

## **SYBSIOMED Workshop on Challenges for SB**

2 October 2008, Costa Adeje, Tenerife (Spain)

### **Participants**

<b>Alves, Rui</b>	Universidad de Lleida	Spain	ralves@cmb.udl.es
<b>Auffray, Charles</b>	CNRS, Villejuif	France	auffray@vjf.cnrs.fr
<b>Cloutier, Mathieu</b>	Hamilton Institute	Ireland	mathieu.cloutier@nuim.ie
<b>Girolami, Mark</b>	University of Glasgow	UK	girolami@dcs.gla.ac.uk
<b>Eufinger, Jan</b>	DKFZ Heidelberg	Germany	j.eufinger@dkfz-heidelberg.de
<b>Fell, David</b>	Oxford Brookes University	UK	dfell@brookes.ac.uk
<b>Goryanin, Igor</b>	University of Edingburg	UK	goryanin@inf.ed.ac.uk
<b>Gross, Thilo</b>	Max-Planck-Inst. Magdeburg	Germany	gross@pks.mpg.de
<b>Hornberg, Jorrit</b>	Schering-Plough, Oss	The Netherlands	jorrit.hornberg@spcorp.com
<b>Mendoza, Eduardo</b>	LMU Munich	Germany	mendoza@lmu.de
<b>Plahte, Erik</b>	CIGENE Aas	Norway	erik.plathe@umb.no
<b>Torres, Nestor</b>	Univ. de La Laguna	Spain	ntorres@ull.es
<b>Schürrie, Karsten</b>	DECHEMA e.V.	Germany	schürrie@dechema.de
<b>Wellstead, Peter</b>	Hamilton Institute	Ireland	peter.wellstead@nuim.ie
<b>Wernisch, Lorenz</b>	University of London	UK	lorenz.wernisch@mrc-bsu.cam.ac.uk
<b>Wolkenhauer, Olaf</b>	University of Rostock	Germany	ow@informatik.uni-rostock.de

The participants discussed which 'grand challenges' – be it scientific or conceptual/organisational – they expect medical systems biology (MSB) to meet in the coming years. The underlying notion of systems biology is that it is about mathematical modelling and computer simulation – a predictive science that generates testable hypotheses. Modelling is believed to 'make a difference' analogous to 17<sup>th</sup> century science<sup>1</sup> when Johannes Kepler – analysing large amounts of data which Tycho Brahe had carefully measured before – built his celestial mechanics models which eventually inspired Isaac Newton to derive the general law of gravitation.

The nature of a 'grand challenge' was considered also worth discussing: Is it a 'hard' problem like pertinacious mathematical conjectures (e.g. the Riemann hypothesis) calling for novel ways of thinking? Or is it a matter of efficient, very clever combination of available technologies to achieve an ambitious goal (e.g. Apollo lunar missions)?

### **Scientific challenges**

In February 2008, researchers at a workshop held in Tokyo announced a bold project: to

---

<sup>1</sup> an analogy emphasised earlier by Stas Swartsman and Jim Ferrel

create, over the next 30 years, a 'virtual human'. This molecule-based computational model would describe the systems or networks of interactions between the tens of thousands of genes and proteins that underpin biological processes in both health and disease. In the workshop it was consensus that such 'virtual human' would not be a very realistic perspective. They expect that every big scientific challenge in MSB will eventually result in a diversity of models, not necessarily mathematically equivalent, but each one able to reliably predict certain parameters. A 'grand challenge' would be to push MSB beyond the achievements of the current Physiome Project with defined medical applications and diseases being addressed. All workshop participants agreed that the best (disease) model system is the human body. However, a system's behaviour can only be understood through stimulus-response experiments. This will for ethical reasons not always be possible. In many cases it is also impossible to generate a stable, reproducible experimental system from human derived samples. Despite known deficiencies of established animal models, e.g. mouse model in cancer, well defined mutations can only be studied in such model organisms.

The experts present at the workshop clearly identified cancer and neuro-degenerative diseases calling for MSB approaches in order to make significant progress. Some participants pleaded to also consider infectious diseases in addition to the disease areas looked at so far within SYSBIOMED (cf. workshop reports at [www.sysbiomed.org](http://www.sysbiomed.org)). Infectious diseases pose a real, growing threat to public health. A lot of knowledge is being generated by current microbial SB projects and will be eventually complemented by insights from systemic studies of host-pathogen interactions and the immune system. Important funders of medical research like the Bill and Melinda Gates Foundation and the WHO identified infectious diseases as major challenges to humankind and strive to fill the gap the pharmaceutical industry created by neglecting anti-infectives research. The workshop participants welcomed the Innovative Medicines Initiative (IMI), to which industry and EC equally contribute €2bn total funding, as a good framework to implement MSB projects which address the above mentioned medical challenges.

With regard to the dramatic threat of malnutrition and famine to almost a billion humans the participants agreed that funders and researchers should let loose on the 'grand challenge' of generating C4 rice plants, an endeavor which is likely to benefit from SB projects in plant biology.

### **Organisational challenges**

Going for 'grand challenges' carries some risks. For example, one could possibly neglect useful model organisms or important pathways. Moreover, the big money big programmes use to come with could attract the wrong people who sell conventional research under a new label, or, not unusual in medical sciences, one sees 'empires' of research activities gravitating around key figures who jealously dominate the project while discouraging creative (younger)

scientists. During the discussion it soon turned out that the management of large projects is crucial and a 'grand challenge' by its own right. Although sufficient funding is essential, 'grand scientific challenges' would require the motivation of the 'right' people to co-operate and to combine the 'right' resources. Referring to the well-managed SysMo initiative the participants considered ERA-Nets especially helpful to support integrated large-scale efforts. It was questioned whether top-down managed initiatives would be more successful than programmes run by more autonomous groups with a high degree of freedom - bearing the risk that the partners are tempted to follow their own favourite research priorities.

The workshop participants see room for improvement on the management side. In order to make sure that 'right' qualifications and expertise come on board, one could adopt a 1year grant scheme as implemented in the Swiss SystemsX.CH initiative. Such grants would allow to evaluate "newcomers" and young researchers before closing longer-term contracts. Finally, a rather radical concept was discussed: Why not trust project co-ordinators with the authority to allocate funds (not assigned to at the beginning) in due course according to the performance of the partners? A more extreme view would require the co-ordinator to solely take charge of the grant application and to build the consortium.

## **Education**

Education was unanimously considered a 'grand challenge', too. Tight collaboration of biologists/physicians, engineers and informaticians/mathematicians requires adaption and understanding of concepts from the other disciplines and to overcome 'cultural' barriers built from conventional mindsets and deficits in familiarity with other ways of thinking. Education obviously is key to achieve optimum conditions for supplying this future technology sector with the qualified researchers it needs. Education in MSB deserves special attention. In most countries training in the medical sciences is quite different from training in biosciences and engineering, hampering the introduction of mathematical modeling and quantitative methods. Since many research institutes and labs in the medical sciences recruit doctoral and postdoctoral researchers with a medical training, it is necessary to provide more opportunities of experimentalists working in a medical environment to interact with modellers.

## **Views**

Finally, the participants took the opportunity to give their individual views on some 40 issues listed in a questionnaire. Although having different backgrounds the participants shared similar opinions on what biological systems are, on the deficits of education in biology and engineering/ physical sciences, on experimental techniques needed, and on the main hurdles for model integration. Not surprisingly, most agreed on Dennis Noble's heart model and the Physiome project being examples of SB "success stories" already achieved whereas the

participants' nominations of successful applications of mathematical modeling in the life sciences proved far more diverse. The answers to many questions, e.g. related to grand challenges ('moonshot projects'), models, or the priorities of project preparation, showed a diversity of aspects which, in summary, represent a sound, concise plea for large-scale programmes. The complete feedback to the questionnaire is given below:

## Your View on Systems Biology for Medical Applications

(Please notice: individual answers are separated by // in the text)

**1. Name ---**

**2. Affiliation ---**

**3. What is your current major research topic in Systems Biology (SB)?**

Modeling of GRNS, brain energy metabolism // transcription in *E. coli* // protein (interaction) networks // dynamics of biological networks // modeling of pathways involved in tumour progression // metabolic networks // Parkinson's disease // Computational SB // network reconstruction, design principles // multi-scale integration // chronobiology, systems virology, spatial and stochastic modelling // signal transduction, cancer, and autoimmune diseases //

**4. What does the word "Systems" in Systems Biology refer to?**

Aggregate of interacting entities of processes which produce or exhibit system properties or behaviour // systems properties that arise from the interactions of biomolecules and cannot be assessed by traditional reductionist approaches // level of description is on the scale of systems (rather than molecules or populations) // complex interactions of components coordinated in task-oriented manner // set of components that can be studied as a whole in isolation from other parts of the cell and still behave as they do in toto // given the reductionist (or "component-oriented") thrust in the past, the SB approach can also be viewed as integrative //

**5. Which part of Biology do you think is most relevant to SB?**

Evolution, the ultimate text of any model of living organisms // biological regulation // biophysics // quantitative questions of molecular biology // all (from cells to populations) this is largely gated by available measurement technologies // all beyond structural molecular biology // biology is an application area as are all areas where dynamics and feedback are relevant // molecular biology and physiology // molecular biology // all of biology would benefit from integrative approaches // medicine. Of course, principles have to be discovered and methods have to be developed using simple systems, but ultimately SB will only receive the attention and input it needs if it has major impact on society //

**6. What is a systems approach?**

analyse a problem by looking at multi-parameters instead of single cause -single effect relations // orchestration of experiment and theory // approach integrating different experimental and theoretical techniques possibly on multiple scales // studying emergent 'systems' properties (systematic collection of quantitative information, how the system is controlled // considering a system in a holistic rather than reductionist view defining relationships between components // bridging between different levels of explanation // understanding by showing how properties at one level can be accounted for, logically and quantitatively, in terms of lower-level interacting components typically requiring a mathematical model // „The consideration of the behaviour of a process by considering the interactions and feedback that produce the observed behaviour“ (Arnold Austin) // iterative process combining question-driven inquiry, precise measurements and perturbations, modeling and generation of hypotheses // a systems approach studies a collection of components and their interactions in terms of the structure(s) and dynamics, formalizes the insights in a computational model and validates through experimental data its validity //

**7. What is systems-level understanding?**

modeling of partial reactions and prediction // coarse-grained understanding that describes causal mechanisms on a systems scale // recognising patterns of behaviour at the higher level that arise from specific interaction patterns between the components // Ockham's razor at dynamic behaviour level: analysis should be sufficiently detailed to expose the features of inter-

est – but not more detailed // kinetic modeling and pathway analysis // understanding how different components interact // understanding biological processes and behaviour that cannot be derived solely from component properties or more global levels of biological organisation // understanding of a set of biological processes uncovers design and operating principles underlying the mechanisms of the processes and their function //

#### **8. Should SB be taught as a separate undergraduate study programme?**

Yes, there is a successful programme at Harvard (David Botstein)

Not necessarily, SB benefits from people with different backgrounds

No. a few courses are sufficient for students having a good background in mathematical physics //

No // No, better train specialists then let them collaborate (using a common language) // include more mathematical modeling topics in molecular biology education //

No – it could be a component of natural and physical sciences education //

Not in detail, systems biologists need a specific area of expertise to bring to interdisciplinary collaboration //

No. We should give biologists the same mathematical and analytical training as engineers and physicists – in addition to their current syllabus // Yes // No // Yes, should be combined with at least one advanced course // No, we need courses which enhance the understanding of the different scientific and engineering disciplines for complex systems and methods to study them and which enable these disciplines to better communicate with each other, and especially with biologists and clinicians // Yes, already early on. We frequently see Systems Biologists who are unable to communicate because they came from different backgrounds, e.g. Theoretical Physics vs Molecular Biologists. The theoreticians often do not have the goal or application in mind and may focus on phenomena like robustness or oscillations as a goal. Biologists on the other hand do not understand the analyses methods, the type of data they need to produce or how to actually integrate their data. As they have different interests, they set different goals for their research. Therefore I think especially the first years of undergraduate studies should focus on different aspects of Systems Biology, after which students will have to choose their specialization direction //

#### **9. Should Masters Programmes in SB specialise or cover a broad spectrum of topics?**

Both. Specialisation is needed for research (PhD courses), broad basis is needed for collaboration // broad spectrum // both // broad education can help prevent fragmentation // broad spectrum plus engineering disciplines // broad – specialisation comes at PhD level // probably only viable if a broad spectrum of topics is covered, specialisation best offered via summer schools // make them conservation courses in 3 classes: systems theory and maths for biologists, biology for physics, maths and biology for control engineering graduates // 2 x broad spectrum // broad, and some case studies // first year with a broad spectrum of topics and specialization in the second year // broad spectrum, but there should be research projects within research departments of at least 1 year in which students specialize on a topic //

#### **10. State the most important change required in training //education:**

##### **a. w.r.t. the biological sciences:**

basic training in math, statistics, data analysis, computing // apply this knowledge to simple models // learn to think in terms of formal models that can be proven wrong // more maths // expose biologists early to conceptual models // quantitative measuring and data processing // additional mathematical training is vital // improved mathematical and computational education (use of mathematical computational models in teaching) // informatics, bioinformatics

and modeling // basic knowledge of mathematical and engineering computing formalisms // integration of math // computational aspects in teaching of all biological areas (e.g. new textbooks with equations and models instead of just pictures and “cartoons”) //

**b. w.r.t. to the engineering and physical sciences:**

adapt curricula to the needs of biology (e.g. nonlinear systems in basic ODE courses) // introduce experimental level biology // develop critical skills in dealing with uncertain models and interpretation // more biology // (re)introduce research oriented courses of study // wet lab experience and familiarity with modern technologies in molecular biology // more biology, not just biochemistry // awareness of the richness of biological systems // biochemistry, bioinformatics and modeling // focus on uncertainty of biological circuits // awareness of biological variation and evolution // inclusion of biological system aspects in as many courses as possible, offering “bio-tracks” //

**11. Name the most important quantitative methods which biologists should be familiar with:**

linear algebra, simple calculus and ODEs, modern data analysis methods // mass balances and ODEs // writing and solving ODEs numerically // TIRF // FLIM // find a common language for communication with modellers // statistical interference, stochastic processes, mathematical modeling // (any would be an improvement) sensitivity analysis // dynamical systems methods and feedback systems analysis // ODEs, PDEs, graph analysis, statistics, Bayesian optimisation, parameter fitting // quantitative „omics“ methods // differential and vectorial calculus // the collection in the book by Klipp *et al* is a good first approximation // should know kinetic rate equations and be able to use those to do some simple modeling, to get some kind of hands on experience. They should know the principles of control or sensitivity analysis //

**12. Name the most important biological concepts which modellers should be familiar with:**

basic biology, chemical and physical "reality" of the cell // evolution and that nothing is as simple as written in the textbooks // phosphorylation // First, modellers should understand the meaning of words (terms?) in biology // cellular function in general // variation, evolution and population genetics // variation and evolution // basic concepts of molecular biology // depends on the problem, a first reading of a Molecular Biology Textbook is useful, the collection in the book by Klipp *et al* is a good first approximation // should become familiar with some cell biology (what is a cell composed of), some molecular biology (including gene expression regulation), some histology (for instance to make sure they know there are different cell types) and basics of signal transduction. Furthermore they should be aware of the pathology related to several diseases. In principle, it is most important that they will be able to understand the biological problems they are working on and also why they are working on it //

**13. Considering pathway //network diagram drawn by biologists, which vocabulary do you consider most important in the diagrams description? (e.g. modulation, activation //inhibition, feedback, ...)**

activation – inhibition // biologists should use standards like SB Graphical Notation, BioPAX, etc. // common vocabulary (e.g. feedback is clear for engineers, biologists use it differently) is essential // clear distinction between different interaction types (non-covalent complex formation, catalysis, covalent modification) // SBGN // interaction, activation, inhibition // develop an interactive method to capture the vocabulary the partners use and clarify them in a stepwise manner //

**14. What features //functionality do you miss with existing software tools?**

easy ways to export or compile results into a simulation report // ways to track changes in models, manuscripts and simulations // Nothing's missing // standard for spacial ML // I write my own software // difficult to easily construct toolboxes for a particular problem // I hardly use software tools, I am mostly developing methodology // automated network drawing from mathematical network description // Matlab is sufficient // model building // intuitive integration of bioinformatics tools // too many // in general. professional-grade (robust, well maintained) software // software tools should be accessible to people who do not know computer science or programming language. I therefore like windows-interfaced programs such as Copasi //

**15. State the difference between the notion of a “networks” and a “pathway”:**

Pathway implies a flow of material, network could be related to more functional interactions // network=multiple interacting networks // no difference // a pathway, despite possible feedbacks, implies a directionality // pathways are thought as linear branching networks with few feedbacks, networks are thought as interconnected graphs. They correspond to different strategies of exploring mechanistic systems behaviour // pathway is a linear modular directed acyclic graph, a network is a cyclic graph composed of multiple pathways // the difference between a roadmap and a motorway: pathway is a quantitatively dominant route through the network (and hence misleading label if applied to cells other than those for which it has been demonstrated to be appropriate) // pathway is a form of network for transmission of specific signalling information // there is no definition for both in biology // pathway is a special case of a network (which integrates many pathways) including genes, proteins and metabolites that are involved in a specific cellular process or response // pathway: mostly linear set of reactions, network: linear + multiple branching sets of interactions // “pathways” tend to be sequential, i.e., there is a main line of interactions, networks are more general graphs of interaction // the difference is merely semantic //

**16. State the difference between the notion of a “module” and a “motif”:**

module: cell subsystem producing a specific function // motif: pattern arising from many possible interactions // module is specifically characterised – motif is general // a module is almost isolated from surrounding entities, a motif can still be identified if it is densely linked with surrounding // module represents a subnetwork with well-defined function and interface with the rest of the network, motif is an empirical observation of frequent network patterns // a motif is structural a module is functional // a motif is a recognisable pattern of component interactions which may have a functional significance, but not necessarily isolable from the system. A module is a subsystem with many more internal interactions than with the rest of the system – definable input //output characteristics // no difference // module: set of elements involved in a given process – motif: regularities found in comparable modules // module: set of components associated with physical, functional property, motif: set of interacting components with specific arrangement // "Motifs" tend to be smaller, repetitive (statistically significant) and are characterized by “information processing” functions ; “Modules” are larger, usually unique and fulfil “biological function”. However, it is very difficult in several cases to really differentiate, e.g. in the bacterial chemotaxis system with only 6-7 components //

**17. Define “cross talk”:**

necessary to introduce as pathway models tend to be too simplified // molecular signal activating many subsystems // interaction between pathways // cross-talk frequently destroys quantitative predictions // interactions normally missed when studying the pathway under a particular condition when it can be considered isolated and independent from other pathways // exogenous inputs //outputs to the system representation // a term needed by those who insist on using a pathway paradigm with reference to a network where it is inappropriate // interaction

between primary transmission channels // capacity of different pathways // modules to exchange information and to regulate each other // interaction between motifs, modules or systems // a (normal) network connection between pathways (usually discovered later) // cross talk is used to indicate the connections between pathways. It is a term that remained from the time period in which biologists were discovering that pathways are not isolated and not linear //

**18. Define the meaning of “cell function”:**

the ultimate c.f. is reproduction // evolutionary purpose of a module or processes essential for survival, fitness and reproduction // should read „cell functions“ like mitosis, meiosis, stress response etc. // in multicellular organisms cells may have immediate functions (skin cells, endothelial cell), higher level functions are hypotheses about properties arising from interactions with other cells // physiological state of a cell // ultimately survival and reproduction // contribution to a biological process // function fulfilled by some processes within a cell // the process a cell performs (or a product it produces) for a larger system (e.g. organ, body, ecosystem) //

**19. Name sources of complexity in SB:**

the need to integrate knowledge from many diverse fields // non-linear feedback // non-linearity, multi-scale, size and identifiability problem // ... not sure whether current mathematical tools are appropriate for describing life // wide dynamic range // non-linearity, dynamics, feedback and scale // evolving self-replicating machines // environmental changes, individual variation and hierarchical levels of organisation // besides the system (and its context) itself: inappropriate data, difficult communication between disciplines, inadequate computational methods // organizational complexity comes from communication problems and access to biological data. Systems complexity arises when a system is too complex to handle in your brain //

**20. Give an example of a success story of mathematical modeling in the life sciences:**

modeling of a complete organism (genome-scale model) // experimentally confirmed prediction that one of the leu\_tRNAs had to read another codon // Tyson&Novak // modeling budding yeast cell-cycle (Chen et al) // optimisation of bioengineering reactors based on flux analysis // Hodgkin-Huxley model of transport across cell membranes, axon models of Fitzhugh-Nagumo, heart model of Dennis Noble // mathematical cell cycle models, metabolic control analysis // Hodgkin-Huxley and Physiome project (Peter Hunter) and heart model (Dennis Noble) // Hodgkin-Huxley, Watson-Crick // Demand theory and synthetic biological circuits // Eric Davidson’s modeling of the development of sea urchins // “Modelling biological rhythms” paper (Mendoza *et al.*) with a whole section on predictive successes //

**21. Give an example of a success story of SB in drug discovery and medicine:**

yet to come, however, Physiome project is a nice existing example // Gleevec // hard to tell as all projects I know about are currently ongoing and do at best look promising // 2 x heart model (Dennis Noble), Merrimack growth pathway inhibition in clinical trial // TB ARO pathways // Entelos and Merrimack Pharmaceuticals, and the work by J. Keasling on artemisinin //

**22. State the most important advance necessary in SB:**

**a. w.r.t. drug discovery:** trying to integrate the psychotropic effects of drugs (important but currently neglected) // small molecules for molecular therapy // sufficiently predictive models of disease mechanisms // sufficiently exclusive models (e.g. cell cycle + apoptosis + growth factor // cytokine signalling) with sufficient detail and molecular targets represented

// working with clinicians // integration of structural biology and bioinformatics with modeling and control analysis of responses // integration of multiple types of data to predict drug efficacy and patient response // disease-focussed virtual patient models (as early testbed for drug candidates) //

**b. w.r.t. technologies:**

happening by advances in computing power and HT technologies // even better labelling technologies // better integration of imaging technologies // HT measurement of enzyme activities // single cell and single molecule technologies // devices with biological behaviour //

**c. w.r.t. methodologies:**

scale (time and space) // multi-scale models // measuring kinetic parameters in single cells // stochastic modeling and analysis // agent-based modeling //

**d. w.r.t. formal concepts:**

scale (time and space) // more efficient extraction of insights from diagrammatic representations // biochemical formalisms are taken too literally, we miss adequate techniques as ODEs and stochastic ODEs adapted to deal with systems performing functions in highly variable conditions (robust evolvable systems), a situation not reducible to the classical chemical kinetics approach // formal framework to describe life // multi-scale modeling // understand evolution // formalism for integration of diverse types of modeling //

**e. w.r.t. experimental techniques:**

new technologies must yield quantitative results // single molecule technologies for eukaryotic cells // large-scale data on states and localisation of cell signalling proteins // improved temporal and spatial resolution // single cell and single molecule technologies // real-time non-invasive quantitative measurements at single-cell, single-molecule resolution //

**f. w.r.t. research funding**

apply SB to problems relevant for society // easy funding of small high-risk short-term projects (1y) would be great // trans-atlantic funding // cooperation between national funders // more integrated activities // decrease size (no.?) of groups that can apply for funding // support of cross-disciplinary training of researchers //

**23. Which standardisation effort in SB do you consider most important (and why)?**

of model formats (no one else would do it) // development of SBML // metabolic reactions // SBGN + SBML + BioPAX because we need standards for model exchange // experimental systems will continue to be too diverse // model structures and languages // SBGN, SBML2 // SBML and CellML modeling language standards // representation of experimental data suitable for modeling // ontologies for consistency in further formalisation //

**24. What is in your view the main hurdle for the integration of models?**

Scales (time and space), a unified framework will be required // models are built for different purposes // analysis of large models // different scales and different problems for which they are constructed // finding a formal framework for integration at different spatial and temporal scales and levels of abstraction // nomenclature of molecules, components, states of molecules // model annotation // models developed for different conditions and assumptions are hard to integrate, makes it more laborious than building new models // lack of common mathematical framework // lack of compatible concepts, initial conditions, generally insufficient information to reproduce results and validate //

**25. Rank the following decisions w.r.t. to importance in preparing a project: (a) model organism (b) cell line (c) network //pathway (d) genes //proteins (e) technologies (f) data management (g) design of experiments (h)type of model**

c-(afd)-h-e-g-f // a-g-d-b-c-e-h-f // h-g-e-c-a-b-f // g-a-e-c-d-h-f-b // c-a-d-b-e-g-h-f // a-f-g // e-d-a-b-g-e-f-h // (gh)-(ef)-(abcd) // when preparing an SB project, the most important point is a consensus (as clear as possible) on the question to be addressed and the solution envisioned. The problem formulation (from the experimenters viewpoint) might already point to the model organism, cell lines, genes //proteins, network //pathways to be studied. An initial design of experiment would follow to give a sense of the data that could be obtained. These factors – problem and data - will then constrain the type of model to be built. Data management will help with information from literature to build the model. Depending on the insights and predictions (hopefully) produced by the model, the experiment design can be refined to validate the model predictions; i.e. follow the SB cycle à la Kitano (2002) as much as possible // a (this directly follows from your hypothesis or problem definition)-b-g-f-e-h-c-d //

**26. How important do you think is the definition of “grand challenges” in SB:**

'Grand challenges' are social, political, economic and scientific // important as it will have impact on society // doubt whether a project of that size can be effectively coordinated at present // showcase projects carrying the risk of rejection of other projects by funders // very important // driving force of community efforts // useful at certain points of the development of the discipline. Should however be from credible //accepted authorities, otherwise not effective. Examples of successful GC efforts are Hilbert's Problems (1900) and the Millennium Problems (2000) in Mathematics. Not aware of similar successes in other sciences // not important //

**27. State a time frame in which a grand challenge should be solved //realised:**

The time it takes. Question whether mega projects are the right answer // decades // 10years // 10-15years // 30years (scientists lifetime) // 10-20years // 10years // 30years // not feasible to give a time frame // weeks to decades and more // 15-20 years //

**28. Would you prefer a “virtual cell” project over a “virtual human” project?**

prefer virtual human, as the virtual cell is already there (Tomita et al 1997) // prefer virtual cell (however, it is too large, what question would it answer) // 2x yes // could be done in parallel // no // both are complementary // no preference – the question(s) to be answered and the data for validation to be generated need to be specified as clearly as possible // both are important //

**29. “Moonshot projects” in SB – which would be your favourite?**

"Virtual human" would have lasting impact on society // understanding of immune dynamics // model of tumour progression explaining tissue specificity // virtual yeast cell – would be more clear-cut whether you'd arrived // get beyond Physiome project with sufficient resources // virtual human // enzyme projects to determine accurate kinetics of individual proteins // one that have impact on a broad spectrum of diseases // remedy for tuberculosis //

**30. How many individuals //groups //institutes would you expect to be involved in large scale research (“moon shot”) projects?**

Many in many different projects // About 20 max. groups from all around the world // 42 // no. of participants according to needs, should be dynamic // 20-40 // A whole institute devoted to the project // as many as required, manageable size 20-50 groups // at least two levels – the core project and the broader research community. The core project would comprise 100-200 persons // if organised well, there can hardly be too many groups //

**31. What funding volume would be required to realise moonshot projects?**

€ 0.5m //year\*group for 10years // funding comparable to pharmaceutical projects // > € 10bn // € 100m // several 100m € // 2 x depends // depends on expertise required, probably 10-20 times the usual grants // make sure that the researches will do what they promise. Then fund based on what the goals, milestone and deliverables are //

**32. Which existing large-scale research project would you consider a success story?**

CERN, Physiome project // Human Genome Project // Heart project or the Flagellum project // Physiome project has that potential // both whole genome metabolic reconstruction and Physiome project are partial successes // Physiome project // sequencing of the human genome (and other species) //

**33. If you could award a prize to a living scientists for the greatest contribution to your work //field, who would that be and what should be prize be awarded for?**

René Thomas for his investigations in the role of feed-back in biological systems // Otto Berg for stochastic gene expression, facilitated diffusion, diffusion limited dissociation etc. // Alfred Wittinghofer // Mark Newman&John Tyson // Spiegelhalter (statistics) Sejnowski (neurobiology computation) // Hans Westerhoff for tirelessly motivating and integrating European efforts in SB // 2 x Dennis Noble for the virtual heart // Michael Savaggan for contribution to biochemical systems theory and studies in design principles // Leroy Hood for enabling technologies and education in SB //

**34. What would you consider is the role of a (mathematical //statistical //computational) model?**

repository of knowledge and hypotheses of the biological system it emulates // to clarify thinking and deduction // to quickly rule out wrong hypotheses and to extract functions from static //structural data in order to generate new hypotheses // unambiguous description of hypotheses and mechanisms, predictive, intelligent guess when developing new concepts and ideas // testing verbal //qualitative explanations // strategic as framework for analysis and repository of quantitative knowledge // generation of plausible hypotheses // allows verification of concepts, suggest alternatives, prioritise wet lab experiments, and to test hypotheses that cannot be studied experimentally (thought experiment) // integrating multiple types of data, predict systems response to perturbations // role in communication with experimentalists particularly in a stepwise process of formalizing ideas, hypotheses and moving the iterative SB cycle forward // help integrate the available information and to identify general principles //

**35. What is a “phenomenological model”?**

All models in biology are phenomenological // model that describes experimentally observed phenomena // non-mechanistic model focused on describing observations // model describing a phenomenon far from ,the first principles’ using ad hoc hypotheses and generic functions fitted to observations // which represents observable properties // reproduces input-output relationships with functions //parameters that do not relate to physical characteristics of the components of the system // no particular details // model that uses kinetic functions chosen to fit the data // phenomenological models take an external view of a system, in particular with respect to the dynamics of its behaviour //

**36. What is a “mechanistic model”?**

A reductionist model // model that targets a simple mechanism // quantitative model including interacting physical objects // constructs a function which mimics biochemical mechanisms known about the system // which represents underlying mechanisms // a model whose functional form and parameters are derived from physical properties of the components // mecha-

nistic models (as usually used by biologists) reflect molecular and some systemic properties of the system //

### **37. What is a “physical model”?**

Based on explicitly formulated physical laws, questionable, whether this is sufficient // model that describes a known physical phenomenon with very few hypotheses // same as mechanistic model // includes physical mechanical processes // constructed from the very first principles (atoms and molecules) // mathematical equations model // model where individual steps of the process are known and kinetically characterised // take into consideration physical properties of the system (somewhat more detailed than mechanistic models) //

### **38. What is a “large-scale model”?**

A genome-scale model of a cell // a model containing >100 different non-identical components and cannot be easily treated by standard methods // simply that they are large and highly detailed // large no. of components spans temporal and //or spatial scales // >100 elements // large number of dependent variables // 38. “Large-scale” is relative – one needs to state “large relative to what”. Examples: large relative to what can be dynamically simulated with ODEs (say 1700 ODE’s used in a Hepatitis C infection model of Vertex Pharmaceuticals vs. the 1-2 hundreds usually considered; or large relative to the relevant part of the genome e.g. genome-scale metabolic network meaning that the majority of the genes encoding enzymes are included in the network ... //

### **39. What is a “useful”, “good”, “practical”, “predictive” model?**

models that answer questions, predict the outcome of new experiments (within a certain range) // descriptive models require further analysis to achieve explanations // predictive model is tautological // one that produces results that are contrary to expectations // predictive, otherwise no good // providing insights and suggesting experiments // usefulness has several levels: show consistency of information, explain available data, show new relationships, make predictions, suggest new experiments to validate,... //

### **40. What is a “descriptive” vs. “explanatory” model?**

descr. fits the data – expl. makes you understand why a systems operates as it does // recapitulates existing facts in mathematical equations, expl m explains s.th. that cannot be explained from already known facts // descriptive=phenomenological explanatory=mechanistic // descriptive: how things work, explanatory: why things work

### **41. What distinguishes a biological system from a physical system?**

evolves and eventually dies // level of knowledge we have on both // purpose in the context of evolution // complexity // p.s. contain fundamental interactions (gravitation, electrodynamics ...) and b.s. contain emergent interactions (activation, phosphorylation, predation, ...) // evolution // alive vs. not alive // autonomous freedom of new entities // evolution // a biological system is a particular case of a physical system // all biological systems are physical // a physical system can be everything including a biological system. A biological system is necessary (or at least has a function, cf 18) for a living system //

### **42. What distinguishes a cell from a computer?**

Evolution // cell is probably more intelligent // complexity // comp. are designed to solve huge range of problems, cells have evolved to solve the problem of survival // totally different design: von Neumann architecture vs. undirected evolution process // a cell has life // cells haven’t got keyboards // self-replication // number of components and interactions, chemical composition, self-replication // lipid membrane, polynucleic acids as information storage de-

vices, chemical energy carrier ATP // serious question? A cell is (part of) a living entity, a computer is not. A cell is far too complex to be understood by the human brain. A computer is thought up by human brains //