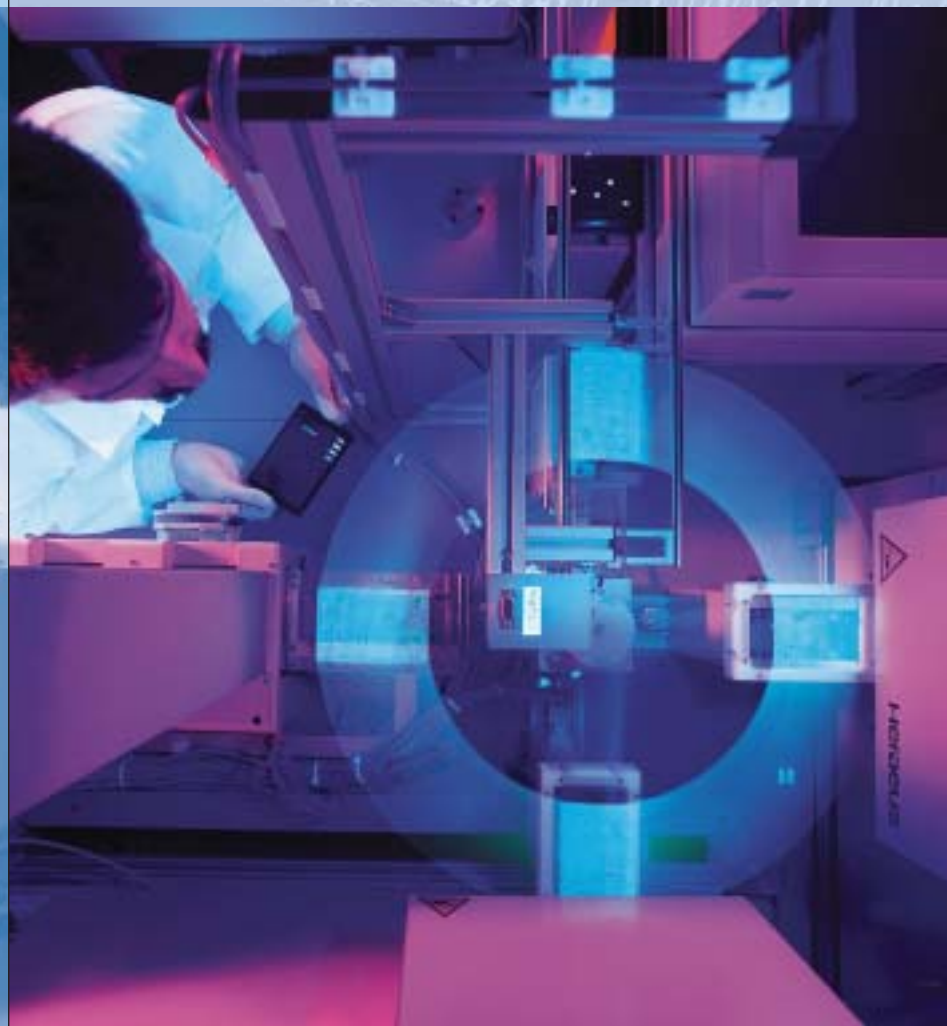


HIGH-THROUGHPUT TECHNOLOGY IN GENOME RESEARCH

The prerequisite for all 'omics approaches' is the availability of quantitative high-throughput methods and equipment. About 15 years ago automatic DNA sequencing equipment for these purposes was presented for the first time; in the meantime it has been developed into high-performance machines sequencing about 500,000 base pairs per day. The international genome research centres of the Genome Sequencing Consortium that are equipped with these sequencing machines thus achieve a sequencing capacity of 172 million bases per day or 2,000 bases per second – a throughput which was utterly inconceivable only a few years ago.

The methods of functional genome analysis are not yet as far advanced as those of DNA sequencing. The DNA array technology for transcriptome analysis is

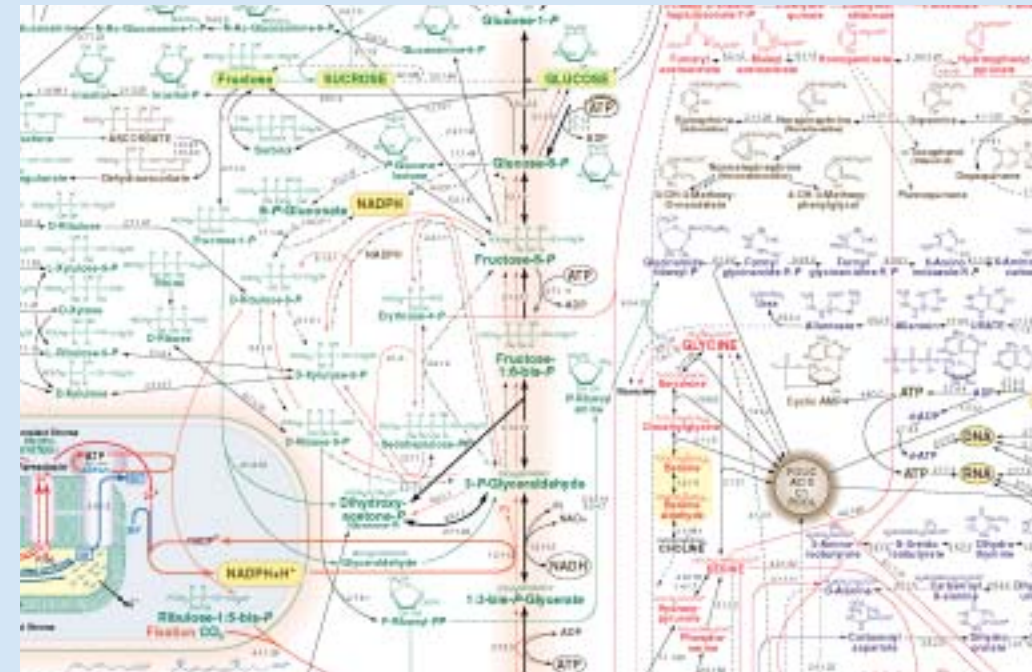
now approaching the production stage, while the methods for protein analysis still require considerable development efforts. With the best equipment available today it is possible to determine the expression levels of 1,000 proteins per day (Ideker et al. 2001a; Palsson 2000). The progress to be expected in protein analysis is reflected in the results published in January 2002 by a Heidelberg team of researchers (Gavin et al. 2002) who go far beyond cataloguing a protein inventory. The scientists succeeded in elucidating the numerous interactions of the protein complexes in yeast cells (in addition to discovering 231 new proteins). Thus a functional network of cellular 'tools' was recorded for the first time.



The latest sequencing robots - indispensable for high-throughput experiments



A map of yeast protein interactions



A network of metabolic pathways (excerpt)

All 'omics approaches' share a reductionist view. The overall biological system is divided into smaller detailed sub-systems whose individual components are then described as accurately as possible. In this way a vast amount of data on individual components or functions of a cell or an entire organism is amassed: information on all genes, proteins and metabolic processes. All these data are then compiled in gigantic gene maps, protein databases and complex metabolic networks.

Handling this host of data offers great potential. The challenge is to develop – based on the knowledge of all molecular components of a biological system - an understanding of an entire system, i.e. the principle underlying the function of a complete organism or an individual cell. How can the immense wealth of information be made accessible and translated into useful knowledge? Will we be able in the future to predict how a yeast cell will behave when a particular substance is added to the culture medium? Will we know what the consequences are when a certain gene in a mammalian cell is turned off? Are we sure that a new drug will really develop its medical

effect via a particular receptor? So far we have not been able to provide convincing answers to these and similar questions.

Especially when it comes to describing and understanding the properties of complex systems such as regulation, control, adaptation, robustness, redundancy or evolution, precise knowledge of the molecular characteristics of individual components is not enough. Rather, it is necessary to collect information on interactions between individual components of the system and between different levels of complexity as well as on the dynamics of these interactions. This means that the previously predominating reductionist approach needs to be complemented by a holistic view and an integrative method. ■

The challenge is to develop an understanding of an entire biological system, based on the knowledge of all molecular components of this system



2. Systems biology - attempting to understand the complexity of life

Complexity is a characteristic of living things. It cannot be captured only by cataloguing the components. The reductionist view needs to be extended to include systems-level approaches. Using realistic *in silico* models it may soon be possible to simulate many vital functions, perhaps even entire cells.

Complex systems exist everywhere in nature. But increasingly they are also created by humanity in the transport sector, in industry and in the information and communications sectors. Cases in point are the climate and its changes, management and control systems for complex production processes (e.g. in chemical plants) or for modern wide-bodied aircraft.

In spite of the epoch-making successes of genome research it will remain a long-term challenge to explore and understand the high level of complexity of biological processes. Every form of life is based on protein, fat, carbohydrates and on genetic information (DNA) which has a determining and controlling function. These macromolecules act as structural elements, transport systems, biocatalysts, energy converters, regulatory mechanisms, information systems etc. The constant flow of their production and degradation (i.e. their turnover) – man, for instance, produces about 60 kg per day of adenosine triphosphate (ATP), an important energy source – and their dynamic cross-linkage which results in a wide range of functional elements form the basis of the 'systems of life'.

The previous approach of biology which concentrated on molecular detail has to give way to a systems-level view. This requires approaches of the mathematical and physical sciences to be integrated into biology.

Now, what exactly is systems biology? Systems biology investigates the behaviour and relationships of all the elements in a particular biological system while it is functioning (Ideker et al. 2001a; Palsson 2000). Systems biology thus wants to help understand biological processes at the level of systems. To this end, it is necessary to take a closer look at the following four key characteristics of biological systems (Kitano 2000):



Modern wide-bodied aircraft are complex technical systems consisting of millions of components

- (1) **Systems structures.** They include the individual components of biological systems and their structural relationships.
- (2) **Systems dynamics.** It focuses on the question of how a system behaves under different external and internal conditions over time.
- (3) **Systems control.** The mechanisms used within a biological system for its control need to be studied.
- (4) The **design principles** of biological systems have to be identified and used for systems analysis.

Now, what does a systems approach in biology look like in concrete terms? What does systems biology encompass? Based on data on biological systems or sub-systems and the knowledge of the function of, and relationships within, the biological system of interest, a mathematical model of this system will be developed which allows hypotheses about the characteristics and behaviour of the system to be derived. These hypotheses can then be verified in parallel *in silico* and *in vitro* experiments. Based on the mathematical model, the *in silico* simulation will produce new data that will permit verification of the hypothesis. Parallel *in vitro* experiments will generate new biological data which will also help to validate the hypotheses developed.

"Systems biology investigates the behaviour and relationships of all the elements in a particular biological system while it is functioning"



A large part of the current knowledge of biological processes is quantitative in nature

This iterative process generates new biological knowledge. The *in silico* and *in vitro* loops shown in the chart can be performed not only in parallel, but also sequentially. In this case *in silico* simulations would be used as a first step

to identify the experimental strategies which would be best suited to solve the problem in hand. In this way the experimental design can be optimised so that at the next stage only the really crucial experiments will have to be conducted.

Data capture

Systems biology research begins with **the capture of data**. A major part of the current knowledge of biological processes consists of qualitative data. This includes, say, the description of the function of biological molecules, the representation of interactive metabo-

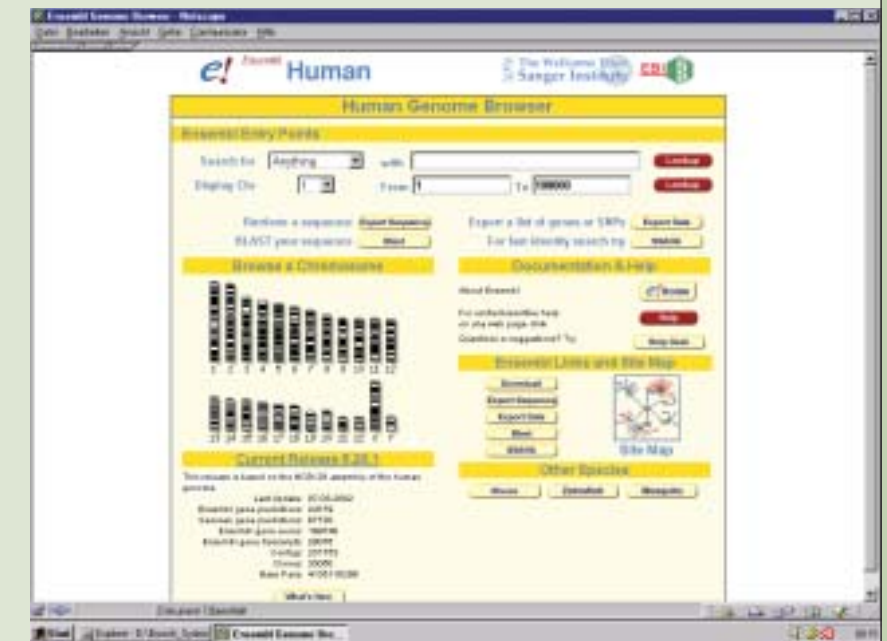
lic pathways, the identification of molecular interactions, the description of the individual development of an organism or findings regarding the evolution of certain species. But the modelling and simulation of biological phenomena which are at the very centre of systems biology require quantitative data at all hierarchical levels of the biological system of interest. This means that quantitative information on DNA, mRNA, proteins, protein interactions, other biological molecules, information pathways, information networks, cells or tissues, organisms or even entire populations is needed. This information includes the concentration and absolute number of molecules present in cell compartments, cells or tissues, material flows, the rate constants of biological reactions, affinity or dissociation constants for interactions. These data should not only represent a snapshot of a certain state of an individual biological system, but also take into account variations between individuals of identical genetic constitution, dependencies on different physiological states and developments over time.

A crucial aspect of data capture is **data quality**. Experimental data have to be amenable to modelling. This is why data generation and modelling have to be designed as parallel and interactive processes. This helps to avoid generating a host of data which may eventually turn out to be unsuited for mathematical modelling. Also, it is necessary to ensure the comparability of biological data that were generated by different working groups at different levels of complexity for different states of systems. For this reason **standards** for data generation are needed which cover both the methodology used and the data quality desired.

To ensure that the data captured can be used for systems biology research it is necessary to organise the filing of these data in suitable

databases and provide access to such facilities. As a result of the genome sequencing activities, numerous databases have been developed in the last few years which contain both genome sequences and the amino acid sequences of proteins. Of primary interest for systems biology, however, are more recent database generations which contain, for instance, interactions between proteins or proteins and DNA, metabolic pathways or information on gene expression. Quantitative data in particular will require the development of additional database formats.

Standards for data generation are needed which cover both the methodology used and the data quality desired.



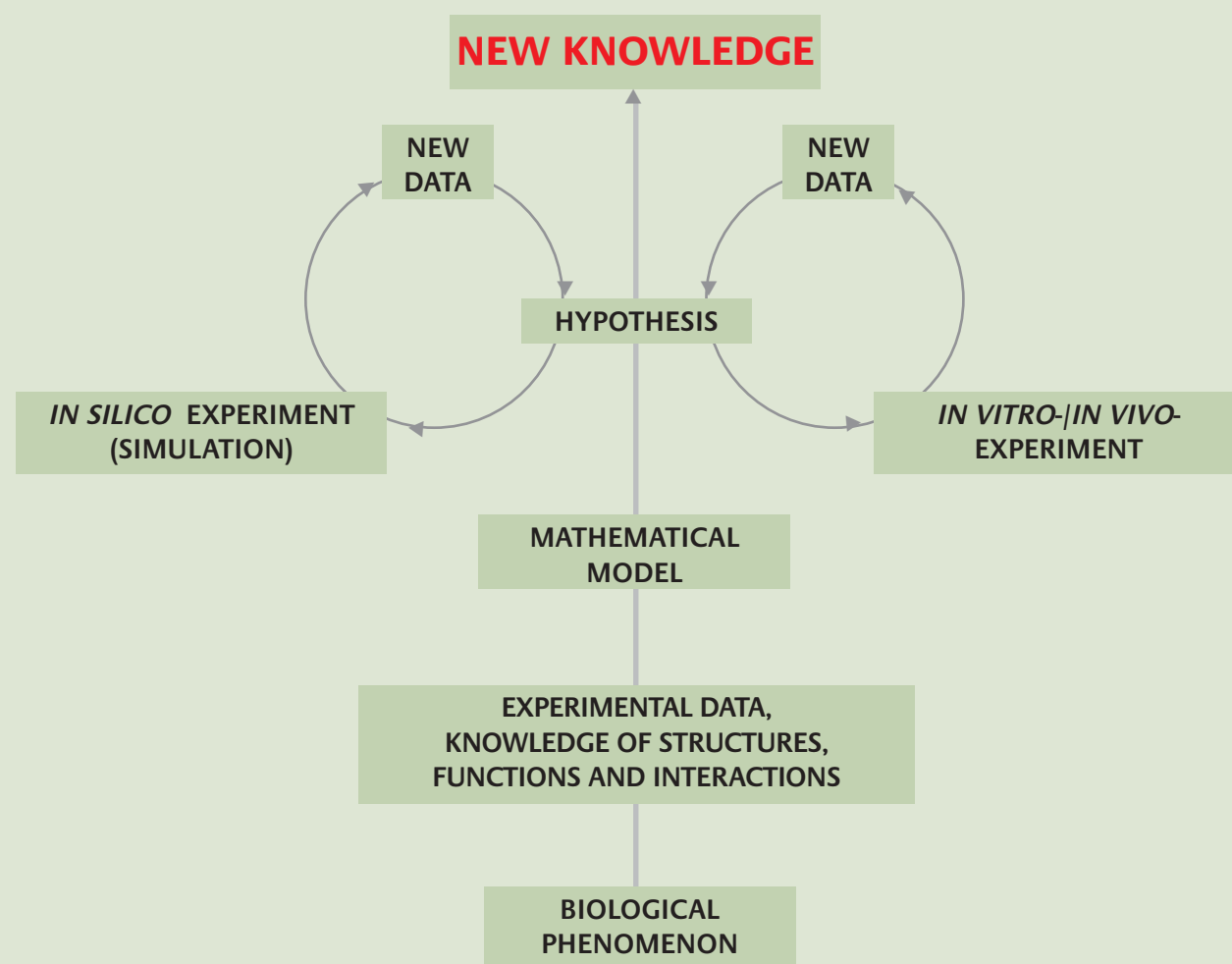
Standardisation will also play a key role in data filing. Binding rules are needed for both the way in which data are collected and the volume of data filed.

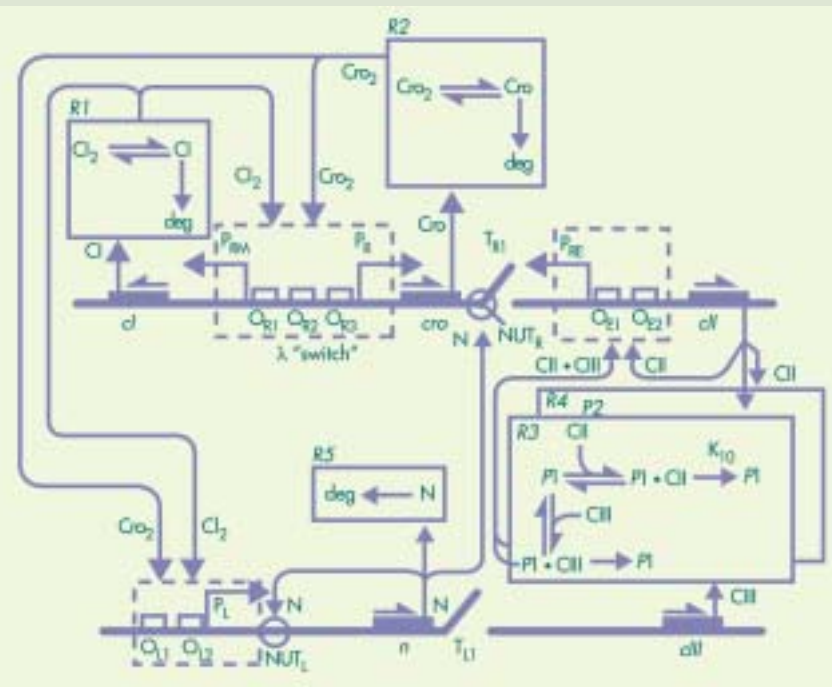
Clearly structured: human genome data on the web

Modelling

In the biosciences, a model is the attempt to provide an abstract representation of the information gained from experimental observations on the structure and function of a particular biological functional element. This functional element may assume very different levels of complexity and range from simple enzyme reactions to specific sections of a

The process of knowledge generation in systems biology





Kinetic model of a 'circuit' of the lambda phage

with an abundance of data and a level of complexity that can no longer be modelled with the help of simple graphics alone and without any additional tools. In systems biology, **computer-aided modelling** is indispensable. Computers serve to store, systematise and compile the data available. Furthermore, automatic tools should be provided which combine data to form network models and allow simulations, thus predicting network behaviour. Such computer models are standard practice in the engineering and physical sciences, but only now taking off in the biosciences. Well-known examples are the so-called E-Cell and Virtual Cell projects which attempt to model all molecular interactions within a cell as an integrated computational process.

Computer models

Computer models of biological systems can be divided into two groups: **chemical kinetic models** and **discrete circuit models**. Chemical kinetic models attempt to represent cellular processes as a system of chemical equations, with the concentrations of molecules of interest and their interactions in chemical reactions determining the instantaneous state of the system. Chemical reactions can be represented mathematically as differential equations where the changes in the concentrations of reactants and post-reaction products are recorded based on the reaction rate. Normally such a system of differential equations is too complex to be solved explicitly. But given an initial network state it is possible to determine the quantity of the various molecular components and thus model – in a step-by-step approach - different states adopted by the system over time.

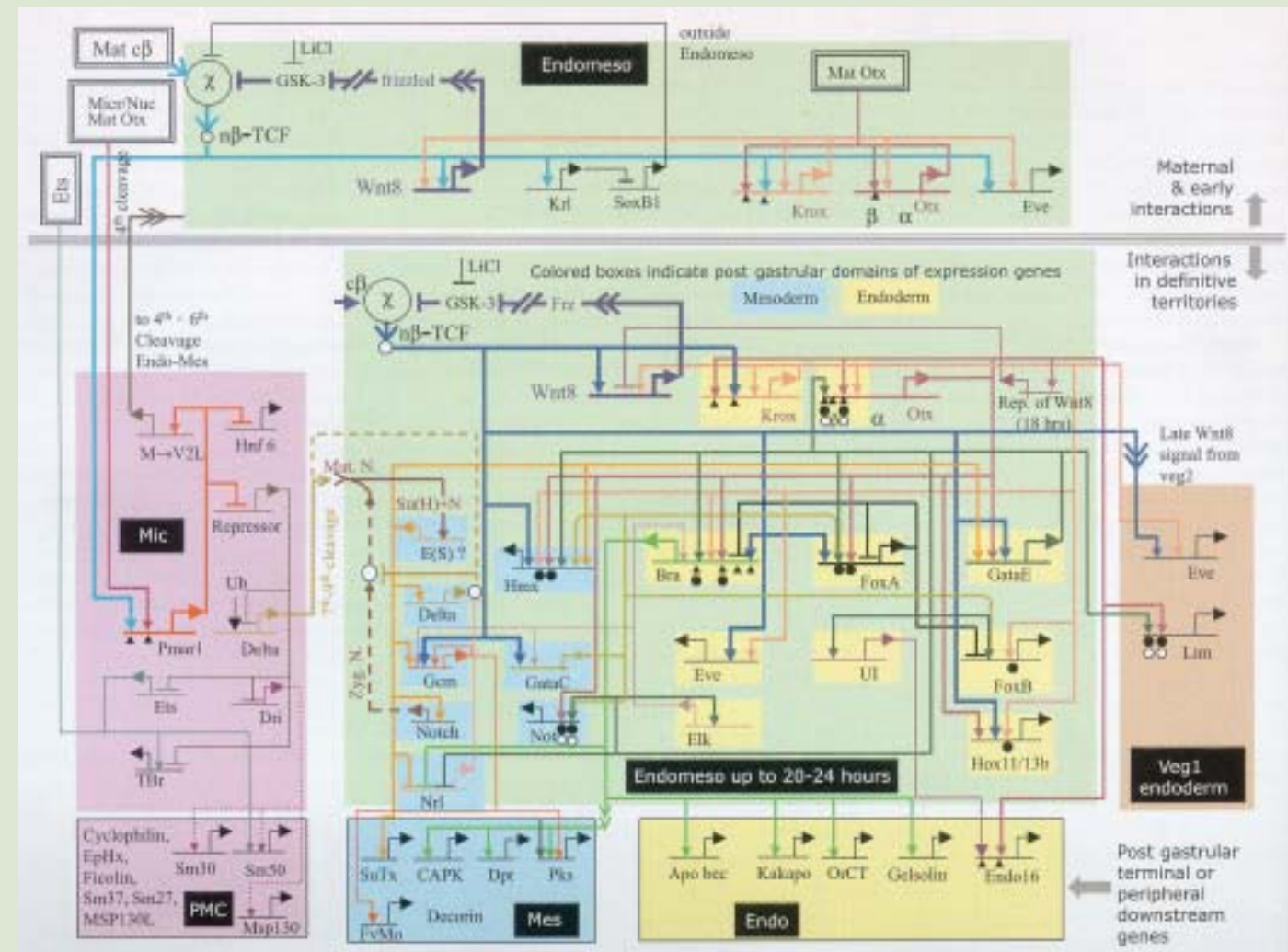
Not all biological processes behave in a deterministic fashion; this is why it is necessary to model a number of central biological processes such as transcription or translation as ran-

metabolic pathway, to signal transduction pathways, cell division, more complex tissue or organ functions to whole organisms. A biological model often offers a simple graphic representation of the postulated relationships.

A model has three important functions. First, it is a tool which helps better to understand the biological phenomena studied. Second, it enables experiments to be specifically designed by permitting predictions of certain characteristics of the functional system represented by the model, which can then be experimentally verified. Third, a biological model summarises the current body of knowledge in a format that can be easily communicated. A model thus provides the possibility to codify biological knowledge and hence is an important prerequisite for communicating the gain in scientific knowledge.

Systems biology aims to capture all components and interactions of a given functional system and their underlying dynamics. Consequently, systems biology is confronted

Systems biology aims to capture all components and interactions of a given functional system and their underlying dynamics.



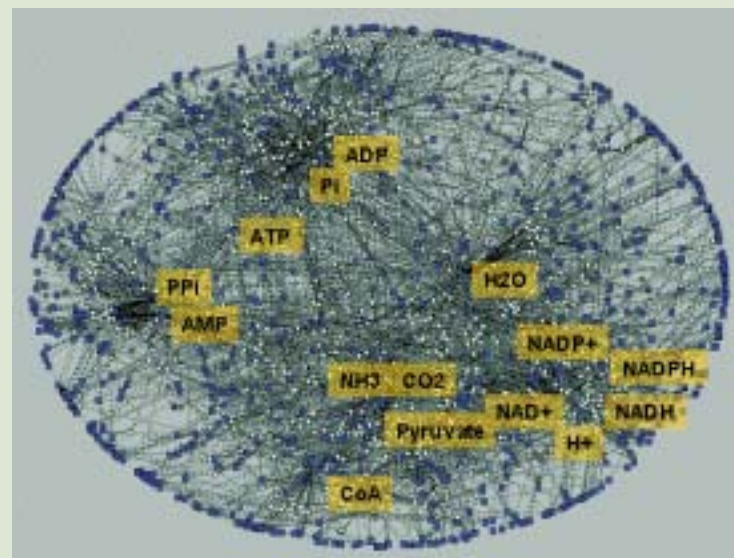
Gene regulation network of sea urchin cells

dom events. Such models are based on stochastic relations rather than differential equation systems.

The second fundamental modelling strategy does not represent processes in a continuous dynamic fashion; rather, it models discrete events as feedback loops. The representation of the model forms a network consisting of

nodes and directed edges between these nodes. In this network nodes stand for the quantity or level of a certain molecule, while the relationships between the nodes describe the effect of the level of a given node on neighbouring nodes. For each node a function is required which describes how all the influences acting on the node of interest determine its level. In a simplified model, nodes may assume one of two discrete states, indicating the presence or absence of a certain molecule or whether a gene is turned on or off. At the beginning of the modelling process there are given starting states of levels for

The representation of the circuit model forms a network consisting of nodes and directed edges between these nodes.



Prediction of all metabolic reactions expected for yeast

all nodes. Subsequent levels can be directly determined from the respective node functions. In this way the network state over all nodes evolves over a series of discrete time steps, where the state of the next step is computed from the current state. Discrete circuit models greatly simplify the real situation. Especially the assumption that all node states change simultaneously is hardly realistic, just as real biological systems can often assume more than two possible states.

not only at the transcription stage, but also through post-transcriptional mechanisms like alternative splicing or translation control. And regulation through protein modifications has to be considered as well. In addition to these different levels of complexity, the spatial compartmentalisation of biological processes within the cell plays a central role in the modelling approach. It is obviously always desirable to develop as comprehensive a model as possible, but pragmatism imposes limitations, including limitations of the parameters mentioned. Consequently, the first step consists in identifying those characteristics which are to be described or simulated in a model, and then all components which might presumably influence these characteristics are included in the model as well. Another constraint is that the number of model parameters should be compatible with the quantity and quality of the experimental data available.

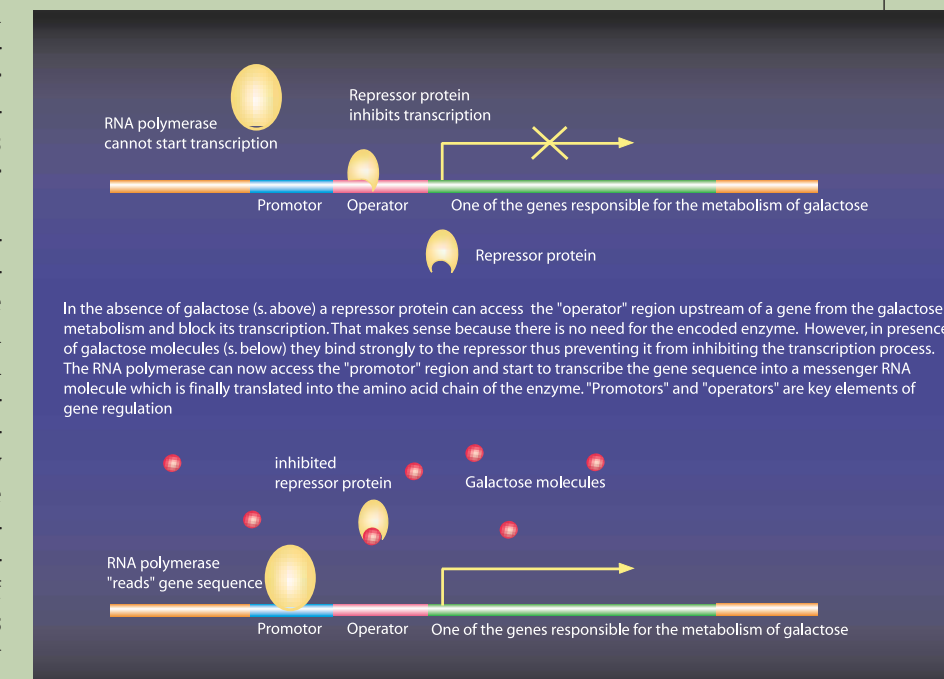
A computational model of a biological unit is not the final product of systems biology, but rather a tool used in the **holistic approach of systems biology research**. This approach is characterised by a close interaction between *in silico* modelling and simulation and *in vitro* or *in vivo* experiments. The aim of the systems biology approach is eventually to develop a model which represents the real situation as completely and correctly as possible (see examples in the box).

A computational model of a biological unit is not the final product of systems biology, but rather a tool used in the holistic approach of systems biology research

Not only the selection of the modelling strategy, but also the definition of the level of detail of the model is a major challenge. Say, a decision has to be taken as to which genes should be studied, which proteins should be included, which other molecules should be considered as well and which relationships between the individual components should be included in the modelling exercise. In the case of gene regulation, for instance, it has to be borne in mind that regulation can be achieved

THE RESEARCH APPROACH OF SYSTEMS BIOLOGY

The metabolic pathway of galactose in yeast is a classic example of a regulatory switch where enzymes needed for the transport and catabolism of galactose are only expressed if galactose is present in the medium and other sugar molecules such as glucose are absent. The genes, gene products and metabolic substrates required for the metabolic pathway are now known, as are the interactions between the individual components that lead to the activation or repression of this metabolic pathway. It was possible to develop a computational model and check its validity in experiments where systems were systematically perturbed. In this exercise gene expression and protein abundance were quantified by means of parallel microassays and proteomics methods, and databases containing known physical interactions were analysed. As a result almost 1,000 different mRNA molecules could be identified which react to 20 systematic perturbation conditions of this metabolic pathway. It could be demonstrated that about 15 out of 289 identified proteins were regulated by posttranscriptional mechanisms. By refining the model it was possible to explain a host of previous observations and gain a number of unexpected additional findings regarding the regulation of the galactose utilisation pathway.



Part of the galactose metabolism: regulation of lactose degradation



An example of signal transduction (Schoeberl et al. 2002)

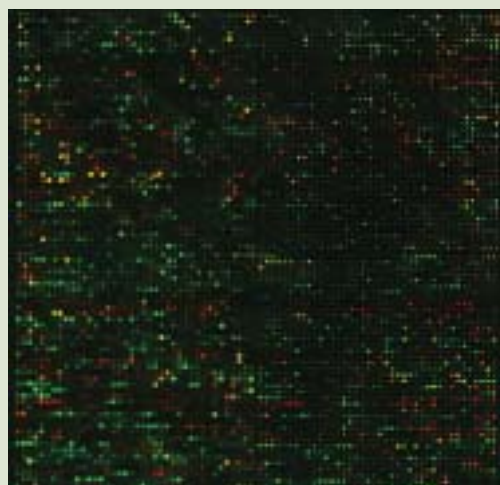
The activation of the MAP kinase cascade by the epidermal growth factor (EGF) constitutes a central signal transduction pathway in mammals. EGF binds to an extracellular receptor which activates an entire network of different proteins and/or induces intracellular shifts. Recently scientists have successfully used a computational model to simulate this complex network which consists of 94 individual components. Almost 100 differential equations and a similar number of parameters were used for modelling. Based on this model a number of physiological reactions could be predicted which agreed closely with the experimental data. For example, it was inferred from this systems biology-based analysis that the initial rate of receptor activation determined the physiological reaction of the signal chain.



The basis of systems biology

The interplay of different scientific disciplines is imperative for the holistic approach of systems biology

The **interplay of different scientific disciplines** is imperative for the holistic approach of systems biology. Close interaction between the generation of experimental data and theory-based computational modelling and simulation is required in any case. Systems biology centres on mathematical models of biological units. These models, however, can only be developed and implemented if biology, mathematics, computer science and systems studies co-operate on an interdisciplinary basis.

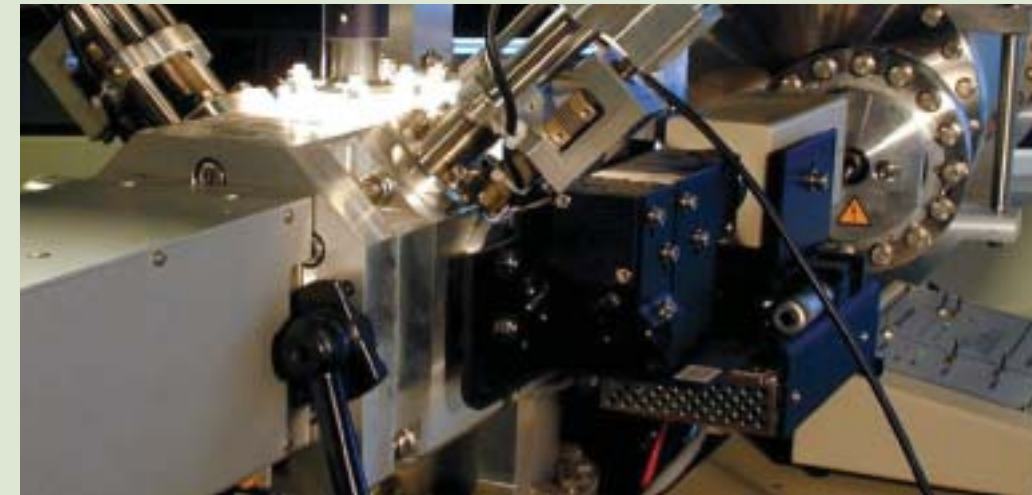


DNA chips have revolutionised molecular genetic research

Biology provides the objects to be studied as well as preliminary concepts of structural and functional relationships. It offers methods such as genetic modification which are valuable tools for studying these systems with a view to verifying models. Finally, it provides the quantitative analysis approach which is of key importance for data generation. The contribution to be made by systems studies consists of providing the modelling concept which may be used to study the stability of a system, discovering the principles underlying regulation and analysing and/or influencing the controllability of the system. Computer science and mathematics finally contribute the necessary modelling tools. They integrate a wide range of different biological data, store them in databases and make them available for modelling. Another central task is the development of user-friendly software that can be handled and understood by both biologists and systems engineers. Finally, the complex relationships and simulation results have to be visualised in a comprehensible fashion.

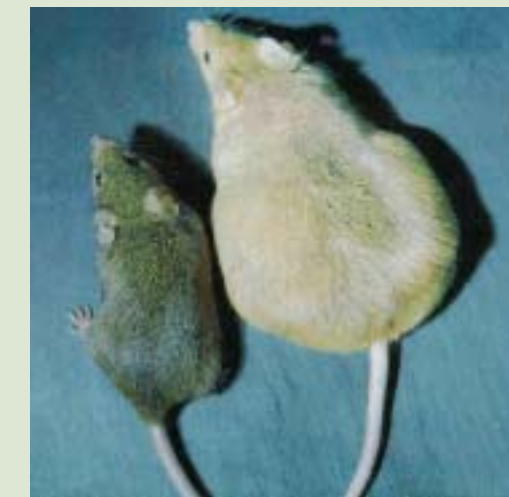
The research approach of systems biology requires appropriate **methods**: first, for characterising the biological systems of interest, second, for systematically modifying the system, and third, for system modelling. Tools for characterising the system will permit a quantitative high-throughput analysis of systems components to be performed with the highest accuracy possible. This requires measurements both of individual molecules and of molecular interactions. In addition, systems characterisation should be performed as fast as possible and in real time. In some cases non-invasive methods are required to keep the overall system intact.

The range of methods needed should cover all levels of complexity of the functional system of interest. This implies that methods are required at the genome, transcriptome, proteome, metabolome and physiome levels. In the course of the worldwide genome sequencing activities highly efficient sequencing methods were developed at the genome level. There are different variants of DNA arrays that may



be used for transcriptome analyses. The challenge for proteome analysis is to identify proteins and gain information about their abundance, processing, possible chemical modifications, interactions with other molecules, their intracellular compartmentalisation and structure as well as their synthesis and degradation rates (turnover time). The immense range of protein abundances present in living organisms - from one molecule per cell to more than one million molecules - illustrates the methodological challenges to be faced in this field. It is also necessary to develop quantitative methods for high-throughput measurements of cellular metabolites.

Genetic methods seem to be best suited for the **systematic perturbation of systems**. Basically, this means that certain genes are deleted, inserted or modified. For instance, it is possible to induce a loss or change of function or add new functions. For these purposes genetic engineering provides a wide range of methods. Mutations may occur spontaneously or they can be induced by radiation. The generation of transgenic organisms offers numerous potentials, and so do gene targeting, gene tapping or chemically induced mutagenesis. For systems biology it will be necessary to optimise these and other



methods in order to allow fast and efficient mutagenesis.

The challenges for computer science and mathematics consist first and foremost in the development of robust algorithms for quantitative analysis. These include methods for multi-dimensional image processing, parameter estimates, statistical methods for model evaluation and the handling and processing of complex data. Second, it is necessary to develop mathematical models and simulations, and finally IT has to accomplish the essential task of organising data. What is needed in particular are standardised processes and methods for the generation, analysis and filing of data, for data access and modelling.

Transgenic organisms lend themselves to the systematic study of the impact of genetic defects and hence of gene functions



Comparison of two proteomes

The benefits of systems biology

It is already clear today that both basic research and the application of biological findings in medicine or plant breeding will benefit considerably from systems biology. This is illustrated by the following selection of examples:

1. Research in the biosciences

The use of mathematical models of biological units will make it easier to provide information on a given biological unit that has been compiled in the literature and in various databases, in a structured and integrated format; this will help to advance understanding and permit the immediate use of these data. Such models can also help teachers to communicate and illustrate complex facts.



under varying conditions. This is why limitation to a selection of possible combinations is inevitable. Mathematical models permit the simulation of those combinations that have not been tested experimentally, and/or provide a more rational basis for the selection of those combinations that appear to be most significant for experimental testing. The model approach thus makes it possible to prioritise the experiments to be conducted which eventually may help to save time and resources.

Computational models can also be used to develop experimentally verifiable hypotheses. Using the mathematical model to go through various options that are all possible in principle and comparing the results with the measured values also contributes to verifying and optimising hypotheses.

Basically, systems biology represents a new

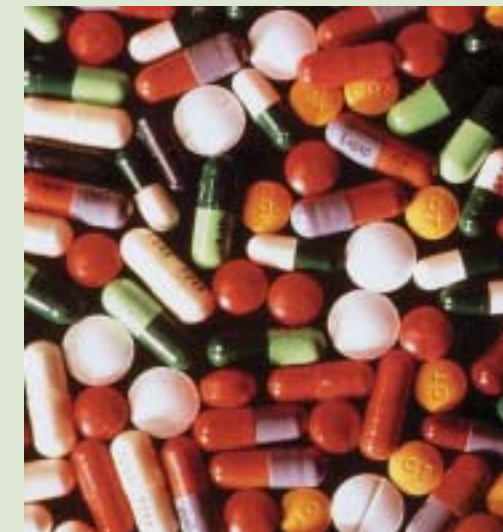
concept which encompasses the further conceptual development of biology into a predictive, quantitative, theory-based science aiming to achieve a holistic understanding of biological systems. Systems biology thus introduces guiding principles to structure the deluge of data to which the biosciences have been exposed as a result of the high-throughput research activities of the last few years.

A model-based approach can also be used to conduct biological experiments. Mathematical models cannot replace real experiments; but they can contribute towards better experimental design and planning and a more efficient and targeted use of resources. For example, it takes a major effort to generate a large number of different mutants and test them

A model-based approach can also be used to conduct biological experiments

2. Medical applications

When new drugs are discovered, computational models can contribute to identifying those target proteins ('targets') for specific metabolic pathways or signal-transduction chains whose repression promises the greatest possible therapeutic effect. Based on this information, specific inhibitors of these proteins can be developed as potential new therapeutic



agents. Such a model-based approach to discovering active principles is particularly promising if it is not a single molecule that can be identified as the active site, but if the therapeutic effect is determined by the behaviour of a network of different target proteins. Modelling the network behaviour allows sites for therapeutic intervention to be identified so that suitable drugs or combinations of drugs can be developed. Finally, modelling also permits a quantitative analysis of the effects of different drug concentrations on the therapeutic result.

As well as discover active substances, systems biology can also make essential contributions towards the process of drug development and in particular towards the performance of pharmacological and toxicological tests. A suitable cell model, for example, might be used to simulate extensive toxicological test series,

thus reducing the number of animal experiments required. Finally, it is also conceivable that a systems biology approach is taken to select the optimal therapeutic strategy. For example, a computational model of cell regulation could be used to identify those genes or proteins in a particular tumour whose repression promises to provide the best control of tumour proliferation.

A computational model of cell regulation could be used to identify those genes or proteins in a particular tumour whose repression promises to provide the best control of tumour proliferation

3. Plant breeding

In principle the systems biology approach also offers numerous possible applications in plant breeding. One of the major challenges in breeding is to identify those genetic markers which influence the desired performance characteristics of a crop. A mathematical model describing the plant characteristics of interest and their dependence on molecular components and interactions of the plant system studied can make it much easier to identify the genetic parameters that are relevant for breeding. ■





3. International activities in systems biology

With the triumphant advance of computer technology and the success of high-throughput processes the modelling of biological systems has more and more become a distinct possibility. This is why in recent years the major industrialised countries launched several research projects and research initiatives regarding systems biology.

Early isolated approaches to systems biology date back to the 1960s when the idea of applying the concepts of systems theory to biology was promoted. But it is only in the age of modern biosciences that has just begun to dawn that systems biology has been internationally recognised as a promising new area of research. Systems biology research groups are establishing themselves in Japan, Canada and the USA, and also in some European countries. In some cases the first publicly funded research initiatives have already been launched. At least in Europe the German 'Systems of Life – Systems Biology' research and funding programme is playing a pioneering role. In no other European country has there been a comparable initiative with similarly ambitious objectives.

The International Conference on Systems Biology (ICSB) has become established as an important forum for systems biology. The first ICSB took place in Tokyo in the autumn of 2000 and was attended by approximately 250 scientists. Only 350 participants could be admitted to the second ICSB meeting in Pasadena in November 2001 and nearly 500 scientists gathered at the ICSB2002 in Stockholm. The interested audience will certainly exceed the capacities of the ICSB 2003 in St. Louis and in 2004 when the annual international conference will be held in Heidelberg, Germany.

Systems biology in other European countries

At the *Université Libre de Bruxelles* in **Belgium** several working groups are looking into issues of systems biology. For example, they are developing an object-oriented modular database which links metabolic pathways, protein-protein interactions, regulatory information and signal conduction details. Other activities focus on modelling gene regulation processes.

Research conducted at the *Centre for Biological Sequence Analyses* of the *Technical University of Denmark* in Copenhagen centres on the complex relationships between gene sequence information and the structure and function of macromolecules. The Centre also develops new computational methods for classifying, predicting and comparing biological networks by concentrating on so-called adaptive data-driven bioinformatics methods that rely heavily on learning structures. The Centre is funded by the *Danish National Research Foundation*.

The International Conference on Systems Biology (ICSB) has become established as an important forum for systems biology



Heidelberg will host the ICSB in 2004

In its 2001 – 2005 five-year research plan the Japanese Council of Science and Technology Policy identified systems biology as one of the priority themes

France has recently launched the *In silico Cell Simulation Initiative* which intends to develop bioinformatics methods for modelling signal transduction and regulatory networks within the cell. The SiliCell Initiative is jointly funded by several CNRS laboratories in Montpellier and Villejuis.

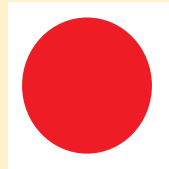
Analysing and modelling metabolic networks is one of the priorities of systems biology research in the **Netherlands**. Working groups at the *Amsterdam Biocentre* and at the *Technical University of Delft* are working in this field. The *Department for Information Technology and Systems* of the Technical University of Delft conducts bioinformatics activities related to the analysis of genetic networks.

Several working groups in **Sweden** are currently shifting the focus of their activities to systems biology. The working group on *Complex Systems* of the *Department of Theoretical Physics* at *Lund University*, for example, is developing models of regulatory genetic and metabolic networks. Other important players in this field are working in the *Department for Cell and Molecular Biology of the University of Uppsala*, at the *Microbial and Tumour Biology Centre of the Karolinska Institute* in Stockholm, at the *Stockholm Bioinformatics Centre* and at the *Intelligent Systems Laboratory* of the *Swedish Institute for Computer Science*. In Sweden, systems biology research is funded by the *Swedish Research Council* and the *Swedish Foundation for Strategic Research*, among others. That the *Third International Conference on Systems Biology* was held in Stockholm in December 2002 reflected Sweden's growing interest in systems biology and certainly provided another boost for this field of research.

In the **United Kingdom** mathematical models for systems biology activities are being developed at the *Control Systems Centre of UMIST in Manchester*, among others. A working group set up at *Cambridge University* looks into the computer simulation of biochemical circuits underlying bacterial chemotaxis. The *Biocomputation Group* at the *University*

of Hertfordshire specialises in the development of bioinformatics methods for modelling and analysing molecular processes; in co-operation with US researchers the Group has developed a *Systems Biology Workbench*.

Systems biology in Japan



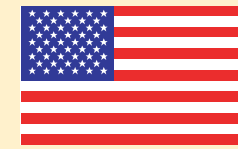
The most comprehensive initiative launched in **Japan** in the field of systems biology is the *ERATO Kitano Symbiotic Systems Project* which was started at the *Sony Computer Science Laboratory* in 1998 and co-operates with another group working at *Caltech* in *Pasadena*, California. Under this project genetic, cellular and metabolic mechanisms in various organisms are analysed and simulated by means of computer models with a view to developing a holistic understanding of the biological systems of interest. The second systems biology working group was set up at the *Institute for Advanced Biosciences at Keio University*. This is where in 1996 the E-Cell Project was launched which aims eventually to represent an entire cell in a computer model. The importance of systems biology is expected to increase in Japan as well over the next few years. In its 2001 – 2005 five-year research plan the *Council of Science and Technology Policy* identified systems biology as one of the priority themes.

Systems biology in Canada and the USA



In **Canada** it is primarily the two networks of centres of excellence, the *Canadian Protein Engineering Network (PENCE)* and *Mathematics of Information Technology and Complex Systems (MIT-ACS)*, which conduct activities in the field of systems biology. PENCE consists of more than 20 institutes from all over Canada which look into different aspects of protein engineering. They also use bioinformatics tools to study protein functions and protein networks. MIT-ACS is a network of centres of excellence funded by the federal government which was set

up in 1998 and is based on the co-operation among three mathematical institutes. Their projects also address the mathematical modelling of biological phenomena.



In the **USA** several research institutions have been active for some years in research areas that are of relevance for a systems biology approach. However, only few institutions have already implemented the interdisciplinary integration and holistic approach required by systems biology. Among those that have done so are the *Alliance for Cellular Signaling (AfCS)*, a *Dallas*-based network, the *Institute for Systems Biology* in *Seattle* and the *Molecular Sciences Institute* in *Berkeley*.

The *Alliance for Cellular Signaling* was founded in September 2000 and is a co-operative venture of seven research laboratories located in the USA and the UK. It is headed by a steering committee under the chairmanship of Nobel prize winner Alfred Gilman. The AfCS receives public funding totalling about US\$ 5 million per year from the *National Institute of General Medical Sciences (NIGMS)* of the *National*

Cancer Institute (NCI) and the *National Institute of Allergy and Infectious Diseases (NIAID)*. Several pharmaceutical companies and numerous biotechnology firms contribute an additional US\$ 7 million per year. The goal of the AfCS is to decipher the interplay of cellular signal transduction pathways in precise quantitative terms and to explain how signals can be correctly identified in different settings. These issues are being studied in two mammalian cell systems – cardiac myocytes and B-lymphocytes.

The *Molecular Sciences Institute* in *Berkeley* was founded in 1996 by Sydney Brenner, one of the winners of the 2002 Nobel prize for medicine; its aim is to develop a predictive biology based on a systems-driven understanding of biological processes. To this end, the institute develops new methods for generating and analysing biological information and for computational modelling.

In the USA several research institutions have been active for some years in research areas that are of relevance for a systems biology approach



Nobel laureate Al Gilman is the driving force behind the Alliance for Cellular Signaling (AfCS)



ling genetic circuits. Together with the Japanese Kitano Project, *Caltech in Pasadena* is developing mathematical modelling tools. A working group at *MIT in Boston* has taken a systems biology-based approach to studying signal transduction in mammalian cells. The *Centre for Biomedical Imaging Technology* at the *University of Connecticut* has developed a software package for modelling cellular systems, called 'Virtual Cell'. The M-Cell Project is a super-computer simulation of the synapsis between nerve cells and muscle cells; it was developed at the *Pittsburgh Super Computing Centre* together with the *Salk Institute in San Diego*. And finally, in March 2002 the *Bauer Centre for Genomics Research* was opened at *Harvard University in Boston*; it is an interdisciplinary research centre which will also use a systems biology approach in its work.

The *International E. coli Alliance (IECA)* plans

to develop a computational model of the *E. coli* bacterium and for this work will receive US\$ 100 million over ten years.

In addition to these and other research institutions, a number of biotechnology firms was set up in the USA in recent years which focus on systems biology aspects. These include *Physiome Sciences* in *Princeton*, *Genomatica* in *San Diego*, *Tissue Informatics* in *Pittsburg*, *Entelos* in *Menlo Park* and *Integrated Genomics* in *Chicago*.

In addition to funding from NIGMS, NCI and NIAID, systems biology research in the USA has recently also received financial support from the *Department of Energy (DOE)*. The DOE has launched two complementary schemes with systems biology objectives. The *Microbial Cell Project (MCP)* is aimed at understanding the function and regulation of all genes as well as

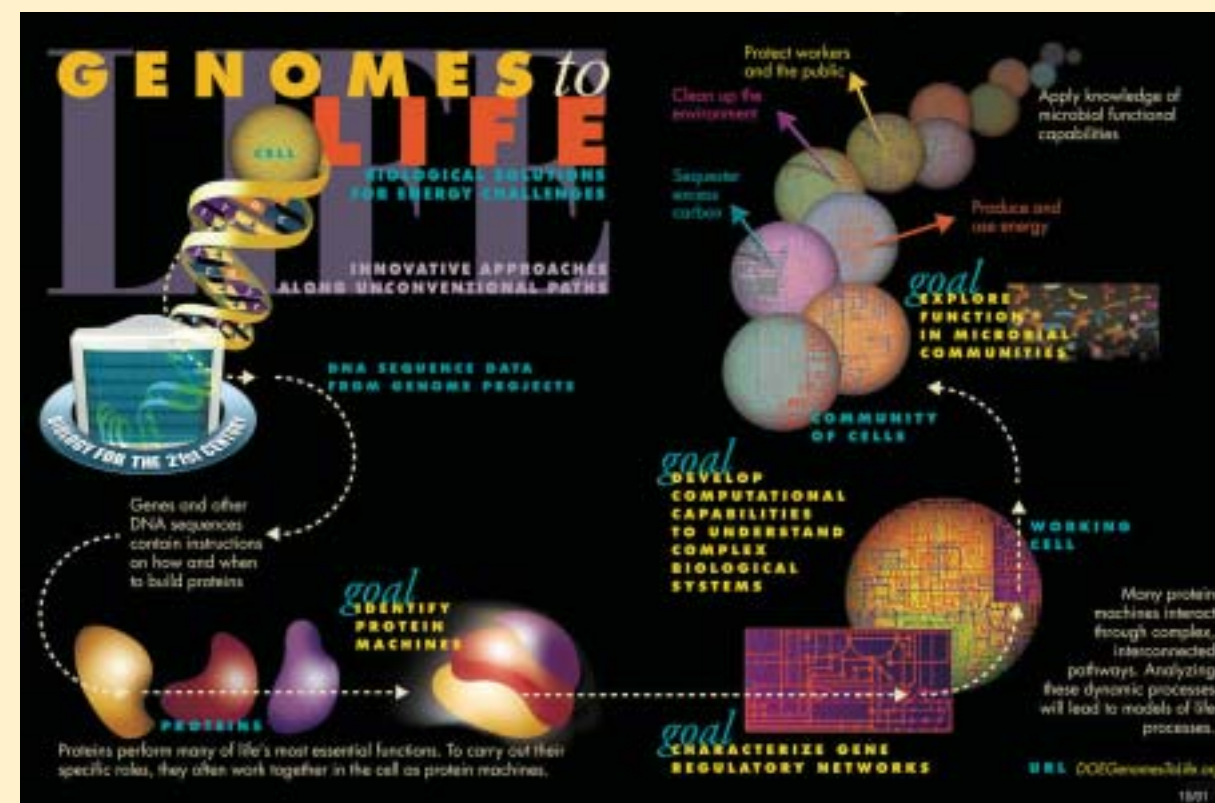
the protein networks of individual microorganisms. To date (*i.e.* early 2002) a total of US\$ 15 million has been allocated to project funding.

The *Genomes to Life Programme (GTL)* was initiated in 1999 and will begin to fund the first projects in 2002. It covers a period of ten years and has an annual budget of about US\$ 22 million. GTL is designed to contribute to understanding the composition and function of the biochemical networks and pathways underlying essential processes of life. Four goals are pursued in particular:

- ▶ Identification and characterisation of molecular machines,
- ▶ analysis of gene regulation networks,
- ▶ characterisation of molecular factors and mechanisms of complex microbial communities in their natural environment,
- ▶ development of computer science methods and skills that are required for understanding complex biological systems and predicting systems behaviour.

GTL makes strategic use of the infrastructure available in the DOE environment, *i.e.* especially the high-throughput sequencing capabilities, the super-computer facilities, synchrotron and neutron radiation sources, the resources of the Human Genome Project and the skills developed under the DOE nanotechnology initiative. ■

The International E. coli Alliance (IECA) plans to develop a computational model of the E. coli bacterium and for this work will receive US\$ 100 million over ten years



Together with the Japanese Kitano Project, Caltech in Pasadena is developing mathematical modelling tools

Numerous other research institutions in the USA are studying other aspects of systems biology. The *Department of Bioengineering* at the *University of California in San Diego*, for instance, is working on analysing and model-



4. The Systems of Life – Systems Biology research and funding programme

From the idea to a well-funded research programme – or how to bring together scientists from different disciplines to achieve a great goal: the *in silico* model of a human cell

A new funding priority emerges

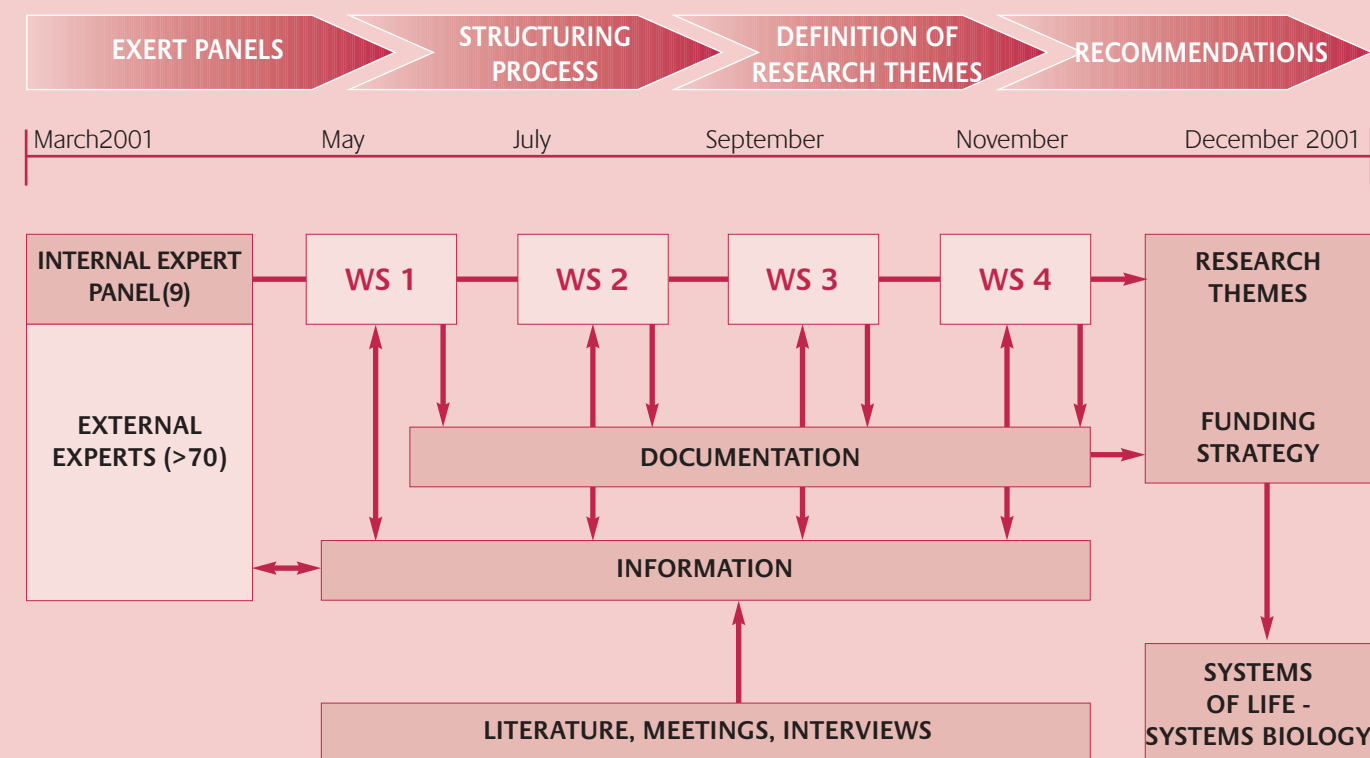
In early 2001 the Biological Research and Technologies Division of the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) initiated a forward-looking process under the name of BioFora21, with a view to developing new funding strategies to meet the future requirements of modern biosciences. The basic question was: Can we develop a cross-disciplinary research programme which is innovative and attractive and can be implemented in the medium term (i.e. in the next 5 to 10 years), which meets international standards and complements the ongoing research schemes in the biosciences that are funded by the BMBF? From March to November 2001, BioFora21 explored these ambitious goals, with the Fraunhofer Institute for Systems and Innovation Research (Fraunhofer Institut für Systemtechnik und Innovationsforschung, ISI) acting as a moderator in the process. The core element of BioFora21 was a panel comprising nine renowned experts from scientific institu-

tions and from industry who developed the new research programme in four two-day workshops.

This internal group of experts was supported by more than 70 scientists from publicly funded research institutions and from industry. The external expert panel fed its assessment into the process by way of personal interviews and written statements; the experts participated in the BioFora21 flow of communication via an intranet platform. Other sources of information were international conferences, the relevant scientific literature and international research activities. In order to give adequate consideration to the interests of industry more than 350 German biotechnology firms were asked to provide an assessment in writing. Information provided by pharmaceutical companies was also included in the process. International activities were assessed not only on the basis of written documents, but also by way of interviews with leading experts from the USA and various European countries. The results of the most important internatio-

The basic question was: Can we develop a cross-disciplinary research programme which is innovative and attractive and can be implemented in the medium term (i.e. in the next 5 to 10 years), which meets international standards and complements the ongoing research schemes in the biosciences that are funded by the BMBF?

THE BIOFORA21 PROCESS



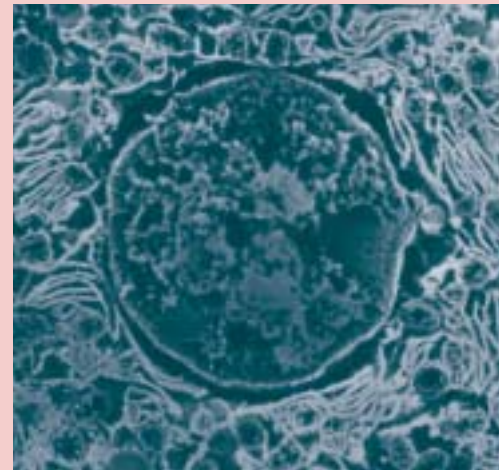
left: Fluorescence micrograph of hepatocytes (liver cells)

national systems biology forum – the Second International Conference on Systems Biology (ICSB 2001) which was held in Pasadena in November 2001 – were also included in the assessment. On 13 December 2001, the BMBF published the result of BioFora21, the new Systems of Life – Systems Biology research programme.

‘Systems of Life – Systems Biology’

Under the ‘Systems of Life – Systems Biology’ funding programme the BMBF will provide grants of up to € 50 million over five years for systems biology research activities

The ‘Systems of Life – Systems Biology’ programme represents a new approach to research funding with a view to establishing systems biology in Germany and tapping its potential for future research and development in science and industry. To this end, an interdisciplinary network of centres of excellence will be developed which will weave the biosciences together with systems studies, engineering sciences, computer science and mathematics to form systems biology. To ensure the medium- and long-term focus of the network of centres of excellence the individual research activities should be conducted in collabo-



Research under the Systems of Life – Systems Biology programme will focus on the hepatocyte system

orative projects (i.e. projects involving science and industry) and concentrate on the hepatocyte as a model system. This type of cell was selected because of the central function it fulfills in higher-order organisms and its great importance for medicine. This includes the numerous essential metabolic processes that take place in the liver, the catabolism of drugs and toxic substances, metabolic adaptations, the transport and processing of substances as well as differentiation and the ability to regenerate. The expectation linked to the ‘Systems of Life – Systems Biology’ programme is that in future methods of quantitative biology will be used more extensively in all disciplines of the biosciences.

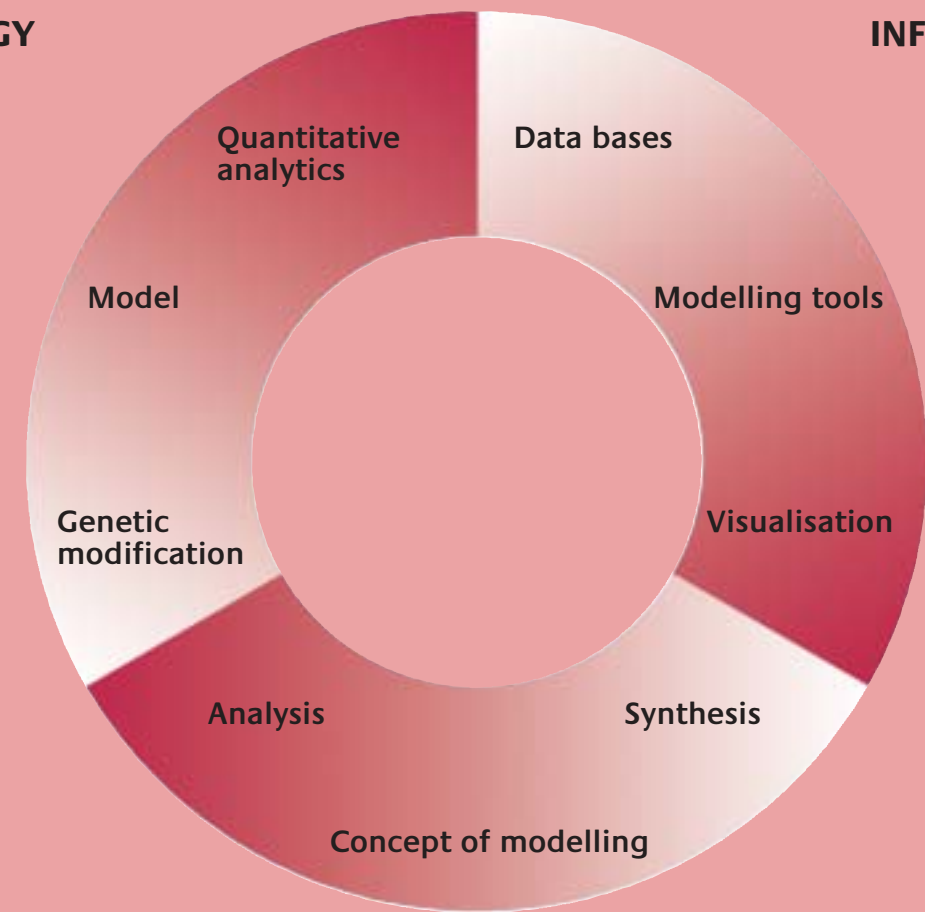
The understanding of systems biology underlying the ‘Systems of Life – Systems Biology’ programme is outlined in the box.

The activities of the ‘Systems of Life – Systems Biology’ network of competence are divided into three thematic **modules**. In **module 1** – *development of methods* – tools and methods will be developed for generating quantitative data for systems analysis. In this process it is particularly important that actual experiments and mathematical and computer-assisted modelling go hand in hand right from the beginning. Another key aspect is the

SYSTEMS BIOLOGY

BIOLOGY

INFORMATICS



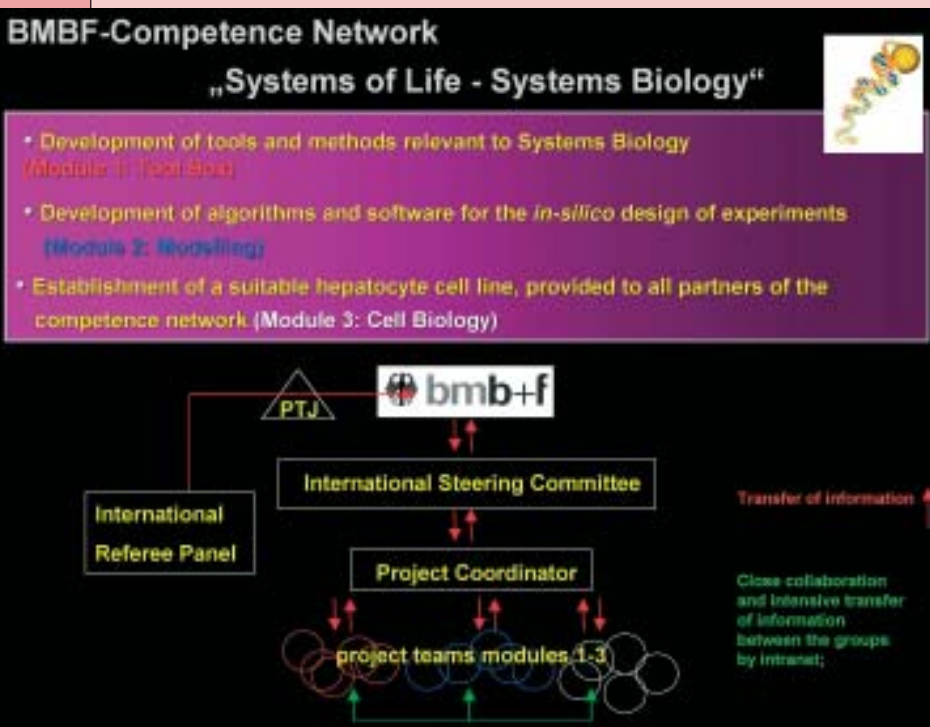
SYSTEMS STUDIES

Over the past 30 years biology has been characterised by a qualitative and descriptive approach designed to investigate molecular details. But in order to understand complex systems properties such as optimal regulation, adaptation and memory both the components of a system and their interactions need to be studied.

Systems biology has chosen this holistic approach. It aims to look into biological processes at the systems level by analysing the complex networks which produce complex functionality. This task requires a shift towards quantitative biology. It is only by means of a quantitative description of a system's components and their interactions that the behaviour observed can be understood.

To attain this objective it is also necessary to take a theory-based approach to a complexity which can no longer be understood intuitively. This is why mathematical modelling plays such a central role in systems biology. Consequently, the effort to develop a quantitative understanding of biology at the systems level is based on an interdisciplinary approach combining concepts borrowed from biology, computer science and systems studies.

The ultimate goal of systems biology is to develop virtual representations of cells and entire organisms. These representations will then enable computer experiments to be conducted in analogy to experiments involving real biological systems. This can pave the way towards predictive biology which – among other things - will help to understand and treat diseases in man.





Kick-off of the 'Systems of Life - Systems Biology' initiative

establishment of standard conditions. And finally, free access has to be ensured to the data generated within this network of centres of excellence.

In **module 2 – modelling** – the algorithms and computer models needed for systems biology will be developed. These include the establishment of a central bioinformatics platform which, among other things, will have to ensure standardised data filing and accessibility.

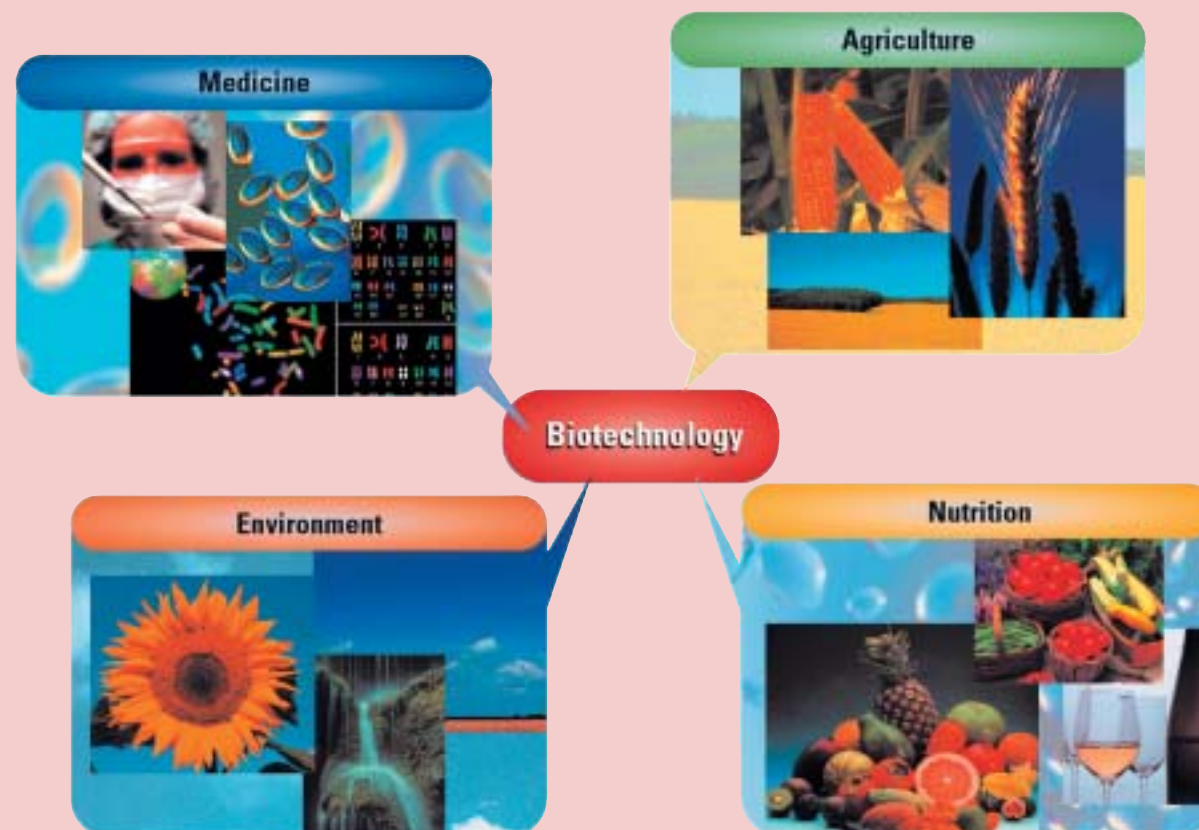
Finally, in **module 3 – cell biology** – the biological systems studied by the network of centres of excellence will be established. Here, the focus will be on developing suitable cell lines which in the medium term will permit concentration on the hepatocyte system.

One of the great challenges and, at the same time, one of the essential success factors of the research programme is the integration of the expertise of different disciplines into the various collaborative projects. Traditionally, research has been structured mostly along the lines of different disciplines, while interdisciplinarity requires special efforts. A second major

success factor is close interaction between standardised data generation and computational modelling. A special **organisational structure** was chosen to take up these challenges and to ensure a targeted and efficient implementation of the programme in the medium and long term. The activities of the network of competence are controlled by an international steering committee supported by a project co-ordinator whose task it is above all to ensure the exchange of information among the working groups receiving grants. Applications for project funding under the programme are assessed by an international referee panel.

Complementing other BMBF research activities

The 'Systems of Life – Systems Biology' programme complements already existing BMBF research activities that were set up under the 'Biotechnology – Using and Shaping its Opportunities' framework programme.



The German Human Genome Project which was launched in 1996 as a joint initiative of the BMBF, the DFG and industry is now in its second phase (1999 to 2002) and has shifted its focus to functional genome analysis. It centres on the functional analysis of medically relevant human key genes and of the genes of model organisms that are necessary to understand the human genes. A National Genome Research Network (NGFN) is being built by further pooling, cross-linking and expanding the capabilities of the best-performing scientific, clinical and industrial partners. In this network five core institutes will closely co-operate with five disease-oriented genome networks. In addition, bioinformatics and proteome research are integrated into the network as platform technologies. The methodology developed for functional genome analysis – e.g. high-throughput technologies for functional analysis – and the knowledge thus gained of the function of genetic and physiological networks form an important basis for systems biology. Additional networks of centres of excellence in genome research are created under

the *GenoMik* research initiative which focuses on microorganisms, and the *GABI* plant genome project 'Genome Analysis of the Plant Biological System'.

The development of techniques and methods for genome and proteome research is complemented and optimised by BMBF's priority funding of proteomics research. The methodological findings gained here can be immediately used in protein network analyses which are required for systems biology. This also applies to the Nanobiotechnology research priority which covers the interface between physics, biology, chemistry and the engineering sciences and which, among other things, aims to develop completely new measuring technologies for the biosciences. Finally, the Bioinformatics Training and Technology Initiative contributes to broadening the computerscience knowledge base needed for systems biology in Germany.

The 'Systems of Life – Systems Biology' initiative will also provide an important methodological foundation for exploring and developing new approaches to the prevention and treatment of human diseases. This research exercise also complements activities carried out under the BMBF's Health Research Programme.

The Systems of Life – Systems Biology research and funding programme is also meant to help shift systems biology more into the focus of modern bioscientific research in Europe. The interdisciplinary approach of systems biology seems to be ideally suited for the creation of European research networks. ■

International Conference on Systems Biology		www.icsb2004.org/
Systeme des Lebens - Systembiologie (Systems of Life – Systems Biology)	Germany	www.systembiologie.de
Genomes to Life (DOE)	USA	http://doegenomestolife.org/
Microbial Cell Project (DOE)	USA	http://microbialcellproject.org/
Complex Biological Systems Initiatives (NIGMS)	USA	www.nigms.nih.gov/funding/complex_systems.html
Service de Conformation des Macromolécules Biologiques et de Bioinformatique, Brussels	Belgium	www.ucmb.ulb.ac.be/SCMBB/
Center for Biological Sequence Analysis, Kopenhagen	Denmark	www.cbs.dtu.dk/
MPI Magdeburg	Germany	www.mpi-magdeburg.mpg.de
Intelligent Bioinformatics Systems, Heidelberg	Germany	www.dkfz-heidelberg.de/ibios/
Theoretical Biophysics, Humboldt University Berlin	Germany	www.biologie.hu-berlin.de/~theorybp/
SiliCell-Initiative, Montpellier	France	http://ibph.pharma.univmontp1.fr/ibphwebsite/team/pepi3.htm
Control Systems Centre, Manchester	UK	www.csc.umist.ac.uk/people/wolkenhauer.htm
Cambridge University	UK	www.zoo.cam.ac.uk/zoostaff/bray.htm
Biocomputing Group Hertfordshire University	UK	http://strc.herts.ac.uk/bio/
Kitano Symbiotic Systems Project	Japan	www.symbio.jst.go.jp/topicE.html
Canadian Protein Engineering Network	Canada	www.pence.ca/pence/english/index.html
Mathematics of Information Technology and Complex Systems	Canada	www.mitacs.math.ca/?section=overview
Delft Systems Biology Centre	Netherlands	www-ict.its.tudelft
The Amsterdam Silicon Cell Programme	Netherlands	www.bio.vu.nl/hwconf/Silicon/
Complex Systems Division, Lund	Sweden	www.thep.lu.se/complex/
Department of Cell & Molecular Biology, Uppsala	Sweden	www.icm.uu.se/
The Molecular Science Institute, Berkeley	USA	www.molsci.org/welcome.shtml
Institute for Systems Biology, Seattle	USA	www.systemsbiology.org/
Alliance for Cellular Signaling	USA	www.cellularsignaling.org/
Cell Systems Initiative, Seattle	USA	http://csi.washington.edu/about/
C. elegans, DKFZ		http://mbi.dkfz-heidelberg.de/mbi/TR/TR64/diplom.Buch.html
A-Cell		www.fujixerox.co.jp/crc/cng/A-Cell/
E-Cell project at Keio University		www.e-cell.org/
GENESIS		www.genesis-sim.org/GENESIS/
M-Cell		www.mcell.cnl.salk.edu/
Modelling Tools		http://mmtsb.scripps.edu/
Molecular Simulation Virtual Lab		http://polymer.bu.edu/vmdl/
Silicon Cell ready to use		www.jjj.bio.vu.nl/
The Harvey Project		http://HarveyProject.science.wayne.edu/development/
The yeast silicon cell		www.bio.vu.nl/hwconf/FP6/ysic/ysicab.html
Virtual Cell (Plant)		www.life.uiuc.edu/plantbio/cell/
Virtual Cell Portal		http://nrcam.uchc.edu/servlets/VCellPortal

DNA	desoxyribonucleic acid, genetic material of all organisms (except for some viruses), consisting of linearly linked nucleotides whose sequence provides the genetic information
EGF	epidermal growth factor, stimulates the growth of (mostly) epidermal cells
Gene Targeting	experimental method used to identify or modify a specific gene
Gene Trapping	experimental method for the direct identification of gene functions; it is based on the insertion into the DNA of a marker which indicates the activity of the gene of interest
Genome	full complement of genes in an organism
Genomics	attempts to analyse the entire genome of an organism
HTS	High Throughput Screening, methods used in gene and protein research
<i>in silico</i>	in the computer
<i>in vitro</i>	in the test tube
<i>in vivo</i>	in the living organism
MAP-Kinasen	mitogen-activated protein kinases; enzymes present in plants and animals which are involved in growth factor signal transduction from the cell surface to the cell nucleus
mRNA	messenger RNA, contains the information required for protein synthesis
Proteomics	analysis of the expression, functions and interactions of all proteins of an organism
RNA	ribonucleic acid, formed by DNA transcription
Transkriptomics	analysis of all genes active in an organism

Gavin, A.-C. et.al. (2002):

Functional organization of the yeast proteome by systematic analysis of protein complexes. In: Nature 405, pp. 141-147

Hattori, M. et.al. (2000):

The DNA sequence of human chromosome 21. In: Nature 405, pp. 311-319

Ideker, T.; Galitski, T.; Hood, L. (2001a):

A new approach to decoding life: Systems Biology. In: Annu. Rev. Genomics Hum. Genet. 2

Ideker, T.; Thorsson, V.; Ranish, J. A. et al. (2001b):

Integrated genomics and proteomic analyses of a systematically perturbed metabolic network. In: Science 292, pp. 929-934

International Human Genome Consortium (2001):

Initial sequencing and analysis of the human genome. Nature 409, pp. 860-921

Kitano, H. (2000/2002):

Perspective on Systems Biology. In: New generation Computing 18, pp. 199-216
Computational systems biology. In: Nature 420, pp. 206 - 210

Palsson, B. (2000):

The challenges of *in silico* biology. In: Nature Biotechnology 18, pp. 1147-1150

Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D., Müller, G. (2002):

Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptor. In: Nature Biotechnology 20, pp. 370-375

Tomita, M. (2001):

Whole-cell simulation: a grand challenge of the 21st century. In: Trends in Biotechnology 19, No. 6, pp. 205-210

Venter, G. et.al. (2001):

The sequence of the human genome. In: Science 291, pp. 1304-1451



This publication is made available free of charge as part of the public relations work of the Federal Ministry of Education and Research. It is not to be sold on a commercial basis. It may not be used by political parties or candidates or electoral helpers during an election for the purposes of campaigning. This applies to all elections at local, Land and federal level and also to elections to the European Parliament. Misuse shall in particular be constituted by distribution at electoral rallies, at information stands of parties, overprinting or attaching of information or campaign material for a particular party. It is also forbidden to hand this document to a third party for use in election campaigns.

Regardless of when or by what route and in how many copies this publication reached the recipient in the form of printed matter or on-line, it may not be used, even outside an election campaign, in a way that could be misunderstood as support by the German Federal Government for individual political groups.