LMS Durham Symposium : 9 - 17 July 2012

Grand Biological Challenges for Mathematicians: From Cells and Microbes to Brains

Organizers: Kasper Peeters (Durham), Anne Taormina (Durham) and Reidun Twarock (York)

Scientific advisers: Paul Chazot, Nigel Robinson, Roy Quinlan (BSI, Durham)

1 Aim and scope of the symposium

The goal of the symposium is to open and consolidate channels of communication between biologists at the forefront of their field, and mathematicians who are either already engaged in various modelling efforts or curious to learn about the challenges faced by biologists at a time when powerful technologies provide a plethora of data and images requiring urgent scientific interpretation. This symposium will cover a wide range of modelling techniques, including computer science, statistics and biophysics as well, and we expect that the interactions with biologists will inspire the development of new mathematical techniques. We anticipate that it will contribute to bridging the culture gap between biologists and mathematicians/physicists unfamiliar with the handling of complex systems, and will attempt to develop a better sense of which mathematical models are either over-simplified to the point of being useless, or over-complex to the point of being intractable.

We envisage the symposium to be biology-driven, and therefore we have invited very high calibre biologists and biochemists to deliver keynote presentations on far reaching challenges posed by recent experimental discoveries in their field of expertise, focusing on those aspects where it is recognised that mathematical modelling would be beneficial and actively sought. These speakers have been chosen for their clear vision of what to expect from mathematical modelling (in terms of predictions) and their ability to communicate with mathematicians, often thanks to their involvement in influential transdisciplinary institutes abroad. A third of the participants will be biologists, and two-thirds will be mathematical modellers (applied mathematicians and theoretical physicists).

The symposium is centred on modelling challenges articulated around three broad themes in the areas of Cell Biology, Virology and Neurobiology.

Theme 1: How the cell works

Theme 1 explores how the cell works through the dynamics of microtubules and actin filaments, and will attempt to set out a mathematical model of regulation of metal availability within the cell through metal-sensing mechanisms.

1. ***Structure, dynamics and function of the cytoskeleton and DNA***

   The cytoskeleton and DNA chains are large molecular complexes which play a crucial role in various functions of living cells.

   DNA is an assembly of nucleic acids which encodes all the proteins that a cell must manufacture to live. DNA interacts with various macromolecular complexes made out of proteins and RNA to fold or unfold DNA and reveal the genes that must be transcribed at any given time. Understanding and modelling how all these macromolecular complexes interact with each other to select and transcribe the correct genes is a very topical issue, which is of primary importance for stem-cell research, amongst others.

   The cytoskeleton is made out of three types of macromolecular assembly of protein complexes that form a scaffold for the cell. The roles of these complexes range from the separation of

chromosomes during cell division, the transport of large cargo within cells, the "crawling" displacement of cells within tissues, to the conversion of chemical energy into mechanical energy in muscles, to name but a few. Modelling the cytoskeleton and the associated motor proteins is of primary importance to cell biology and is directly related to the development of several drugs, including cancer chemotherapies.

Traditionally, biologists adopt a qualitative approach to describing the complex systems they study. Advances in experimental technology makes it now possible to collect enough detailed information and quantitative data to develop mathematical models of many biological processes. Biologists are now able to describe in detail how DNA and the cytoskeleton interact with other complexes, making it possible to develop physical models that can be tested against experiments.

Mathematically the modelling of the cytoskeleton involves Langevin's equation for complex dynamical systems as well as Monte Carlo simulations.

Participants include[1] : Anja Geitmann (plant biologist, Montreal), Fred Mackintosh (theoretical biophysicist, Amsterdam), Michel Peyrard (theoretical biophysicist, ENS Lyon).

2. ***Metalloproteins and metal sensing***

By binding to metals, proteins can catalyse a greater range of chemical reactions than would be possible with carbon, nitrogen, oxygen and sulphur alone. Nearly a half of all enzymes, perhaps a third of all gene products, are estimated to need one or more metal atoms in order to function. Ensuring that the right metal binds to each protein is, literally, an elemental biological challenge. Some metals, such as copper and zinc, have an inherent tendency to form tight complexes with organic molecules such as proteins, while others such as magnesium and manganese tend to form weaker complexes. It could be argued that the greater challenge is ensuring that the wrong metals are excluded from each protein. An emerging view is that this is achieved in cells by limiting the availability of the most competitive metals. By maintaining high (buffered) levels of the least competitive metals but low levels of the most competitive metals, cells counteract the inherent preferences of metals to bind to proteins. In this model, the metal sensor proteins that control these metal levels become especially important. The fidelity of the sensors, their ability to discern the right metal from the wrong one determines metal occupancy of other proteins. Some metals are delivered to the correct destinations by specialised metallochaperones: other types of protein-protein interactions guide metals to preferred destinations. The extent to which the cell biology of each metal is influenced by such kinetic factors remains to be defined.

The field is ripe for mathematical modelling to help define the parameters that determine how each metal-protein partnership arises. How the actions of the metal sensors are integrated such that the right one in a cell's set responds to the right metal is an obvious target for interrogation through mathematics. With metals playing a crucial part in the functional roles of many proteins, breakthroughs in this area will have wide implications across the biosciences and multiple applications throughout biotechnology and biomedical science. It is anticipated that knowledge of how to optimise metal supply to rate-limiting enzymes has the potential to transform the efficiency of biotechnological processes. Also, an understanding of the pathogens' delicate metal-balance may inform the design of new types of anti-microbial compounds. Aberrations in metal supply are also linked to a growing catalogue of pathologies including neurological disorders such as Alzheimer disease. An understanding of how cells help proteins acquire the correct metals is also a pre-requisite to comprehending how these processes fail in ageing and disease.

Although there have been no attempts at designing mathematical models in this area so far, it is anticipated that dynamical systems in and out of equilibrium could play a role. It is one of the aims of the symposium to propose novel mathematical modelling strategies for this biological challenge.

[1] We anticipate about 10 participants in each subtheme

Participants include: Tom O'Halloran (chemist, North Western, Chicago), Nigel Robinson (biochemist, Durham)

Theme 2 : Self-organisation and emergence

Theme 2 investigates how single cells develop in multi-cellular organisms and how proteins self-assemble to form structured entities, in particular in the case of viruses.

1. ***How capsid proteins assemble into viruses***

   An important stage of the viral replication cycle is the formation of the protein containers that encapsulate the genomic material. A better understanding of the mechanisms underlying the assembly of these containers from their protein building blocks is a first step towards the development of new anti-viral strategies that block their formation. Over a decade of work in this area has focused on the development of models and simulation techniques that consider the assembly of these containers from their protein building blocks, mostly in an idealised test tube environment. Whilst these models have provided important insights into virus assembly, they are in most cases too simplistic to account for assembly in the crowded environment of a cell, or of assembly in the presence of other gene products such as the viral genome. Recent work by Stockley and collaborators has shown that far from being a passive passenger as previously assumed, the genomic material of ssRNA viruses can play important cooperative roles during virus assembly. In this part of the workshop, we explore new modelling approaches for these cooperative roles, and in particular for assembly in the cellular environment. For a number of viruses, the assembly of the protein containers is followed by a structural transition that is a prerequisite for infectivity of the particles. The mathematical and biophysical basis of such structural transitions will therefore also be explored in the context of this symposium.

   Mathematical techniques relevant for the modelling of virus structure and assembly include group, graph and knot theory, differential equations based on the law of mass action, molecular dynamics as well as stochastic differential equations and simulations.

   Participants include: Peter Stockley (biochemist, Leeds), Robijn Bruinsma (biophysicist, UCLA)

2. ***Cellular architectures***

   Real-time fluorescence imaging is combined with classical genetics, molecular biology, biochemistry and biophysics to model the dynamics of cellular systems and morphogenesis. The symposium will provide opportunities to discuss strategies for the modelling of eye lens formation in zebrafish and how cell division and morphant gradients contribute to the lens epithelium in the mature lens as a case study. It will benefit from the expertise of several biologists working on vision, and mathematical modellers working on the spatio-temporal dynamics of nonlinear biological systems.

   Interesting modelling tools in this area are discrete Markov chains to describe the emergence of geometric order in proliferating epithelia, reaction-diffusion equations and the cellular Potts model to simulate cell sorting behaviour.

   Participants include: Roy Quinlan (biologist, Durham)

Theme 3 : Neurobiology: from single channels via synapses

Theme 3 looks at some of the current challenges in Neurobiology: the exact molecular nature of the ion channels that occur in the excitable cells of the nervous system, the evolution of synapses that contain these ion channels and their plasticity, as well as the mathematics of neural networks and pattern recognition.

1. ***Structural basis for single channel function***

   The function of the nervous system is based on electrochemical transmission and the molecular basis of this process is an ion channel, residing in a membrane. Ion channels are highly allosteric protein complexes which rapidly open and close in response to local voltage changes or binding of selective agonist ligands. Furthermore, ion channels, like enzymes, can also discriminate between specific 'substrates', namely monovalent cations and anions which diffuse through their respective pores. Without ion selectivity, electrical signal production in biology does not occur. Improvements in numerical methods, based on clues gained from enzymes and other allosteric proteins are required to optimally model experimental single molecule electrophysiological recordings to produce ideal descriptions of single channel function. This will improve the rational development of new pharmacotherapeutics, which target these ion channel targets.

   Participants include: David Colquhoun (Pharmacologist, UCL), Paul Chazot (neuropharmacologist, Durham), Lucia Sivilotti (pharmacologist, UCL)

2. ***Origin and Plasticity of the synapse proteome and phosphoproteome***

   The origins and evolution of synapses, which are the fundamental structural and functional units of the nervous system, are not known to date despite their intense study over several decades. A recent model has been proposed, whereby synapse formation would have evolved in unicellular organisms before other stages in neural development including neuronal migration. This may not be great surprise as unicellular organisms, and the mammalian central nervous system have a fundamental role to sense the local environment and react appropriately to that sensory information. Therefore, the emergence of the synapse would be a crucial step in the origin of the nervous system. This unit, coined the "protosynapse", evolved from comparative genomics and proteomics. Proteomics and phosphoproteomics of the human synapse has generated an impressive array of molecular components, revealing one of the most complex functional systems currently known to cell biology. Clusters of structural, regulatory and signalling molecules have been identified, but how these components interact in a co-ordinated fashion remains unknown. The challenge is to design and develop thoughtful mathematical models to make sense of the vast array of data available, which can then be tested and further optimised via Biological experimentation. Improved iterative mathematical models are required to provide the complete picture depicting how the synapse works, is regulated, and can adapt to change.

   Participants include: Seth Grant (molecular biologist, Sanger Institute Cambridge), Stephen Eglen (computational biology, DAMTP Cambridge)

3. ***Neural networks and the visual cortex***

   Progress in understanding the brain mechanisms underlying vision requires the construction of mathematical models that not only emulate the anatomy of the brain and its physiology, but ultimately match its performance on visual tasks. Increasingly, our understanding of the biological structure of the visual cortex is driving attempts to improve in fundamental ways upon traditional neural networks. Important biological aspects include neuronal plasticity, the organisation of the cortex into strongly horizontally connected layers and vertical minicolumns, the use of sparse representations for information storage and a hierarchical classifier structure. There is evidence that such models can solve new classes of pattern recognition

tasks and improve upon important aspects such as scale and position invariant representation and robustness to noise, at least in as far as the first few hundred milliseconds of visual shape processing are concerned. Connecting these models to the more advanced stages of the vision process, in which eye movement and shifts of attention play an increasing role and cortical backprojections become important, is among the topics of attention in this symposium. A solid mathematical description, and the formulation of a corresponding statistical learning theory, that is able to explain the performance of biologically-inspired computer vision mechanisms is lacking to date.

Participants include: Kristine Krug (neuroscientist, Oxford), Andrew Parker (physiologist, Oxford), James Bednar (computer scientist, Edinburgh), Tomaso Poggio (computer neuroscientist, MIT),

Individual themes have been chosen for their synergies. Indeed, the models for the cell in Theme 1 will further our understanding of virus assembly in the cellular environment in Theme 2, and the topics explored in the first two themes have high relevance for Theme 3, where the focus is on a specific type of cell and architectures, namely neurons and neural networks.

By bringing researchers in these areas together under the umbrella of this symposium, we hope to initiate new collaborations across the boundaries of these themes. Moreover, the mathematical techniques likely to be used to solve problems across the three themes are similar, covering stochastic models and simulations, graph and group theory, statistical and continuum mechanics, applications of PDEs and ODEs, Bayesian statistics and data mining, which will lead to further synergies between different modelling approaches.

2 Format of the symposium

The format we envisage is eight days of scientific activities, with a one-day break at the week-end for social activities. Each day will start with a keynote talk delivered by a biologist who has experience of working in an interdisciplinary environment with mathematicians, and who is able to identify a significant biological challenge that is ripe for mathematical modelling in the broad sense of the word 'mathematical'. Other talks relevant to the topic discussed in the keynote talk will be presented by mathematical modellers, and brainstorming sessions in the afternoon involving biologists and mathematicians will be organised to stimulate interdisciplinary collaborations. Ideally, we would like to gather scientists who can think outside the box and trigger novel approaches to solving biology-inspired questions.