# Statistical Mechanics of Molecular and Cellular Biological Systems

**19 January to 9 July 2004**

**Report from the Organisers: TAJ Duke (Cambridge), TCB McLeish (Leeds), JE Molloy (York), WCK Poon (Edinburgh),
PG Stockley (Leeds), JA Trinick (Leeds)**

---

## Scientific Background

We are currently witnessing a remarkable period of renewed collaboration between physicists, mathematicians and biologists. Not for the first time are there very good reasons for synergy: the beginnings of molecular biology itself emerged from a previous period in which techniques from physics, and mathematical analysis, demonstrated huge potential to answer biological questions. Famous examples are the application of radio-labelling in Copenhagen in the 1930s and of course the determination of the structure of the "Double Helix" of DNA in Cambridge and London in the 1950s by X-ray diffraction, and the subsequent application to the analysis of protein structures.

This programme arose from the very promising creative tension between the "randomness" of statistical physics and the "specialness" of biology. Two aspects of the growth of these two subjects in the last 20 years created the context for it. In physics, the structure and dynamics of "soft matter" has rapidly become a subject in its own right. Techniques from quantum field theory applied to polymers, membranes, liquid crystals and colloids by Edwards, De Gennes and others reveal a subtle world where gentle forces combine with Brownian motion to create emergent properties such as elasticity, membrane interaction and microphase separation. The field has generated important theoretical tools for studying non-equilibrium and self-organised states, and made contact with increasingly sophisticated experimental data. In biology, it is largely the explosion of accessible data at the level of individual biological molecules and cellular processes that create a ripe moment to provide a context for the two communities to work together. After all, at the cellular level, Biology has chosen the very motifs of polymers and membranes that constitute the raw material for soft matter physics. Furthermore, biomolecules, such as the protein whose structure is illustrated in the figure overleaf, are of the scale that subjects them to continual random Brownian motion - the consequence of heat, and the generating dynamics for the whole approach of "Statistical Mechanics".

However, the very deep methodological differences between physics and biology constitute severe challenges to this hopeful setting. As Fox-Keller has recently pointed out, the very words "theory" and "model" mean very different things to a biologist than they do to a theoretical physicist. The essential step of "coarse-graining" or "renormalization" of a physical system that permits theoretical statistical mechanics (the elasticity of flexible polymers is an emergent property of their connectivity, not of their local chemistry) mystifies a molecular biologist who knows very well that every chemical side-group on a protein, every base-pair in a strand of DNA might matter. Yet even the tools used in current molecular biology suggest that sometimes coarse-graining is of use, if not an essential step in understanding.

Recent mathematical theories have begun to shed light on how nature exploits physics in these special molecules. Leibler pioneered physical models of self-assembly and growth of microtubules; Huxley, Jülicher, Duke, Fisher and others simple models for biological molecular motors. Maggs, MacKintosh and Morse have extended coarse-grained models for ordinary polymers to the case of very high local stiffness met with in the case of F-actin and microtubules, while models for the elastic behaviour of DNA at various degrees have been developed by Callaghan, Marko and Siggia, P Nelson, Ou-Yang and others. A rich statistical mechanics for the thermodynamically-driven self-assembly of mis-folded sections of proteins into fibrils and fibres is due to Nyrkova, Semenov and others, and theories of the more finely-tuned self-assembly of viruses has recently been injected with new momentum from Bruinsma, D Nelson and Twarock in the tradition of Caspar and Klug. The organisation and properties of lipid membranes have been addressed by D Nelson, Andelman and others using developments of the celebrated coarse-grained Hamiltonian of Helfrich. The role of electrostatic charge in biomolecular solutions, and its frequently paradoxical effects, has actually stimulated fresh work in physics with a long and current history. Our goal was to stimulate interaction of this new generation of theoretical tools with core scientific programmes in molecular biology, to explore new problems, develop new models, and to move beyond purely reductionist approaches towards an understanding of the systems that the cell’s active and passive components constitute.

Cambridge provides an ideal setting for a programme of this highly multidisciplinary type, since such ready expertise in molecular biology was on the doorstep of the Newton Institute. The goal of maintaining a near-constant presence of experimental biologists was increasingly met during the programme, and several local experts were extremely generous with their time.

## Structure of the Programme

The overall shape of the programme was designed to allow people interested principally in one of the sub-themes, and unable to participate for the full six months, nevertheless to spend several weeks at the Institute, overlapping with at least one other theme. Long-term participants were able to identify a number of coherent approaches and problems between themes. Visits were guided by a series of four themes that built in logical complexity:
- Single Molecule Biophysics
- Molecular Motors
- Membranes
- Statistical Mechanical Aspects of Gene Signalling

In addition, the programme began with a broad scene-setting week and contained several embedded workshops with satellite meetings in Oxford and Edinburgh. The programme attracted 62 long-stay participants (51 from the UK) and 78 short-stay (45 from the UK). The large overall number of shorter-stay visitors represented the special need for this programme to attract biologists with particular expertise to leave their laboratories and keep the longer-term participants honest! Running experimental work meant that the biologists among the organisers (JE Molloy, PG Stockley and JA Trinick) were not resident continuously, but were able to visit at strategic times when the relevant theoreticians were resident.

The 6-month duration was especially suited to this type of programme that attempts to engage two very different communities. There was in several cases sufficient time to suggest, advertise and run short focused meetings and workshops not anticipated at the start of the programme, but which arose spontaneously from early collaborations and conversations between participants. A prime example was the Oxford satellite meeting on viral self-assembly. The suggestion arose during the opening conference, funding was sought and obtained from both EPSRC and LMS, an international panel of speakers was arranged, and the meeting was run all before the end of the programme!
Other short meetings of this type included a 3-day workshop on processive molecular motors and another EPSRC-funded 2-day meeting on the protein-protein interactions in the cellular environment.

At the heart of the programme lay the NATO ASI - in this case run very successfully as a satellite meeting in Edinburgh. Organised by WCK Poon and D Andelman, it took the core theme of coarse-graining in soft-matter biological physics into a pedagogical setting.

The programme developed another characteristic and very useful activity that we had not foreseen at the start: the 1-day thematic workshop. These were much less formal, and responded to the emergence of interesting questions within the programme themes. One or more biologists would introduce the molecular biology of the topic and present some of the unsolved problems from their perspective. There would then be a plenary discussion, typically provoking a rather different set of questions from the physicists. After lunch, the self-selected group of people keen to develop models and calculations around the topic would then gather around blackboards for the afternoon. Focus workshops of this kind were held on the topics of histone tails (led by A Travers), evolutionary thinking in molecular structural biology (led by S Conway Morris, D Bray and S Laughlin), bacterial chemotaxis (led by D Bray), DNA replication and repair (led by R Laskey and a team of Addenbrookes-based researchers), protein crystallisation (led by WCK Poon, J Doye and A Louis) and unstructured proteins (led by K Dunker).

The Rothschild Visiting Professor was D Frenkel (AMOLF), who contributed at the opening of the programme as well as to the NATO School, and who travelled extensively in the UK. His Rothschild lecture on new sampling techniques in Monte-Carlo simulations was a memorable landmark of the programme.

As well as the core programme support from the Newton Institute, the programme was supported additionally by the BBSRC, the MRC and a further special grant from the EPSRC Life Sciences Interface programme.

---

## Workshops and Conferences

## Towards a Predictive Biology?

Workshop, 19-23 January 2004

Organiser: TCB McLeish

The programme started with a remarkable scene-setting conference framed around position-papers given by world leaders in the fields covered by the programme, and covering not only science but also science policy (Julia Goodfellow from BBSRC and Rodney Townsend from the Royal Society of Chemistry) and the undergraduate physics curriculum (P Nelson). In addition, each of the forthcoming themes of the 6-month programme was approached in dedicated sessions. Single-molecule biophysics was presented in the experimental context of magnetic tweezers by V Croquette and theoretically for the protein-folding problem by A Finkelstein and K Dill. Linear motors were discussed experimentally by JA Trinick and theoretically by ME Fisher, and the famous rotational ATP-ase by J Walker. Membrane structures such as lipid segregation and curvature were raised by RH Templar and simple models for ion channels by J-P Hansen. Aspects of systems, networks and signalling were presented bioinformatically by JM Thornton and RM Durbin, and using systems modelling by BA Ogunnaike and D Bray, who launched a considerable effort on chemotaxis during the programme. P Nelson nicely summarised the week as a gathering of “people who wanted to meet each other but didn’t know it!” The topics spawned the first of several discussions of the mathematical “coarse-graining” of problems in biology as contrasted with soft-matter physics generally, where the spatial levels are less coupled.

## Protein Folding

Workshop, 15-19 March 2004

Organisers: TCB McLeish, S Radford and A Fersht

Perhaps the subfield with the longest tradition of a fierce interaction of theory, simulation and experiment is that of the kinetics of protein folding. Two “grand challenge” problems are outstanding:

- How can one predict the folded “tertiary” structure of a protein from its primary sequence of amino acids?
- What are the kinetic pathways and associated rates that characterise the folding dynamics?

The meeting was organised around a structure of two focused sessions each day, introduced for one hour by two contrasting speakers, then followed by a chaired discussion. The nature and investigation of the “transition state ensemble” was a major topic of some controversy, presented by A Fersht, E Paci, E Shackhnovich and D Thirumalai. The new insights made possible by single-molecule mechanical studies were discussed by JM Fernandez and PM Williams. Fast folding and early events, including a discussion of recent “record holders” in folding speed, were presented by WA Eaton and T Kiefhaber, and more extended dynamics on the folding landscape by M Oliveberg and SE Radford. New analytic and computational approaches to the folding problem that attempt to explore the implications of the extremely high effective dimensionality of the search were suggested by D Wales and TCB McLeish. The week was memorably concluded by reflections from CM Dobson, which in a salutary way reinstated the context of protein function around the folding issue. This offered the opportunity to think deeply about the research directions of this field.

## Theoretical and Experimental Understandings of Processive Molecular Motors

Discussion Meeting, 23-24 March 2004

Organisers: JE Molloy, TAJ Duke, RA Cross and TCB McLeish

This two-day meeting, part-funded by Hewlett-Packard, addressed our current understanding of the molecular mechanisms of processive motors. The aim was to catalyse a dialogue between experimentalists and theorists interested in this fast-moving scientific area.

Processive molecular motors are critical for a wide variety of life functions, and recent advances in experimental techniques and theoretical approaches have contributed greatly to our understanding of their molecular mechanism; e.g., experimental observations and theory on the stepping mechanism of myosin V. There is a pressing need to define new, experimentally testable models and ideas for how processive motors "walk" along their cellular tracks.

The meeting brought together a large number of UK and also several international experts in the field. The bulk of the delegates gave short 5-minute presentations describing their recent work. This was followed by discussion of each presentation. The meeting then divided into break-out groups which consisted of a balance of Experimentalists and Theorists. These teams addressed a set of detailed questions and reported back in a general session. The meeting ran over two days and allowed sufficient time for informal discussion around the chalk-boards and coffee machines. The meeting was generally accepted to have been successful and to have fulfilled its original aims.

**Soft Condensed Matter Physics in Molecular and Cell Biology, NATO ASI & SUSSP 59 Satellite Conference at the University of Edinburgh,**
29 March-8 April 2004
Organisers: WCK Poon, D Andelman, F MacKintosh and TCB McLeish
This Advanced Study Institute formed the 59th Scottish Universities Summer School in Physics (SUSSP). Its key objective was to introduce participants to the way 'coarse-graining' (a judicious neglect of details) can solve old biological problems and point to new, fruitful avenues of interdisciplinary investigation. Our method was to examine in some detail the application of recent concepts and techniques from 'soft condensed matter physics' - the study of colloids, polymers and amphiphiles - to the investigation of biomolecules and cells.

The opening lecture by WCK Poon examined how 'coarse-graining' at various levels could give information on the structure and function of a typical globular protein, lysozyme. The rest of the lectures were divided into three groups. First, there was a set of lectures introducing soft condensed matter physics: D Frenkel on colloids, PB Warren on polymers, PD Olmsted on amphiphiles, D Andelman on electrostatic effects and TCB Mcleish on barrier crossing, all emphasizing biological applications.

The second set of lectures dealt with specific topics in biological physics. These ranged from rather detailed treatments of specific biomolecular systems (M Rief on molecular motors and DB Bensimon on DNA and DNA enzymes) through rather generic discussions (such as R Podgornik on electrostatic effects, F MacKintosh on stiff biopolymers, R Elber on protein evolution and JC Smith on protein dynamics) to the movement of whole cells (J-F Joanny).

A final set of lectures was focused on experimental techniques: CF Schmidt on optical tweezers, C Seidel on fluorescence single molecule measurements and SU Egelhaaf on scattering.

In their end-of-workshop Questionnaire, students gave an average of 4.1/5 for the quality of the lectures; in answer to the question on whether they had benefited from the programme as a whole, they gave the same average score (4.1/5). We also asked students to name the 'scientific highlights' for them. It is interesting that many of them singled out the general introductory lectures to soft condensed matter physics in answer to this question. Most students contributed research posters, two of which were awarded prizes. All of the lectures are now available to the participants in electronic form on the internet.

**Mathematical Virology, Satellite workshop at the University of Oxford,**
14-17 June 2004
Organisers: P Maini, R Twarock, P Stockley and TCB McLeish
The focus of this workshop was on mathematical models for vital stages of the viral life cycle. Particular emphasis was placed on the structure and assembly of viral capsids, i.e., protein shells encapsulating the viral genome, and the packaging of the viral genome inside the capsids. New approaches from symmetry, tiling and elasticity (SR Twarock, P Nelson and RF Bruinsma) contrasted with specific protein-interaction approaches (RW Hendrix, JA King and DI Stuart). The role of dynamics was a central theme (CL Brooks, D Wales and IF Thorpe). The workshop took place at the Mathematical Institute of the University of Oxford with funding from EPSRC, the LMS and the Isaac Newton Institute. It combined a broad spectrum of international participants from mathematics, mathematical physics, biophysics and biology in an interdisciplinary working environment, and created awareness of open problems and challenges in mathematical virology. The proceedings of the workshop will be published as a special issue of the Journal for Theoretical Medicine, edited by TCB McLeish, P Stockley and SR Twarock, as an interdisciplinary state-of-the art account of the field.

**Protein-Protein Interactions in Vitro and in Vivo**
Workshop, 21-23 June 2004
Organisers: J Doye, A Louis, R Sear and M Vendruscolo
This workshop was separately funded by EPSRC, which enabled us to invite 6 international speakers and to pay the expenses of the UK-based participants. In all, 93 attended. Most of the speakers, and those participants who were not members of the host programme, were biologists. The resulting mix of biologists with the physical scientists attending the programme led to real dialogue across the biology-physics divide. The talks dealt with most aspects of protein-protein interactions, from the damaging, such as the interactions causing the formation of amyloid fibrils (CM Dobson and AL Fink), to the exquisitely engineered interactions that allow highly specific protein interactions (JM Thornton, SA Teichmann, J Janin, JL Harden and L Pearl). Also, models for the highly-crowded interactions both in vitro and in vivo were discussed (AP Minton), as well as their effect on important processes such as protein crystallisation. The ubiquitous question of applicability of coarse-grained models appeared again (SA Deem and TM Truskett). The workshop contributed to three of the four themes of the host programme: protein interactions are central to self-assembly in biological systems, to gene expression, and of course they are what is measured in single-molecule measurements of proteins.

---

## Outcome and Achievements

The programme's central aim was to start radically new approaches to research across biology, mathematics and physics in the four main themes of the programme. A successful outcome from such a goal will be very long-term by its nature, but there were several real achievements made during the six months and there are already signs that the new research collaborations begun are yielding fruit. A dedicated issue of an interdisciplinary journal is in preparation for 2005, as are the proceedings from the workshop Mathematical Virology.

One unexpected and pleasing aspect of the programme was the congenial relationship we enjoyed with the parallel programme on Random Matrix Approaches in Number Theory. One of our participants (R Sear) actually developed a random matrix approach to protein phase separation in the cell, while another (ME Fisher) spent some time in discussion with the parallel programme on the use of the Yang-Lee theory for ferromagnets.
Another “unorthodox” but creative partnership between Z-C Ou-Yang and P Nelson is leading to the translation (and publication) of Nelson’s book Biological Physics: Energy, Information and Life into Chinese.

Many collaborations have begun, and some results are already available in detail. At least 30 current preprints or publications contain work done during the programme. We identify some of these outcomes by theme in the following, extending the first category to include aggregates, self-assembly and complex fluids of biomolecules.

### Single Molecule Biophysics and Self-assembled Biomolecules

RF Bruinsma began a new formulation of the problem of structure and kinetics of “b-amyloid” self-assembled fibres after discussions with JL Harden, ME Fisher, R Sear and TCB McLeish. JL Harden and R Sear began another novel approach to the problem from an evolutionary standpoint. RH Colby made critical progress on the synovial fluid system of reversibly aggregated proteins. H Flyvberg, TCB McLeish and BS Khatri sharpened understanding of their complementary work on analysing the Brownian power-spectra from biomolecules in optical tweezers or atomic force microscopes. R Golestanian and TB Liverpool derived a successful theory for the phase behaviour of DNA and actin in the presence of multivalent counterions, and with F Mohammad-Rafiee made a prediction for the under-twisting of condensed actin. S Licher, TB Liverpool and AJ Levine derived a new model for non-linear protein elasticity. Motivated by discussion with M Oliveberg, TCB McLeish found a way to calculate non-native contributions to the curvature of protein folding “chevron plots” from coarse-grained models. WCK Poon and D Dryden solved a long-standing problem in the slow collapse of DNA under flow in the presence of multivalent counterions. R Sear completed a model on protein-protein interactions. RS Twarock completed a model of tubular structures in (papova) virus assembly.

### Molecular Motors

H Flyvberg and J Taylor initiated a collaboration on the F0 motor. Collaborations on microscopic models for motor dynamics began between T Golestanian and TB Liverpool, and on the consequences of Brownian searches in landscape models between TCB McLeish and JA Trinick. A Simha, TAJ Duke, A Vilfan and D Andor developed a model for active travelling waves in the cochlea, where the cilia are coupled hydrodynamically.

### Membranes

D Andelman, PD Olmsted and S Komura completed a model for the formation of lipid “rafts” by phase separation within the membrane, now in press. AL Frischknecht completed a calculation on the lateral pressure in a bilayer membrane, prompted by discussions with RH Templar, S Mayor and V Gordon. JL Harden, F MacKintosh and PD Olmsted extended work on budding in membranes. WCK Poon and V Gordon were able to rationalise a large data set on binary lipid vesicles after wide discussion with programme participants in terms of tilted domains.

### Gene Regulation and Signalling

D Dryden began a collaboration with JA Trinick to achieve the first high-resolution electron microscopy images of a type I restriction enzyme, and with H Flyvberg derived an analytical model for specific binding of a protein to DNA in the presence of a non-specific binder. ME Fisher and M Howard made progress on proving the existence of a limit cycle in the governing reaction-diffusion equations that M Howard had derived for positioning the mid-cell division plane. D Frenkel continued to develop novel very rapid Monte- Carlo sampling methods and an approach to agent-based modelling. WCK Poon, II Potemkin and TCB McLeish formulated an initial theory for the role of histone tails in control of nucleosome structure, stimulated by discussions of experiments with AA Travers. RJ Hawkins and TCB McLeish completed a calculation of entropy-driven allostery in the lac repressor and formulated extensions to this theory to treat the chemotaxis receptor cluster (with D Bray and TAJ Duke), the met repressor (with PG Stockley) and the dynein coiled coil (with PJ Knight). M Howard and P ten Wolde developed an approach to pattern formation in Drosophila. BA Ogunnaike completed work on the control loop in DNA damage repair, and began work on coarse-graining differential systems.

## Conclusion

So much for the “trees”. The success of the programme in regard to detailed progress is clear. Of rather more importance in this case is the “wood”. Several long-term participants wrote that their stay had sharpened for them how statistical physics might and might not play a role in biology. Several said that they had come to very useful conclusions about what possible research programmes they were not now going to follow! During the six months there was a growing realisation of just how different the approaches of physics and biology are: the questions asked, the definition of a “solved problem”, the role of theory, what a “predictive biology” might look like. It was very interesting and gratifying to hear time and again how the visiting biologists found the questions asked of them quite different to those from their own communities, often of course naïve, yet apparently refreshing and stimulating. “We need your approach as well” was often the conclusion of interdisciplinary discussions. There were some huge surprises, such as the neglect of evolutionary thinking and discussion in molecular biology (we were expecting the opposite). This gave rise to a focus day on that topic, and to several new theoretical approaches. Above all, the unpressured time and repeated exposure that the context of the Newton Institute gave to the programme illustrated uniquely how much mutual listening and explaining needs to be done, as well as contributing greatly to the growing realisation that it is going to be worth it.