Modelling of complex biological systems in the context of genomics: an account of a multidisciplinary thematic seminar held in Montpellier (France) in April 2005

Alain R. Thierry,[1] Francois Kepes,[2,3] Patrick Amar,[2,4] Georgia Barlovatz,[5] Gilles Bernot,[2,6] Marie Beurton-Aimar,[7] Marie Dutreix,[8] Jean-Louis Giavitto,[6] Janine Guespin,[9] Jean-Pierre Mazat,[7] Vic Norris,[2,10] Vincent Schafter,[11] Philippe Tracqui,[12] Christophe Godin[13] and Franck Molina[14,*]

[1] *Laboratoire des Défenses Antivirales et Antitumorales, UMR 5124 CNRS, CC 086, Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier, France.*
[2] *Epigenomics Project / Genopole, 91057 Evry, France.*
[3] *Ateliers de Génomique Cognitive, CNRS UPR 2355, Gif, France.*
[4] *Laboratoire de Recherche en l'Informatique, Université de Paris Sud, Centre d'Orsay, France.*
[5] *Dynamic / INSERM, Créteil, France.*
[6] *Laboratoire de Méthodes Informatiques, CNRS UMR8042 / Genopole, 523 Terrasses de l'Agora 9 000 Evry, France.*
[7] *Labri-INSERM U688, Université Bordeaux 2, 33405 Talence, France.*
[8] *Institut Curie, CNRS UMR 2027, 91405 Orsay, France.*
[9] *Laboratoire de Microbiologie du Froid, Université de Rouen, Faculté des Sciences et Techniques, Université de Rouen, 76821 Mont-Saint-Aignan, France.*
[10] *FRE CNRS 2829, University of Rouen, 76821 Mont Saint Aignan, France*
[11] *Genoscope, CNS Genopole, CP5706 Evry, France*
[12] *CNRS, TIMC/DynaCell, Faculté de Medecine, 38706 La Tronche, France*
[13] *INRIA, CIRAD, INRA, CNRS, Université Montpellier 2, Botanique et Bio-informatique de l'Architecture des Plantes, Boulevard de la Lironde, 34398 Montpellier, France*
[14] *Centre de Pharmacologie et Biotechnologie pour la Santé, UMR 5160 CNRS, Faculté de Pharmacie, 34093 Montpellier, France.*

Introduction: challenges in modelling complex systems
François Houiller, Bernard Pau

Plant morphology: creating plant morphological ideotypes with respect to genetic and environmental factors. Although 3D structure and morphogenetic processes are now numerically simulated, the capacity for capturing both environment and function is less explored. Modelling might make it possible to link plant structures, functions and morphogeneses.

Scrutinizing the integrated function of a plant via a modelling approach is a challenge in plant biology. For instance, better understanding of nitrogen management by plants is of crucial interest when lowering agricultural nitrogen consumption is of concern. There is a need for a global vision for nitrogen management with respect to, for instance, the various plant organs and coupling of nitrogen with carbon. Use of transcriptomic tools is envisaged to implement modelling to different levels of plant function.

Elaboration of strategies for enhancing fruit quality might be facilitated through modelling. Thus, parameters such as organoleptics, wall fruit, and saccharose import would be linked to plant global functions.

There is no experimental plant model for studying the interactions of plants with bio-aggressors. Thus, modelling is needed for linking structure, function and morphogenesis of a plant with respect to bio-aggressor attacks.

Manipulating unicellular organisms is easier, allowing the use of a dynamic model. When considering pluricellular organisms, the metabolic network can be linked to the gene network. However, such types of modelling do not address the global integration of plant functions.

Investment in bioinformatics at the Institut National sur la Recherche Agronomique (INRA) particularly focuses on metabolic, proteomic or genomic representation. However, quantitative modelling is poor. It seems necessary to implement a program where, first, horizontal integration would make sense of the numerous data provided by proteomic, transcriptomic and metabolomic studies. Second, the program should, as well, address vertical integration, for instance through the links between plant organization and systems. Finally, the program would benefit from studying species diversity relative to the modelled species of agronomic interest, ensuring transversal integration.

The challenges and future of the "life complexity" approach in the health domain was introduced by B. Pau. The quest for the infinitely small is not sufficient to address life's complexity, which appears fractal relative to the monomolecular scale. Modelling might help in combining the monomolecular aspects and the complexity of life. For instance, neuroscience should, in addition to synaptic transmission, take into consideration the steric and temporal dimensions of neurones and their organization.

*To whom correspondence should be addressed. Centre de Pharmacologie et Biotechnologie pour la Santé, UMR 5160 CNRS, Faculté de Pharmacie, 34093 Montpellier, France. E-mail: franck.molina@cpbs.univ-montp1.fr

Journal of Biological Physics and Chemistry **6** (2006) 103–107
Received 27 April 2006; *accepted* 2 May 2006

One of the ultimate goals would be to understand the elaboration of thoughts. Furthermore, integration of complexity would facilitate the rate of drug discovery, which at present is very low (typically only one in a hundred successfully tested drugs is ultimately successful).

The big concepts of life have not yet been discovered. Who knows about the origin, the organization and the future of life? The zone of collaboration between disciplines is not sufficient, we must bring together means and structures to address life's complexity and understand life.

Artificial chemistry and the virtual cell
Peter Dittrich, Martyn Amos, Andrew Griffiths

To understand and deal with complex dynamical systems consisting of many components that interact with and produce each other continuously, we need theory. Peter Dittrich introduced a theory of chemical organization as a closed and self-maintaining set of components. This concept allows us to map a complex (reaction) network onto its set of organizations. By looking at a set of molecules instead of a set of states, such an approach provides a new view of the system's organizational structure. One can relate to classical dynamical systems theory by connecting dynamics with the set of organizations. Rather than merely using a network analogy or other convenient approach as a useful technique for clarifying our understanding of complex systems, it may now be possible to harness the power of such systems through computation.

Martyn Amos reviewed several such proposals, focusing on the molecular implementation of fundamental computational elements. Since Adleman's [1] original experiment several other attempts to implement computations using DNA have been reported. However, these experiments, although different in many ways, were all implemented *in vitro*. It may be argued that this approach is sub-optimal and that the advantages of working *in vivo* are numerous. For instance, by reprogramming part of the cellular machinery to our advantage, the DNA program can affect its own execution by the production of proteins. A couple of "single cell" experiments have already demonstrated the feasibility of implementing artificial logical operations using genetic modification. In addition, Savagneau [2] addresses the issue of finding general design principles among microbial genetic circuits. On the other hand, bacteria can act collectively as a group and a large variety of different patterns can arise as a result of the way that individual cells respond to others in their neighbourhood and to the conditions in their environment. This spontaneous bacterial self-assembly as functional patterning is a form of computation that is common in nature but currently poorly understood and highly unexplored in computer science.

As a step forward Andrew Griffiths proposed a way to perform *in vitro* directed evolution by compartmentalization using nano-droplets. This approach allows the physical linkage of genotype and phenotype in artificially designed droplets thanks to a technology providing up to 10^9 droplets/day. Combining a gene library approach with selection for catalysis by binding in an emulsion of droplets, the methodology provides a way to screen enhanced enzymatic activity droplets using fluorescence-activated cell sorting (FACS).

Morphogenesis and development
Brian Goodwin, Pierre Roux, Przemyslaw Prusinkiewicz

The need to introduce meaning into the study of biological processes was emphasized by B. Goodwin. This implies unity between nature and culture, and leads to the notion of responsible science. It was illustrated via intracellular networks and the rôle of geometry in morphogenesis.

Proteomics shows the importance of protein interactions, which follow a power law in the frequency of interactions between different proteins. What does that mean? It may be useful to compare this with human language, where the frequency of words in a text also follows a power law. Language may be seen as a phase transition due to the tension between a communicator and a listener, the meaning arising from the relationship of the communication to the context. In language, words form a self-referential network, as do protein interactions. In this respect protein networks represent a kind of language. The organism itself is the meaning of the genetic text.

Order for free comes from self-organizing processes at different levels, within organisms, between them and in their ecological contexts. Organisms are appropriate forms of living in ecological contexts.

Morphogenesis illustrates how geometry can give meaning to cells. In the simplest metazoans, (*ascidia*, *hydrae*) budding is the major form of reproduction. Cells and tissue deformations (curvature) are coördinated with changes in the activities of genes, morphogens and other molecules. Geometry, coupled to the dynamics of these intracellular networks, can explain morphogenesis. A part of an organism such as a bud, or a zygote in sexual reproduction, undergoes changes of shape and complexity, so that a part becomes a whole.

Dynamic tension between local (molecular networks) and global (geometry), resulting in different morphogenetic attractors, can be seen as the source of creativity in evolution.

In *Acetabularia* molecular dynamics and geometry are interconnected. In a more general theoretical model of morphogenesis, cell shape changes in response to two

adhesion molecules endowed with "inverse" dynamics (a molecular switch). Local diffusion of small molecular adhesion activators and geometry are enough (no need for reaction/diffusion models). The overall size of the organism acts as a bifurcation parameter on the dynamics, changing the spatial pattern of the morphogen concentrations.

Conclusion: nature as culture. The organism has embodied meaning by the development of its own shape. Organisms participate in culture through the use of proto-languages to read their genetic texts, creating embodied meaning in their forms (morphology and behaviour), which are appropriate to their ecological context. But they have 3.7 milliard years, experience in sustainable living, so humanity should take lessons from them to avoid causing disasters while mingling with the complexity of the world.

This presentation raised numerous questions:

1. The concept of meaning in natural systems

Evolutionary claims that consciousness has arisen ('emerged') in humans but is not shared by other species creates a logical dilemma. Consciousness cannot emerge from nothing, but must have an antecedent.

Meaning is context-dependent. For instance in the inner ear of the chicken, cells are spatially ordered, each with a protein sensitive to a different sound. All these proteins are encoded by the same gene. That represents 576 meanings, but each one only has meaning in its context.

The failure of the human genome project results from neglect of the need to take the context of genes into account. Each context requires specific selection among alternatives, a form of perception. The context is not the same for a molecule, a cell, or an organism. One must find the adequate language for an appropriate description at each level.

We are used to the analytical way of understanding; the subjective way is disqualified. Trying to understand the quality of the experience of other beings is not included in our training.

We should adopt a more sympathetic manner of doing science; this might avoid damaging the complex systems on which the quality of our lives ultimately depends.

2. Curvature and its rôle in morphogenesis

Q. There are positive feedback loops where the curvature increases the curvature. Could this also happen inside a cell?

A. The curvature concept may become inadequate, depending on the level of organization. It is appropriate mainly at higher degrees of organization.

Q. What about counterexamples, like cell division, in which curvature plays a major rôle, since the lipid bilayer is quite sensitive to curvature? By preventing lipid transfer, one blocks cell division; long-binding proteins are also sensitive to the curvature

A. This shows that one has to determine the appropriate solution for different processes.

Morphogenesis and cell motility mechanisms were described with respect to the rôle of p53, the well known tumour-suppressing protein (P. Roux). A new rôle has been evidenced for the pleiotropic protein p53, which has been shown to inhibit cell migration when a locally growing primary tumour transforms into an invasive metastasis.

The central dogma of oncogenesis states that cancer is due to mutations. Rho GTPases are the main factors controlling the transformation of adherent cells (such as fibroblasts) into motile cells. However no mutation of the corresponding genes has been evidenced during this process. In contrast, late mutations of p53 have been evidenced in 50% of tumours: p53 displays frequent mutations in human cells, and is known as being involved in cell cycle regulation. Could these mutations be equally responsible for the morphological modifications of transformed cells, by acting on proteins involved in cytoskeletal actin modification? During colorectal tumorigenesis for example, p53 is active on the cytoskeleton, and experiments have evidenced a regulatory cascade controlling the processes involved in cell motility. Thus p53, which was known as the "genome guardian", now also becomes known as an inhibitor of cell motility. If motility is activated in cells with a mutation in p53, the process can be reversed by inhibiting ROCK and the rho GTPases, but this is not possible if p53 is not mutated. This also provides evidence that mitosis and blebbing, which characterize the amoeboid type of migration of the most invasive cells, alternate during the cell cycle.

Questions were raised about the possibility of generalizing the model, and about the advantages of having a structure-centred approach for decision making. Concerning multifactor aspects, it was stressed that the environment further increases the number of factors, mutations leading to an increased impact of the environment on cells. This might have therapeutic consequences, and might lead to multiple therapies that could be less onerous.

Finally it was noted that modelling might be the best way to tackle multiple dynamical factors.

The talk of Przemyslaw Prusinkiewicz investigated the coupling of two major paradigms in morphogenesis.

As illustrated by the notion of an L-system,

introduced by A. Lindenmayer at the end of the 1960s, the development of modular (e.g. multicellular) organisms can be modelled as a *rewriting system*. In this paradigm, the development of the whole organism is defined by the changes of its parts, and leads to a bottom-up approach of morphogenesis (from the parts to the whole). A very detailed model of the development of *Arabidopsis thaliana* was presented to illustrate this point.

A second paradigm was particularly studied by D'Arcy Thompson. The development of an organism is considered as the result of a transformation of space. The organism is put in a coördinate system and then changes in the organism shape are due to deformation of the coördinate system itself. This top-down process was illustrated by several excerpts from D'Arcy Thompson's well-known book "On Growth and Form".

Then P. Prusinkiewicz introduced the idea that both paradigms could be coupled to lead to a powerful theoretical framework for morphogenesis. The growth of a multicellular organism for instance can be viewed from both perspectives: cells make tissues (bottom-up, Lindenmayer viewpoint) and reciprocally, tissue growth makes cells (by making more space available for creating new cells). In this new unified paradigm, a possibly non-linear stretching of space would impose a growth field on the organism, which would then develop by adding locally new modules (e.g. L-system rules) at the places where the stretching is higher. Simple examples were proposed to illustrate this idea (dynamic generation of a Cantor set, generation of an S-shaped branching structure).

The discussion first addressed the similarity between the proposed approach and the zooming into a structure (which also can be seen as a stretching of space and a generation of new components—the details—as one zooms in on an object). The question of whether this approach could give insight into the general question of integrating continuous approaches (à la D'Arcy Thompson) and discrete ones (à la Lindenmayer) was then debated. The possibility of inverse modelling was also discussed: how to design a model of this sort from experimental observation (e.g. in the domain of leaf growth modelling). A similar question was raised regarding the possibility of modelling cell mobility within such a framework.

Function/redundancy/robustness

P. Tracqui, Jacques Dumais

Understanding morphogenesis is of primary importance for applied and basic research. As for plant morphogenesis, modelling this phenomenon provides a new avenue for elucidating the mechanisms involved. Modelling physical interactions between cells in animal tissues was presented by P. Tracqui. It was suggested that strain propagation within tissues can act as a morphogenetic field that causes cell migration and aggregation. This approach was illustrated in the context of *in vitro* angiogenesis. It defines a mechanical paradigm for morphogenesis complementary to that of Turing, based on biochemical interactions. The proposed theoretical framework enables us to better understand the complex and emerging behaviour of interacting cells due to the coupling of different processes (tissue mechanics, cell forces, cell migration). It is interesting to note that this work is part of a new and growing trend in developmental biology suggesting that the rôle of physical forces and strains is of critical importance in the development of organs and vessels.

Complex networks and evolution

A. Wagner, J. Stelling, D. Fell

Networks in biology have a major rôle in advancing our understanding of complex biological systems. We know that natural selection has influenced many features of living organisms at various level of organization. However, we know little about the influence of natural selection on the intermediate level of genetic networks. The influence of natural selection on the structure of molecular networks was discussed in relation to the scale range of organization (A. Wagner). On the other hand, approaches and issues in analysing the functioning of metabolic networks were described, in particular the use of elementary modes for the structural analysis of metabolic networks (A. Fell). An elementary mode is a minimal set of enzymes that can operate at steady state, with all reversible reactions working in the thermodynamically favoured direction. TCA cycle elementary mode analysis was proposed as an example. Such structural modelling of metabolic networks illustrates that they are aspects of metabolism that can be modelled with little information—just the presence or absence of enzymes in the system.

Systems analysis of robustness was proposed (J. Stelling). There are a few mechanisms conferring robustness, such as redundancy, feedback, modularity, hierarchies and protocols (sets of rules managing relationships between modules, for instance). Analysis of robustness could be performed at multiple levels (metabolic networks, circadian oscillators, cell cycle regulation, etc.). Using a combination of methods applied at various levels may ultimately lead to better understanding of the relative importance of redundancy of components vs. of pathways, the modular organization of cellular networks, the rôle of individual feedback

circuits, etc. Investigation on a common set of mechanisms that contribute to robustness in biology and engineering could lead to cellular design principle identification.

Complex systems modelling and health
Randall Thomas, Bernard Korzeniewski

Modelling and simulation of an organ might provide clues for new hypotheses of organ functions and consequently might have repercussions in medicine. S. Randall Thomas reported on several methods of mathematical modelling of some functions of the kidney. The study of an organ function such as renal behaviour illustrates the need for interdisciplinarity between physics and anatomy. Mathematical modelling affords links in particular between the *in vitro* and *in vivo* experimental settings, providing help in interpreting data obtained from *in vitro* experimental models and in elaborating hypothesis based on the *in vivo* functioning at all levels of organ organization. In addition, he described web resources that provide quantitative databases and simulation resources for kidney models. A quantitative kidney knowledge database eases accessibility to a wider audience. Current active efforts to facilitate the development of novel models originating from the creation of open databases of anatomical data and quantitative measurements were presented.

Workshop on the demonstration and comparison of modelling tools

This workshop, demonstrating and comparing tools from common biological models, pointed out the possibility of translating biology into other languages. In particular, it was emphasized that languages are chosen according to the formalism and the pertinence to the given objectives, or are defined following various modelling philosophies.

Acknowledgment

The organizers of the seminar acknowledge the support of the CNRS, the Epigenomics Project, Genopole, INRIA, the Fondation Scientifique Fourmentin-Guilbert and LAMI.

References

1. Adleman et al., *Science* **266** (1994) 1021–1024.
2. Savagneau, *Chaos* **11** (2001) 142–159.