

## **SYSBIOMED Workshop on Dynamic Principles of Cellular Function**

October, 24<sup>th</sup> and 25<sup>th</sup> 2007 Frankfurt am Main, Germany

### **Participants**

<b>Blüthgen, Nils</b>	The University of Manchester	UK	nils.bluthgen@manchester.ac.uk
<b>Bruggemann, Frank</b>	Vrije Universiteit	NL	frank.bruggeman@falw.vu.nl
<b>Fell, David</b>	Oxford Brookes University	UK	dfell@brookes.ac.uk
<b>Girolami, Mark</b>	University of Glasgow	UK	girolami@dcs.gla.ac.uk
<b>Lunkes, Astrid</b>	ESF	France	alunkes@esf.org
<b>Nimmegern, Elmar</b>	BMBF	Germany	Elmar.Nimmegern@bmbf.bund.de
<b>Schürrie, Karsten</b>	DECHEMA	Germany	schürrie@dechema.de
<b>Snoep, Jacky</b>	University of Stellenbosch	South Africa	jls@sun.ac.za
<b>Thieffry, Denis</b>	INSERM	France	thieffry@tagc.univ-mrs.fr
<b>van Leeuwen, Ingeborg</b>	University of Nottingham	UK	pmzivil@maths.nottingham.ac.uk
<b>Wilkinson, Darren</b>	Newcastle University	UK	d.j.wilkinson@ncl.ac.uk
<b>Wolkenhauer, Olaf</b>	University of Rostock	Germany	ow@informatik.uni-rostock.de

### **Why a workshop on Dynamic Principles of Cellular Function?**

Cell function (growth, differentiation, proliferation, apoptosis etc.) are nonlinear spatio-temporal processes operating across levels of the structural and functional organisation of cells. Understanding the dynamic principles of cell functions is the basis for advances in understanding the onset of diseases.

The recent advances in cell biology research suggest a rather astounding conclusion: the more we learn about cells and complex cell populations the less we are encouraged to generalise. Apart from the inherent flexibility of living systems evolution has obviously invented a large diversity of "systems solutions", which, differently organised, produce similar "readouts". Despite the tremendous technological advances we are far away from direct, quantitative observation of living systems like we are used to in the physical and engineering sciences: We assume the functions of cells by envisioning their components and changes thereof. Instead of asking general questions about the functioning of cells, we interpret data for studying a particular signal transduction pathway, for a particular cell line, cell type generated within the limitations of the measurement devices used (microcopy, transcriptomics, proteomics etc.). Our ability to understand completely how a cell works is probably limited, partly because we may have difficulties in defining the functions being implemented and partly because we may not recognise the collection of components that implement the function. Fortunately, this does not necessarily mean that one cannot predict the responses of a biological system by computer

simulation.

However, cell functions are dynamic processes that can be understood in terms of dynamical systems theory with modelling employed to formulate and test hypotheses while supporting the design of experiments: If there are many pathways relevant to one cell function, why study pathways in isolation instead of focusing on a cell function and then hypothesize about the realisation of dynamic principles? The elucidation of the functional organisation of living systems is obviously indispensable for understanding malfunctions, i.e. disease conditions.

Addressing the implications of this approach the workshop's experts discussed the following fundamental aspects of the dynamics of cell function:

### **Model Integration**

In some cases different parts of a model can be treated separately on the basis of their different time scales. For example, this is common in metabolic models where the reactions constituting the internal processes of the enzymes are assumed to be at steady state on the time scale of changes in cellular metabolite pools. This allows the enzymes to be treated as a module with an input-output function. However, this approach cannot be easily applied in many biological systems, for example where the transient behaviour of the fast system has consequences for the behaviour of the slow system, which then feeds back to the faster system after a delay. This may be the case with signal transduction systems, where the signal may be encoded in oscillations (Ca signalling) or transient movements between cellular compartments (NF- $\kappa$ B). Epidermal growth factor signalling causes transients in the MAP kinase cascade on time scales of minutes and relaxes to a new quasi-steady state within an hour. However, important down-stream consequences take much longer to emerge; entry of cells in to the cell cycle and commitment to division requires several hours sustained signalling, whilst growth factor receptor internalization and recycling and altered gene expression alter the concentrations of the components of the signalling system also on time scales of hours.

Therefore merging models of modules with different time scales leads to the introduction of new dynamic interactions that were not present in the sub-models and techniques for developing such multi-scale models need further development by biologists, computer scientists and mathematicians. One technique could be to run the detailed submodel many times to establish its range of behaviour and then to emulate this with a simpler model. However, there are disadvantages if this model reduction loses the mechanistic/physical interpretation of the system structure and parameter values, since it is then more difficult to incorporate perturbations such as the action of a drug that targets a specific component of the module.

Issues of validation also arise, since even if the models of the isolated modules have been separately validated, adjustments will certainly be needed in the merged model so some re-validation will be needed. The same module may need different adjustments when incorporated in different scenarios, so although the concept of re-usable component modules is attractive, it may be difficult to implement in practice. Certainly issues of selection of

appropriate model structures, and resolution of the conflicts between the need to use sparse data for parameterisation whilst retaining some results for model validation are areas where guidance is needed from experts, particularly statisticians.

### **Merging Signalling, Metabolism and Gene Expression**

There is a clear distinction between the technologies which are available for experimental measurement of cell signalling, metabolic or transcriptional mechanisms and this gives rise to operational divisions amongst experimentalists. The development of technology platforms such as proteomics is providing new tools for experimentalists which provide the means for a more integrated view of these concepts to be taken. On the other hand mechanistic models of signalling, transcription and metabolic control are being developed and studied primarily in isolation and there is growing realization that principled integration of these models will be an important capability in systems biology.

The question is how can this be achieved both from a theoretical perspective as well as experimental and cultural. The levels of conceptual detail available to modellers differ across signalling, metabolism and gene expression and this is viewed as a challenge to the successful merging of these concepts. Some are particularly well developed, for example the availability of detailed molecular interaction maps and software tools such as Cell Designer, whereas notions of cross-talk are conceptually not at all well developed and should be considered in full-network behaviour. This lack of consistent conceptual granularity across these areas presents further challenges in this merging. Yet despite this there is a real need, for example, explicitly modelling gene expression mechanisms within signalling pathway models.

One way forward is to develop the idea of a canonical process model describing the very basic components of a sub-cellular process which may take the form of a consistent set of basic dynamic motifs describing e.g. degradation, transcription and translational mechanisms. The development of such canonical components and their study in both autonomous and embedded mode, in the form of libraries, will provide a valuable resource which can guide the overall modelling process. These canonical components will differ significantly based on the levels of abstraction adopted, assumptions made and approximations required.

### **Crossing Levels**

The highly advanced multiscale heart models developed by Noble, Hunter et al probably constitute the best example to illustrate the potential applications of systems biology in medicine. It is important to explore whether some of the methodology developed specifically for heart modelling can be exploited in building integrative models of other biomedical systems. Moreover, lessons could also be learned from other fields, such as climate and weather modelling.

Cancer is a complex, multifactorial, multistep process. Cancer modelling therefore provides a good case-study for the biological, mathematical, computational and technological problems normally encountered when building an integrative model (e.g. coupling processes across

levels of organisation and exploring their interaction).

One of the most exciting/challenging aspects in multiscale modelling is the possibility to explore the impact of interactions between phenomena taking place at different time and length scales). For instance: (1) how do mutations in signalling pathways translate into changes at the tissue level; (2) how do biochemical and biomechanical alterations of the stroma (environment) affect intracellular networks?

Techniques are required for analysing and understanding large, complex multiscale models, and their underlying mathematical properties. It will be necessary to simplify and/or reduce models, and perhaps also to develop approximate emulators of sub-components of the full model system. Some form of modularity will need to be exploited. Model simplifications will enable analysis of model sensitivity and validity that would be otherwise computationally intractable. Similarly, fast simulations will be required for the development of effective algorithms for optimal experimental design.

There will also be a need to reconcile bottom-up mechanistic models with more phenomenological top-down statistical models of full system behavior. Although this is generally true, it is especially important in the multiscale context. It may also be desirable to use statistical models as emulators for model components.

A proper understanding of the importance of spatial effects in models will also be required. When is it OK to ignore sub-cellular spatial effects, and when is the specific 3d arrangement of cells in a tissue non-ignorable.

Stochastic (hybrid) models are likely to be required for effective multiscale modelling. Noise and heterogeneity occurs at many levels, beginning with intrinsic intra-cellular noise at the level of gene expression, through to genetic and environmental differences associated with individual organisms (people?!) in a population. Proper incorporation and propagation of noise and heterogeneity across scales will be necessary for validation and calibration of multiscale models against available experimental data.

European-wide collaboration in multiscale modelling is currently inhibited by a lack of standard language for describing biological models at multiple scales. Although SBML has proved invaluable for the sharing of single-cell continuous deterministic models, use of SBML for stochastic models is still not widespread. More fundamentally, SBML lacks important features such as modularity and support for arrays, which are necessary for the development of large and complex models. Furthermore, SBML is essentially single-cell based, and therefore modelling cell population behaviour is not directly possible. There are plans to include such features in the future, but progress is currently very slow. CellML includes some of these features, but is not widely used within the systems biology community.

### **Do modules exist naturally or are they organising principles?**

Due to the size and complexity of biological systems it will be necessary to construct models of parts of the system that will then be merged at a later stage. Whereas such modules are

important for us to conceptualize biological systems, i.e. identify pathways such as glycolysis and PP-pathway, it is unclear as to whether such modules really “exist” in the system. Thus, we addressed the question whether modules can be identified in biological systems for which separate models can be constructed. Such a module should be separate enough from the rest of the system, where separation can be made on the basis of 1) structural organisation (e.g. organelles separate reactions in the cytosol from those in the organelle), or 2) a functional separation (e.g. reactions involved in different functions such as differentiation, growth, apoptosis, coordination of cell-cell interaction). In addition classic separation techniques such as 3) time scales can be incorporated for identification of modules (e.g. the cell cycle is affected through treatment over hours, signalling pathways are studied over minutes). One specific example of a functional separation that is often used to modularize cellular systems is between transcription/translation and metabolism. Although reactions in each of these modules are linked it is possible to separate them experimentally by adding an inhibitor of translation.

Since a system is defined as linked components, there will not be any modules that are completely separated from the system, i.e. there will always be some links. Thus, the consideration should be, is the module separated enough from the rest of the system. Effects of links between the modules and the rest of the system must be taken into account. This can be done by measuring time traces of the linking metabolites and studying the effects of such changes on the isolated module. Indeed, such an experimental validation could be used as a tool in the identification of modules, i.e. only define modules where it is possible to measure the linking metabolites.

The participants also took the opportunity to give their views on the scientific challenges, priorities, and bottlenecks of systems biology. The results are summarised below:

### **What is your most important/interesting problem/question?**

- Integration of metabolism, signal transduction and gene expression for an experimentally accessible system and study optimisation at variable external conditions.
- Are there principles physiological evolutionary adaptation?
- Do we really understand all the requirements for a cell to be ‘live’, i.e. an autonomous self-generating entity?
- Through which mechanism does a cell adapt and is robust against all kinds of unforeseen perturbations?
- What general principles underlie cell differentiation and development?
- How do multicellular organisms age and what underlies the interspecies differences in metabolic rate and longevity?
- How can we build, simulate and calibrate (against data) timely multiscale, multicellular hybrid stochastic heterogeneous population models of biological systems with direct relevance to human health?
- What is the overall organising and central mechanism(s) for cellular + organ function

from birth, development to death?

- How do we identify a subsystem (pathway, module) to study it in isolation and then integrate models thereof to understand a complex whole (cell, physiology)?

### **Has it to do with metabolism, cell signalling, gene expression?**

- All, Answer not possible if only one strand followed
- All, metabolism
- All, adapt the mechanism through signalling and gene expression
- Gene expression -> cell signalling -> metabolism
- All; signalling followed by gene expression
- Doesn't matter; Cell signalling

### **To which cell function does it relate?**

- All, growth,
- All, adaptation to new environments
- All, robustness of one function e.g. growth
- Differentiation -> cell cycle -> proliferation -> apoptosis -> growth
- All; proliferation and apoptosis key
- All; differentiation, following by cell cycle (proliferation)
- Doesn't matter; cell differentiation

### **Which experimental system would you choose?**

- Unicellular organism; microorganism (*E.coli*), yeast, parasite, microbial ecosystems
- Lymphoid cells(T-cells); Drosophila embryo
- More than one required (e.g. Drosophila, mice) for interspecies comparisons
- Yeast / human, cell line, most relevant
- Doesn't matter, eukaryotes
- Doesn't (shouldn't) matter

### **Which pathway, network or cellular process would be most important?**

- All
- don't know
- Aim is to understand whether there are more important components
- TCR signalling pathway -> Notch, wnt, itih
- Tissue renewal / cellular senescence
- Not clear but cell proliferation, apoptosis and necrosis pathways
- Signalling pathway, e.g. MAPK
- To be decided

### **What methodology/approach would you use for modelling?**

- One that does the job, starting with the simplest; ODEs and up
- Simple <-> detailed model; deterministic & stochastic, all kinetic models
- This probably can't be encompassed by a single model. Parts may be stochastic, parts by ODEs; multi-scale modelling
- Possibly evolve large ODE systems and ultimately a realistic ODE model

Qualitative modelling; logical multi-leveled models, Petri nets -> stochastic extensions of

- Petri nets -> hybrid modelling (HPN, diff. eqs.)
- Stochastic, delay ODEs
- Multicell multiscale hybrid stochastic models
- All; ODE + stochastic
- Whatever (is best)

### **Which technology would be required?**

- All; quantitative measurement of metabolites, time series
- Cell – ecosystem level: metabolomics (small and large scale), flux analysis
- All - but quantitative and time-resolved; quantitative proteomics
- All Omics first; see whether adaptations effect only small parts and then narrow it down if possible
- All; chip-on-chip, chIP on chip, and transcriptome
- Several to cover different levels of organisation; single-cell technology
- All; transcriptomics, proteomics initially, single-cell for refinement, facs etc. for calibration
- All; single cell technology -> proteomics (transcriptomis, quant. blot.)
- All; proteomics

### **What are the mayor hurdles in answering your question?**

- Experimental data of suitable precision
- Long term endeavour that requires tools (experimental & theoretical) at all stages
- Not clear whether we have the right conceptual framework
- Money and men power, don't know where to start
- Time constraints, interdisciplinarity, money
- Cover different levels of organisation experimentally and theoretically
- Standard language for describing multiscale multicell (stochastic) models
- Technology for observing biochemical system dynamics
- Complexity -> interdisciplinarity

### **On Research Programmes**

The participants agreed on recommending the formation of a European DPCF Network which aims at understanding DPCFs. By creating a respected network the network would attract new

people to the area of Medical Systems Biology. A good start of would be an internet platform supporting young investigators by awards and the publication of projects, dissertations and papers. Research Training Network grants would be helpful since current calls are too specific and do not fit the objectives of DPCF research.