

European Molecular Biology Laboratory
Research at a Glance 2005/2006

For more information, please contact:

Office of Information and Public Affairs
EMBL-Heidelberg
Meyerhofstrasse 1
69117 Heidelberg
Germany
Tel. +49 6221 387 355, Fax +49 6221 387 525
<http://www.embl.org>
info@embl.de

EMBL-Grenoble
6, rue Jules Horowitz, BP181
38042 Grenoble, Cedex 9
France
Tel. +33 4 76 20 72 69, Fax +33 4 76 20 71 99

EMBL's member states are:

Austria, Belgium, Denmark, Finland,
France, Germany, Greece, Iceland,
Ireland, Israel, Italy, the Netherlands,
Norway, Portugal, Spain, Sweden,
Switzerland and the United Kingdom

EMBL-Hamburg
c/o DESY
Notkestrasse 85
22603 Hamburg
Germany
Tel. +49 40 89 90 20, Fax +49 40 89 90 21 04

EMBL-European Bioinformatics Institute
Wellcome Trust Genome Campus, Hinxton
Cambridge CB10 1SD
United Kingdom
Tel. +44 1223 494444, Fax +44 1223 494468

Texts:

EMBL Group and Team Leaders

Layout and editing:

EMBL Office of Information and Public Affairs

Cover Design:

Vienna Leigh

Exposure and printing:

ColorDruck Kurt Weber GmbH, Leimen

EMBL-Monterotondo Research Programme
Adriano Buzzati-Traverso Campus
Via Ramarini, 32
00016 Monterotondo (Rome)
Italy
Tel. +39 06 90091285, Fax +39 06 90091272

European Molecular Biology Laboratory

Heidelberg • Grenoble • Hamburg • Hinxton • Monterotondo

Contents

Foreword by EMBL's Director-General	5
EMBL-Heidelberg, Germany	6
Cell Biology and Biophysics Unit	7
Developmental Biology Unit	20
Gene Expression Unit	29
Structural and Computational Biology Unit	40
Directors' Research	53
Core Facilities	56
EMBL-Grenoble, France	65
EMBL-Hamburg, Germany	72
EBI-Hinxton, England	81
EMBL-Monterotondo, Italy	95
Index	102

EMBL: A flagship for 21st-century biology

More than three decades ago, Europe's scientists encouraged their nations to create a unique international centre for life sciences research: the European Molecular Biology Laboratory. EMBL's missions are to conduct cutting-edge research across a broad spectrum of basic themes, to provide vital services to European scientists, to give advanced training to researchers and others, to develop new instruments and methods and to help move ideas and innovations smoothly into the marketplace to benefit many more scientists and society.

What happens at the five EMBL sites (a central laboratory in Heidelberg, Germany, with Outstations in Grenoble, Hamburg, Hinxton (near Cambridge) and Monterotondo (near Rome)) has a deep and lasting impact on the wider community. Most of our researchers leave to assume key positions in the member states. When they go, they export the unique experience of working in a very energetic, international and interdisciplinary environment. Highly-recognised courses and conferences, technology transfer and an extensive outreach programme ensure that know-how spreads further and involves the public in a dialogue about the impact of history's fastest-moving science.

This document gives a concise overview of the work of our research groups and core facilities. EMBL science covers themes ranging from studies of single molecules to an understanding of how they work together in complex systems to create structures and organisms. Scientists at EMBL have great intellectual freedom to pursue the topics that most interest them. Research is loosely structured under thematic Units, complemented by interdisciplinary "centres" and a growing number of joint appointments between Units. Increasingly, our young scientists come from physics, chemistry, mathematics and the computer sciences – bringing in expertise that is essential in understanding complex biological systems. Freedom, flexibility and a regular turnover of staff allow EMBL to pursue the most exciting themes in molecular biology as they arise.

Coping with the dramatic pace of today's biology requires a critical mass of expertise and resources and organisational flexibility. This has been essential over the last three decades and will be even more crucial in the future. EMBL's success has repeatedly been demonstrated through the impact of the Laboratory's scientific work, the quality of its services and its continued ability to attract world-leading scientists to Europe. It is well poised to be a flagship for biology in the new century.

Iain Mattaj
EMBL Director-General

EMBL-Heidelberg, Germany

A city of about 130,000 inhabitants, Heidelberg is home to Germany's oldest university, as well as leading biological and medical science institutes such as the Centre for Molecular Biology, the German Cancer Research Center (DKFZ) and the Max Planck Institute for Medical Research, making it an ideal site for EMBL's Main Laboratory. Nestling in the wooded hills above the city, the complex is home to five of EMBL's scientific Units, namely gene expression, cell biology and biophysics, developmental biology, structural and computational biology and directors' research, as well as core facilities and the central administration, from which service functions are provided for the use of staff working at any of the five EMBL sites.

Today more than 800 staff members are located at EMBL-Heidelberg, and the close proximity of the other excellent institutes has led to numerous long-term collaborations in scientific projects, training and conferences.

EMBL shares a campus with its sister organisation, the European Molecular Biology Organization. The two share strong historical ties and work together in many ways; for example, to carry out many highly-recognised international courses and conferences.

Heidelberg is also home to EMBLEM, the Laboratory's technology transfer company. Other important central functions can be found on the Heidelberg campus, such as the offices of the International PhD Programme and the Science and Society Programme.



Cell Biology and Biophysics Unit

The cell is the basic unit of life. In terms of scale, this is not surprising, since the cell occupies the precise midpoint between molecular and macroscopic scales. Thus in order to understand how organisms are built and how they function, we need to identify the molecular mechanisms and physical principles that give rise to cellular organisation.

All cells (including prokaryotes) are divided into functional domains, each with different molecular compositions. In addition, eukaryotes have compartments such as the nucleus, the cytoplasm and the secretory pathway. These compartments are permanently renewed by freshly synthesised molecules. They are usually made in one location and have to be delivered to their destination inside or outside the cell by mechanisms that are still poorly understood. Research in the Cell Biology and Biophysics Unit focuses on the mechanisms that permit the organisation of the different compartments in three dimensions and the distribution of specific molecules to the endoplasmic reticulum, the nuclear envelope and nucleus, the Golgi complex, the lysosomes and the cell surface. Molecules find their way to such destinations thanks to stereo-specific self-recognition processes, complex transport mechanisms, self-organisation processes involving energy dissipation and regulatory networks.

Cell biologists and physicists at EMBL are therefore trying to define the role of targeting events, as well as that of more complex mechanisms involving the cytoskeleton in organising cellular space. Novel developments in light microscopy have made it possible to observe what goes on inside living cells. One can follow how membrane carriers move along microtubules from one organelle to another. The cytoplasm is crowded with filaments of different types which not only orient movement but which also undergo continuous rearrangements. As a cell enters mitosis, all the microtubules suddenly depolymerize to reassemble into the mitotic spindle. At the same time, the nuclear envelope disappears, the Golgi complex fragments and membrane traffic ceases. How all these changes are regulated is presently an enigma. During development, when progenitor cells differentiate into new cell types, not only do the daughter cells receive a complement of chromosomes and organelles from the parent cell, but the genetic program is also changed. This change in gene expression leads to a coordinated shift in cell function and structure. New proteins are synthesised in the cytoplasm and are distributed to their sites of function inside and outside the cell. A reorganisation of cellular architecture takes place, guided by rules that are as yet totally unknown. The elucidation of these principles is a major challenge to contemporary biology.

This is the challenge that the Cell Biology and Biophysics Unit is pursuing, both thematically and methodologically. We are developing a wide range of experimental themes within the Unit itself and attempting to integrate these into a functional whole. The four areas that we are presently concentrating on are membrane trafficking, cytoskeletal networks, the cell nucleus and the cell cycle. Together, these themes provide a comprehensive view of cell organisation in time and space. Physicists working together with cell biologists are trying to elucidate the fundamental rules that govern the organisation of these compartments, their dynamics and their function while developing new instruments and tools.



Eric Karsenti

PhD 1979, University of Paris, France.
Postdoctoral research at University of California, San Francisco, USA.
Group Leader at EMBL since 1985.
Unit Coordinator since 1998.

The role of microtubules and phosphorylation reactions in cellular morphogenesis

Previous and current research

Microtubules are key players in the organisation of cellular space. They are involved in the positioning of organelles during interphase, in the polarised transport of vesicles in differentiated cells and, generally, in cellular morphogenesis. During mitosis, their organisation into a bipolar spindle allows chromosomes to segregate into two daughter cells.

In recent years, we have been working on the mechanism of mitotic spindle assembly, which we have used as an interesting example of morphogenesis at the cellular level. Using frog egg extracts and synthetic chromosomes made of magnetic beads coated with DNA, we have shown that a partly functional bipolar spindle can form around these artificial chromosomes in the total absence of kinetochores and centrosomes. We believe that bipolarity results from the following basic principles: (1) mitotic chromatin introduces asymmetry into the mitotic cytoplasm that results in the preferential stabilisation of microtubules around chromosomes; (2) once microtubules start to grow in random orientation around the artificial chromosomes, motors present in the cytoplasm and on chromosomes organize the microtubules into a bipolar array (this is a self-organising process) (3) the steady-state length of the spindle depends both on the regulated dynamics of spindle microtubules and on the movements of microtubules inside the spindle, which are powered by the activity of motors. The preferential stabilisation of microtubules around chromosomes seems to be mediated by phosphorylation/dephosphorylation reactions of microtubule associated proteins (MAPS) and of microtubule destabilizing proteins.

Future projects and goals

Our goal is to understand the fundamental principles involved in microtubule dependent morphogenetic events in the cell. To achieve this goal, we plan to investigate the nature of the effect of chromatin on microtubule dynamics during mitosis, to identify microtubule regulators involved in spindle assembly and to investigate how phosphorylation reactions participate in the morphogenesis of the mitotic spindle. We will also develop new cytoskeleton-dependent morphogenetic systems.

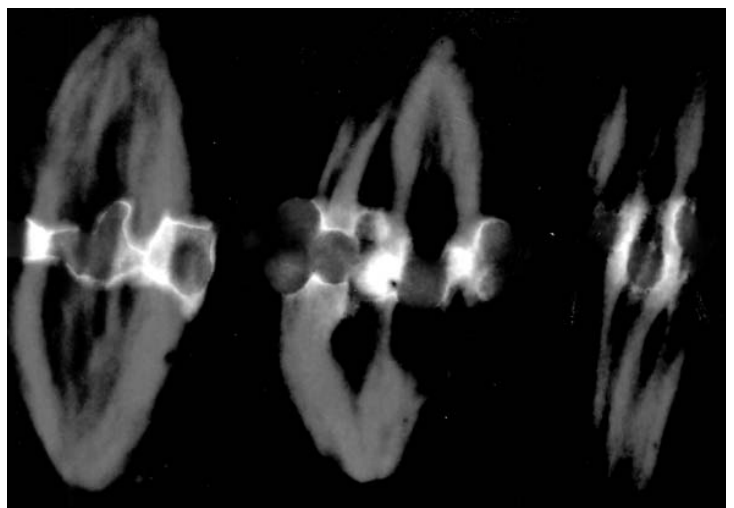
Selected references

Caudron, M., Bunt, G., Bastiaens, P. & Karsenti, E. (2005). Spatial coordination of spindle assembly by chromosome-mediated signalling gradient. *Science*, in press

Niethammer, P., Bastiaens, P. & Karsenti, E. (2004). Stathmin-tubulin interaction gradients in motile and mitotic cells. *Science*, 303, 1862-1866

Nédélec, F., Surrey, T. & Karsenti, E. (2003). Self-organisation and forces in the microtubule cytoskeleton. *Curr. Opin. Cell Biol.*, 15, 118-124

Left: a spindle assembled around chromosomes; middle: a spindle assembled around chromatin-coated beads in a *Xenopus* egg extract; right: microtubules are unable to form spindles in the absence of molecular motors.



Claude Antony

PhD 1984, University Paris VI, France.
Postdoctoral research at EMBL.
Group Leader at CNRS 1994-2003.
Facility Head and Team Leader at EMBL since 2003.



Electron microscopy of microtubules in yeast cells

Previous and current research

Yeast cells are ideally suited for Electron Microscopy (EM) of microtubule arrays because the cells are small and are a relatively simple system to study compared to higher eukaryotic cells. In the budding yeast *Saccharomyces cerevisiae*, kinetochore function is essential to regulate the assembly of the mitotic spindle regarding chromosome attachment and segregation. Some mutations in kinetochore proteins can dramatically alter the organisation of the spindle and can explain chromosome segregation.

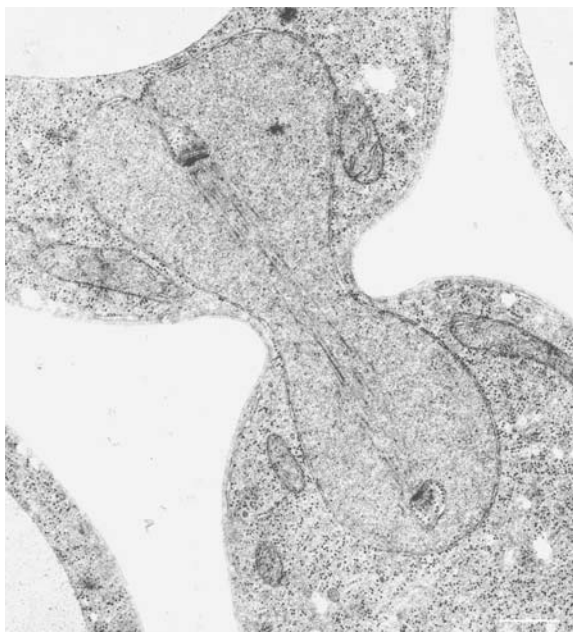
The fission yeast *Schizosaccharomyces pombe* is also an excellent model to study at EM level. The polarised organisation of its cytoplasm is supported by interphase microtubule bundles.

Future projects and goals

Our project is to solve the organisation of microtubule bundles in the fission yeast. In particular, we want to understand the microtubule organisation within the bundles and their polarity in wild type and mutant strains. For this purpose, we are developing electron tomography (ET) analysis of *S. pombe* microtubules. Our aim is to build synthetic 3D models in which all organelles and their various connections can be visualised precisely.

This kind of approach will also be applied to budding yeast mutants displaying aberrant spindles, otherwise difficult to analyse using serial sections. Moreover, we work at developing correlative microscopy (immunofluorescence/EM) by applying photoconversion methods to budding yeast expressing GFP-tagged proteins. We hope to find a way to localise more kinetochore proteins during the cell cycle, possibly using ET.

In the long term, solving in situ microtubule interactions with the various cellular elements would be best achieved in vitreous ice and cryo EM.



Budding yeast mitotic spindle in anaphase (cryofixed cell by high pressure freezing).

Selected references

- Johannes, L., Pezo, V., Mallard, F., Tenza, D., Wiltz, A., Saint-Pol, A., Helft, J., Antony, C. & Benaroch, P. (2003). Effects of HIV-1 Nef on retrograde transport from the plasma membrane to the endoplasmic reticulum. *Traffic*, 4, 323-332
- Müller-Reichert, T., Sassoon, I., O'Toole, E., Romao, M., Ashford, A.J., Hyman, A.A. & Antony, C. (2003). Analysis of the distribution of the kinetochore protein Ndc10p in *Saccharomyces cerevisiae* using 3-D modelling of mitotic spindles. *Chromosoma*, 111, 417-428
- Arrastua, L., San Sebastian, E., Quincoces, A.F., Antony, C. & Ugalde, U. (2003). In vitro fusion between *Saccharomyces cerevisiae* secretory vesicles and cytoplasmic-side-out plasma membrane vesicles. *Biochem. J.*, 370, 641-649
- Mallard, F., Tang, B.L., Galli, T., Tenza, D., Saint-Pol, A., Yue, X., Antony, C., Hong, W., Goud, B. & Johannes, L. (2002). Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform. *J. Cell Biol.*, 156, 653-664



Philippe Bastiaens

PhD 1992, Wageningen University, the Netherlands.

Postdoctoral research at the Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.

Group leader at the Imperial Cancer Research Fund, London, UK.

Group Leader at EMBL since 2000.

Imaging signal transduction processes in single cells

Previous and current research

The objective of the research group is to understand: (1) reaction-diffusion properties of protein reaction networks that generate spatial patterns of protein reaction states in cells, and (2) how these reaction patterns regulate cellular signal transduction and morphogenesis. We therefore develop optical microscopic approaches to image elementary protein reactions, such as interactions, post-translational modifications and conformational changes within the intact cell, thereby maintaining network interconnectivity and spatial organisation. The information on the reaction state of fluorescently tagged bio-molecules is obtained by exploiting the photophysical properties of fluorescence using microspectroscopic techniques. For example, by imaging fluorescence resonance energy transfer (FRET) between fluorescent tags on proteins, molecular interactions or conformations can be mapped with sub-micron resolution. In this way, information on the spatial distribution of signalling molecules (micrometer scale) is combined with their reaction state (nm scale). We have used this approach to show that the cell can generate a local cytoplasmic environment in interphase and mitosis that can orient the microtubule cytoskeleton by the generation of a steady-state phosphorylation gradient of the microtubule regulator stathmin. By recursion between optical experiment and modelling we try to explore the properties of complex multi-component signalling systems. This approach was, for example, applied to study how the interplay between receptor tyrosine kinase (RTK) and protein tyrosine phosphatase (PTP) activities regulate signal initiation, propagation and termination.

Future projects and goals

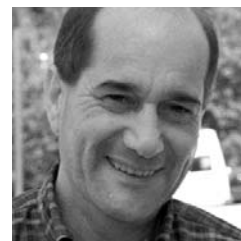
We are studying topology of reaction networks in signal transduction in terms of where in the cell reactions happen as well as the response properties that emerge from the connectivity between reactions. In order to study the spatial organisation of protein activity in cells, optical assays that report on a single protein reaction state are often sufficient. In this context we are further investigating the spatial regulation of growth factor signalling systems using genetically encoded fluorescent sensor systems reporting on growth factor receptor activation to downstream signalling events such as ras or Erk activation. Here, the objective is to understand how the cell maintains heterogeneous distributions of activities and how this affects response and signal specificity. Since the (spatial) regulation of growth factor receptor activity by protein tyrosine phosphatase dictates signal strength and duration, generic optical approaches are being developed for imaging the spatial regulation of protein tyrosine phosphatase and kinase activity. The other goal is to identify networks modules and derive their function (response) from the connectivity (topology) between reactions. In theory, network topology can be derived from systematic perturbation of activities on large sets of proteins, and subsequent read out on multiple reaction states at steady-state. The predicted connectivity in networks can then be tested by imaging coupled reaction kinetics by multiple optical sensors that are simultaneously detected in one cell. So far, the low multiplexing ability of functional fluorescence microscopy has limited us to the detection of a single protein reaction per cell. We have thus started to develop optical proteomic approaches by which we can measure the reaction states of large sets of proteins quantitatively in cell arrays and sample the reaction dynamics of at least two protein activities in the same cell.

Selected references

Reynolds, A.R., Tischer, C., Verveer, P.J., Rocks, O. & Bastiaens, P.I. (2003). EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation. *Nat. Cell Biol.*, 5, 447-453

Christian Boulin

PhD 1976, University Louis Pasteur, Strasbourg, France.
Group Leader at EMBL since 1976.
Coordinator of the Scientific Services
and Core Facilities Unit since 2001.



Microcomputing and data acquisition

Previous and current research

The Microcomputing and Data Acquisition Group concentrates its activities in three areas: (1) development of detector electronics for high speed time-resolved synchrotron radiation scattering experiments; (2) use of microstructures to carry out work on single cells; and (3) development of image processing methods for biological applications.

For fast synchrotron radiation instrumentation developments we rely on collaborations both with our Grenoble and Hamburg Outstations and with CERN. We are currently working on a new fast wire-per-wire detector and its associated readout electronics.

Work on the miniaturisation of biochemical assays includes the establishment of a capillary electrophoresis system with a suitable on-column detector. A method for handling and injecting very small biological samples will also be developed. This will directly address the present biological problem of capturing and detecting very small quantities of DNA-protein complexes.

Alternative assays will be investigated, as well as on-column reactions. At this stage it will be possible to specify a microfabricated device that will efficiently implement a broad range of useful functions.

Image processing is a tool that has established itself in modern molecular biology. Ongoing activities already help in addressing biological problems related to the quantitative analysis and the comparison of images. The automatic detection and counting of particles for immunomicroscopy images using reconstruction of an orthogonal wavelet decomposition has proved to be very successful. The analysis of video-microscopy image sequences has reached a stage where the analysis of movement is reliable.

Future projects and goals

In general our research is geared toward developing new tools and methods to help molecular biologists obtain quantitative data from their experiments.

New technologies such as silicon-pixelated detectors will be tested, with particular emphasis on the development of fast front-end electronics required to read out the detectors.

We will be working to adapt and refine existing capillary electrophoresis and cell and organelle manipulation techniques for use in our particular applications.

Further developments of innovative image processing algorithms will concern time sequence analysis, motion analysis, shape analysis and 3D-image analysis. In particular, the automation of the image analysis in 2D and 3D of microtubules and cell traffic will be an important part of our future work.

Selected references

- Epstein, A., Baumlín, P. & Boulin, C. (1998). A RISC based SCSI interface for a protein crystallography detector. *IEEE Trans. on Nuclear Science*, 45, 1934-1936
- Epstein, A. & Boulin, C. (1998). A fast position encoding system for delay line based gas filled area detector. *IEEE Trans. on Nuclear Science*, 45, 1931-1933
- Briquet-Laugier, F., Boulin C. & Olivo-Marin, J.-C. (1998). Analysis of moving biological objects in video microscopy sequences. *Proceedings of SPIE*, 3642, 1-11
- Herr, S., Bastian, T., Pepperkok, R., Boulin, C. & Ansorge, W. (1994). A fully automated image acquisition and analysis system for low light level fluorescence microscopy. *Meth. Mol. Cell. Biol.*, 4, 164-170



Damian Brunner

PhD 1995, University of Zurich, Switzerland.

Postdoctoral research at the Imperial Cancer Research Fund, London, UK.

Group Leader at EMBL since 2000.

Joint appointment with the Developmental Biology Unit.

Spatial microtubule organisation and cell morphogenesis

Previous and current research

To create a defined cellular morphology, a cell must polarise and correctly orient its polarity axis. Both depend on the proper arrangement of the actin- and microtubule cytoskeletons. We investigate the contribution of interphase microtubules to cellular morphogenesis. Of central importance is the spatial organisation of the microtubules, which varies tremendously between different cell types. Very little is known about how this variability is achieved at the molecular level but it seems to involve the coordinated activity of a number of conserved microtubule associated factors.

We use the fission yeast *Schizosaccharomyces pombe*, a unicellular eukaryote, to investigate the molecular mechanism that drive spatial microtubule organisation. In these cylindrical cells interphase microtubules serve two important functions. They exert forces that position the nucleus in the cell centre, which defines the future site of cell division, and deposit the Tea1p marker protein, which directs the exact positioning of the two antipodal growth sites. The microtubules are arranged into 3-6 bundles each containing up to 10 anti-parallel microtubules that overlap with their minus ends in the cell centre. The microtubule plus ends grow towards the two opposite cell poles. There the microtubules switch from growth to shrinkage, an event termed “catastrophe”. It is unknown how microtubules identify the cell ends but it involves conserved proteins that localise to the growing microtubule plus ends (+TIPs) and that guide microtubules to the cell ends. One of these proteins, Tip1p, a homologue of human Clip170, acts by preventing premature microtubule catastrophe if inappropriate regions of the cell cortex are encountered. This allows microtubules to continue growing along the cell cortex until the cell end is reached. What remains to be discovered is what differences between cortical regions are sensed by the Tip1p system.

Another +TIP is Mal3p, a homologue of human EB-1. We could show that Mal3p promotes microtubule growth throughout the cell. Its removal from microtubule tips is a pre-requisite for catastrophe to occur. In wild type cells this happens at the cell-end cortex because the presence of Tip1p prevents Mal3p removal in other cortical regions. We are also addressing how +TIPs can “surf” with the growing microtubule tips. We found that there are two different mechanisms operating. Whereas Mal3p is constantly and rapidly exchanged, Tip1p is delivered by the kinesin motor protein Tea2p. When arriving at the microtubule tips the Tip1p/Tea2p complex is slowed down and prevented from falling off by Mal3p. At the same time the fast exchange of Mal3p still allows some motor movement, which permits Tip1p/Tea2p to keep up with the growing tip.

Selected references

Busch, K.E., Hayles, J., Nurse, P. & Brunner, D. (2004). Tea2p kinesin is involved in spatial microtubule organisation by transporting tip1p on microtubules. *Dev. Cell*, 6, 831-843

Busch, K.E. & Brunner, D. (2004). The microtubule plus end-tracking proteins mal3p and tip1p cooperate for cell-end targeting of interphase microtubules. *Curr. Biol.*, 14, 548-559

Future projects and goals

In the future we would like to understand how the activity of +TIPs is spatially regulated and also how the proteins interact with the different microtubule targets and microtubule associated organelles (e.g. cell membrane, nucleus, vesicles, mitochondria). Therefore we are identifying and analysing Tip1p and Mal3p interacting proteins. In addition to the fission yeast work, we have started exploring how microtubules are organised in *Drosophila melanogaster*. In multicellular organisms, spatial microtubule organisation differs considerably between different cell types and it often changes dramatically during development. Because the +TIPs are highly conserved they are likely to have comparable functions in flies and yeast. However, in the flies the cells need to coordinate their microtubule organisation, which involves cell adhesion and signalling mechanisms. These must control the activity of +TIPs, but in turn also the +TIPs may influence cell adhesion and signalling events. We hope to be able to exploit what we learn about the basic microtubule organising machinery in yeast to understand how the activity of this machinery is coordinated amongst the cells of a multicellular organism.

Darren Gilmour

PhD 1996, Cambridge University, UK.
Postdoctoral research at the Max Planck Institute for
Developmental Biology, Tübingen, Germany.
Group Leader at EMBL since 2004.



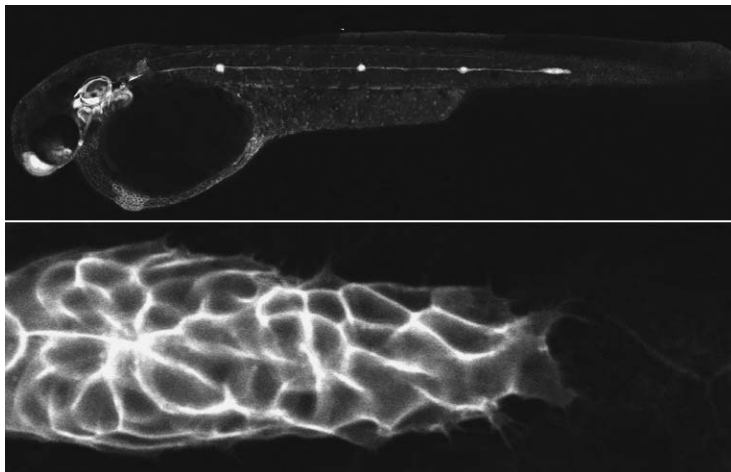
The role of collective cell migration during organ morphogenesis

Previous and current research

The coordinated migration of cohesive groups of cells is a hallmark of both morphogenesis and tumour metastasis. Such “collective” cell migration sculpts the shape of many complex organs and yet very little is known about the logic underlying these coordinated movements. The zebrafish lateral line primordium is a cluster of 200 migrating cells that has several innate features, such as excellent imaging potential and genetic tractability, that allow this process to be studied in the four-dimensional context of the intact developing embryo. We have generated a set of transgenic reporter lines that allow the migration of cells of the lateral line to be captured by multicolour timelapse. In addition, genetic screens have led to the isolation of a number of molecules required for this process, most notably CXCR4, a chemokine-receptor required for the guidance and shape of this migrating tissue. In embryos lacking this receptor or its ligand SDF-1, the cells of the cluster attempt to move in random directions but remain together, resulting in a primordium that is misshapen and immotile. We are currently carrying out a genetic mosaic analysis to determine the role of this chemokine receptor and other key signalling pathways in controlling individual cell behaviours within this moving tissue.

Future projects and goals

Our aim is to understand the mechanisms that coordinate cell behaviour and morphology across a migrating tissue. We will develop probes that allow us to quantify the activity of CXCR4 and other key regulators to determine the extent to which local changes in signalling levels correlate with differences in migratory behaviour in vivo. Dynamic interactions between migrating cells are likely to play an equally important role and we are currently screening for molecules mediating these cell-cell interactions using genetics and expression profiling. Previous data suggests that such mediators encode cell adhesion molecules whose activities are dynamically regulated. It is our hope that 4D-imaging of cytoskeletal dynamics in living embryos will reveal how changes in cell organisation spread across moving tissues during organogenesis.



The zebrafish migrating lateral line organ allows collective migration to be easily studied in vivo

Selected references

- Gilmour, D.T., Knaut, H., Maischein, H.M. & Nusslein-Volhard, C. (2004). Towing of sensory axons by their migrating target cells in vivo. *Nat. Neurosci.*, 7, 491-492
- Gilmour, D.T., Maischein, H.M. & Nusslein-Volhard, C. (2002). Migration and function of a glial subtype in the vertebrate peripheral nervous system. *Neuron*, 34, 577-588
- Gilmour, D.T., Jessen, J.R. & Lin, S. (2002). Manipulating gene expression in the zebrafish in “Zebrafish, a practical approach”. Oxford University Press



Gareth Griffiths

PhD 1975, University of Wisconsin, USA.
Postdoctoral research at the Max Planck Institute for
Biological Cybernetics, Tübingen, Germany.
At EMBL since 1977.

Cell biology of pathogens

Previous and current research

Our group analyses how pathogens interact with cultured cells, focusing on two medically important pathogens, poxviruses and mycobacteria. Using light- and electron microscopy, we established how distinct steps of the poxvirus life cycle interact with cellular membranes and the cytoskeleton.

Mycobacteria enter macrophages by phagocytosis, a process in which the membrane surrounding the host cell unwraps around the bacteria to form a phagosome (see figure). This organelle normally fuses with several organelles of the endocytic pathway of the cell in order to kill the pathogen. *M. tuberculosis* phagosomes fail to fuse with lysosomes and can thus survive. We also use latex bead phagosomes as a model, which enabled us to describe the mechanisms of phagosome maturation in more detail. A possible therapeutical application is our finding that pro-inflammatory lipids help to kill *M. tuberculosis* while anti-inflammatory lipids facilitate pathogen growth in macrophages.

In the poxvirus project, we dissect how the virus modifies the cellular cytoskeleton to promote its infection process. An attenuated vaccinia strain (MVA) that does not affect the cytoskeleton is used as a genetic system to find the molecules involved. Another focus is to unravel the poxvirus structure using cryo-electron tomography techniques.

In the phagosome project, we study the detailed signalling networks regulating actin assembly by phagosomes. This included a systems biology analysis in collaboration with bioinformatic specialists. We then try to link this information to the process in phagosome maturation in macrophages infected with pathogenic and non-pathogenic mycobacteria.

Selected references

Sodeik, B., Schramm, B., Suomalainen, M. & Locker, J.K. (2005). Meeting report: EMBO workshop "Cell Biology of Virus Infection", September 25-29, 2004, EMBL, Heidelberg, Germany. *Traffic*, 6, 351-356

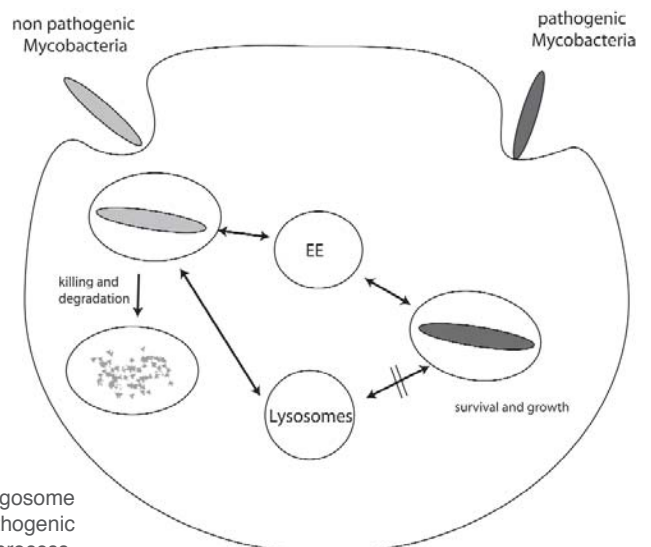
Griffiths, G. (2004). On phagosome individuality and membrane signalling networks. *Trends Cell Biol.*, 14, 343-351

Anes, E., Kuhnelt, M.P., Bos, E., Moniz-Pereira, J., Habermann, A. & Griffiths, G. (2003). Selected lipids activate phagosome actin assembly and maturation resulting in killing of pathogenic mycobacteria. *Nat. Cell Biol.*, 5, 793-802

Sodeik, B. & Krijnen-Locker, J. (2002). Assembly of vaccinia virus revisited: de novo membrane synthesis or acquisition from the host? *Trends Microbiol.*, 10, 15-24

Future projects and goals

- The poxvirus project: we will continue our aim to understand how the virus interacts with its host and unravel the structure and assembly of the virus.
- The phagosome/mycobacterial project: our goals are to understand how macrophages kill mycobacteria with the aim of boosting these processes, and to decipher how *M. tuberculosis* blocks phagosome maturation and find ways to rationally overcome this block.



Summary of normal phagosome maturation and how pathogenic mycobacteria alter this process.

Michael Knop

PhD 1995, University Stuttgart, Germany.
 Postdoctoral research at MPI for Biochemistry, Munich, Germany,
 and the Beatson Institute for Cancer Research, Glasgow, Scotland.
 Group Leader at the Max Planck Institute for Biochemistry, Munich, Germany.
 Group Leader at EMBL since 2001.



Plasma membrane biogenesis during cellular differentiation processes

Previous and current research

Our laboratory is interested in molecular mechanisms of cell differentiation with a focus on the generation of asymmetries, such as processes during membrane shaping and structuring, and spindle-asymmetries.

We use the bakers yeast *Saccharomyces cerevisiae* as a model system.

During meiotic cell differentiation in yeast, a number of processes occur that allows the investigation of the formation of asymmetric structures (see figure). Novel membrane systems are generated at the spindle poles. Subsequently, proper shaping of the membranes towards and around the post-meiotic nuclei has to occur in order to fulfill the faithful formation of the progeny.

The formation of the membranes at the spindle poles can be regulated. The number of membranes depends on the nutritional state. Nutritional limitations are sensed and lead to the inactivation of some specific spindle poles.

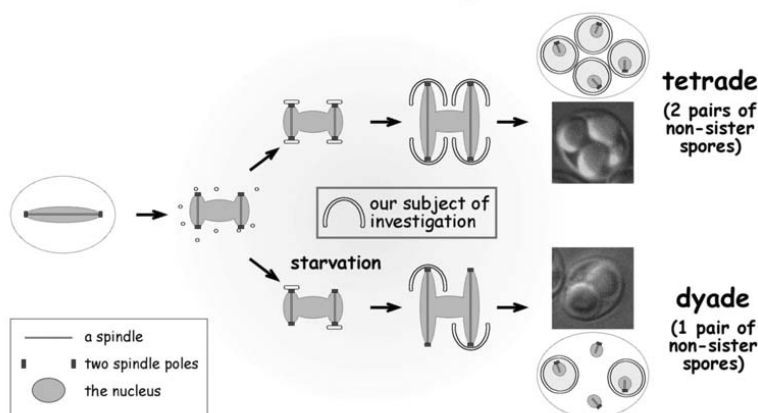
Future projects and goals

We aim to understand the following mechanisms:

- how the cell controls membrane formation at the spindle poles (membrane transport, membrane fusion);
- how specific spindle poles can be selected;
- how the specific shaping of the membrane is achieved.

For this purpose we use genetic, biochemical, molecular and cell biological as well as genome-wide approaches.

Meiosis I.....Meiosis II.....Cytokinesis.....



Progression through meiosis in yeast is associated with the intracellular formation of four daughter cells. The nucleus divides at the end of meiosis two and the post-meiotic nuclei become encapsulated by spindle pole organised membranes, the so-called prospore membranes (PsMs). These membranes are formed from secretory pathway derived precursor membranes, which become directed to the spindle pole bodies. There, the membranes fuse and the formed PsMs subsequently grow around the nuclei.

Selected references

- Moreno-Borchart, A.C., Strasser, K., Finkbeiner, M.G., Shevchenko, A., Shevchenko, A. & Knop, M. (2001). Prospore membrane formation linked to the leading edge protein (LEP) coat assembly. *EMBO J.*, 20, 6946-6957
- Knop, M. & Strasser, K. (2000). Role of the spindle pole body of yeast in mediating assembly of the prospore membrane during meiosis. *EMBO J.*, 19, 3657-3667
- Knop, M., Pereira, G. & Schiebel, E. (1999). Microtubule organisation by the budding yeast spindle pole body. *Biol. Cell*, 91, 291-304
- Knop, M. & Schiebel, E. (1998). Receptors determine the cellular localisation of a γ -tubulin complex and thereby the site of microtubule formation. *EMBO J.*, 17, 3952-3967



François Nédélec

PhD 1998, Université Paris 11, France.

Postdoctoral research at EMBL.

Team Leader since 2002.

Joint appointment with the Structural and Computational Biology Unit.

Cellular architecture

Previous and current research

Modern microscopy has shown us the dynamic nature of biological organisation. Let's look, for example, at eukaryotic cell division and one of its essential parts: chromosome segregation. This is accomplished by a structure called a mitotic spindle, made of chromosomes, microtubules (polar filaments) and numerous associated proteins. All these elements are connected into a solid structure, which remains highly dynamic: the main components – microtubules – are in rapid turnover. The microtubules grow, shrink and disappear in a matter of minutes, while the mitotic spindles can subsist for hours. In fact, none of the microtubule-associated proteins, such as molecular motors, remain for long, yet their permanent stochastic interactions at the molecular level result in a stable overall structure: a spindle conserves its shape and size, and applies the balanced forces necessary to position and segregate the chromosomes precisely.

The spindle is thus a fascinating structure, which illustrates a central question in biology: how can the uncoordinated and inevitably imperfect actions of proteins and molecules result in a structure able to fulfill its biological function with the utmost accuracy?

Obviously, some kind of averaging is going on, but deciphering how multiple proteins cooperate to produce cellular order is not at all straightforward. It is a challenging problem for at least two reasons: 1) there are many different types of protein implicated, and 2) most of their interactions are dynamic and largely uncharacterised.

Within the field of cytoskeletal morphogenesis, we try to address these two fundamental difficulties in practical terms, by developing *in vitro* experiments and modelling tools. The *in vitro* approach allows us to reduce the number of components in the system; we can either remove a specific protein, or start from scratch by mixing purified components. Modelling allows us to recapitulate the process of protein organisation in a framework in which all the interactions are known exactly and can be specified at will. In practice, we develop innovative numerical methods to simulate the collective behaviour of multiple polar fibres and of their associated proteins. Simulations are often used to validate or refute pre-existing ideas, but they can also be used in a more creative way. We generate systematically various properties for the molecules, and use simulations to test their ability to form stable structures. Successful scenarios are identified automatically, leading to the formulation of new hypotheses, which can later be tested experimentally.

Future projects and goals

We want to study systems in which experiments and theory can be synergistically combined. For example, our current work on the mitotic spindle is to further characterise chromosome-microtubule interactions by experimentation, in order to include chromosomes into the simulations. We are generally interested in any cytoskeletal process observed at the cellular level, such as morphogenesis in *S. pombe*, or the generation of asymmetry in the first division of the *C. elegans* embryo.

Selected references

Nédélec, F., Surrey, T. & Karsenti, E. (2003). Self-organisation and forces in the microtubule cytoskeleton. *Curr. Opin. Cell Biol.*, 15, 118-124

Nédélec, F. (2002). Computer simulations reveal motor properties generating stable anti-parallel microtubule interactions. *J. Cell Biol.*, 6, 1005-1015

Surrey, T., Nédélec, F., Leibler, S. & Karsenti, E. (2001). Physical properties determining self-organisation of motors and microtubules. *Science*, 292, 1167-1171

Rainer Pepperkok

PhD 1992, University Kaiserslautern, Germany.
Postdoctoral research at University of Geneva, Switzerland.
Lab Head at the Imperial Cancer Research Fund, London, UK.
At EMBL since 1998.



Membrane traffic in the early secretory pathway

Previous and current research

Transport between two adjacent membranes in the secretory pathway is mediated by coated vesicular carriers and involves at least four basic steps: (1) budding of vesicles from donor membranes; (2) transport towards the target membrane; (3) docking and fusion of vesicles with the target membrane; and (4) recycling of the transport machinery back to the donor membrane. To warrant both the specificity of delivery of cargo and the maintenance of the unique protein and lipid compositions of the organelles involved, these four steps must be tightly regulated.

Our research focuses on how membrane traffic between the endoplasmic reticulum and the Golgi complex is regulated in space and time. To investigate this we have developed light microscopy approaches to directly visualise in living cells the kinetics of secretory and organelle markers simultaneously with vesicular coat molecules (COPI and COPII) and their regulators. We also use fluorescence recovery after photobleaching (FRAP) and fluorescence resonance energy transfer measurements (FRET) in order to study the temporal and spatial regulation of the molecular interactions involved.

We are also interested in the trafficking of a specific class of protein toxins travelling the secretory pathway in reverse from the plasma membrane to the endoplasmic reticulum. Using this system we could obtain evidence for a novel COPI-independent recycling pathway from the Golgi complex to the endoplasmic reticulum.

In order to identify further specific regulators of the alternative transport pathways and potential regulators of COPI and COPII in microtubule dependent transport, we systematically localise novel human proteins in cells in collaboration with the department of Annemarie Poustka at the DKFZ in Heidelberg. Using the high content screening microscopy platform, that we have developed in the past, we use automated microscopy-based assays that reveal a potential functional implication of the proteins under study in secretory transport, Golgi morphology or microtubule stability.

Future projects and goals

We will study the novel proteins, which we reveal in our screens to be involved in the early secretory pathway, in further details. An important question in this context will be if and how they participate in the temporal and spatial organisation of ER-exit sites and their function. We will also exploit our developments on high content screening microscopy and transfections on siRNA arrays to conduct cell based "genome-wide" siRNA screens with the aim of identifying the interaction networks that regulate ER-exit of cargo and its link to subsequent transport along microtubules.

Selected references

- Neubrand, V.E., Will, R.D., Mobius, W., Poustka, A., Wiemann, S., Schu, P., Dotti, C.G., Pepperkok, R. & Simpson, J.C. (2005). γ -BAR, a novel AP-1-interacting protein involved in post-Golgi trafficking. *EMBO J.*, 24, 1122-1133
- Watson, P., Forster, R., Palmer, K.J., Pepperkok, R. & Stephens, D.J. (2005). Coupling of ER exit to microtubules through direct interaction of COPII with dynactin. *Nat. Cell Biol.*, 7, 48-55
- Starkuviene, V., Liebel, U., Simpson, J.C., Erfle, H., Poustka, A., Wiemann, S. & Pepperkok, R. (2004). High-content screening microscopy identifies novel proteins with a putative role in secretory membrane traffic. *Genome Res.*, 14, 1948-1956
- Erfle, H., Simpson, J.C., Bastiaens, P.I. & Pepperkok, R. (2004). siRNA cell arrays for high-content screening microscopy. *Biotechniques*, 37, 454-8, 460
- Simpson, J.C. & Pepperkok, R. (2003). Localizing the proteome. *Genome Biol.*, 4, 240



Ernst H.K. Stelzer

PhD 1987, University of Heidelberg, Germany.
Postdoctoral research at EMBL.
Group Leader at EMBL since 1989.

Light microscopy group

Previous and current research

During the past few years our group has evolved. We started with the development of confocal fluorescence microscopes but always collaborated with scientists inside and outside EMBL, making excellent use of various microscopy techniques for biological and biophysical research. The group was actively involved in the development of many methods and devices that are now commonly used all over the world, and collaborated with industry to make sure that technology became generally available (e.g. Carl Zeiss's LSM 5 series). Our group includes biologists who apply and guide our developments, as well as physicists and engineers who devote their efforts to addressing challenges that are extrapolated from current developments in the life sciences. We develop and maintain a healthy mix of basic and applied research, technical development and applications.

The various types of lasers used for research in the life sciences has remained a central topic. The Photonic Force Microscope (PFM) is used to determine the location of small latex beads attached to single molecules with subnanometer precision at Megahertz rates. The Selective Plane Illumination Microscope (SPIM), on the other hand, is particularly useful for the observation of large biological objects. The diffraction limited UV laser cutter has been used successfully for the manipulation of in vivo systems.

Future projects and goals

SPIM will be further developed by combining it, for example, with FLIM, by adding laser cutters and by offering deep blues and red lasers to provide more excitation lines. SPIM, with its dramatically decreased bleaching (or in more general terms decreased damage), provides an excellent platform for live cell imaging.

We are currently witnessing the transition from a "flat" two-dimensional cell biology to a more complex three-dimensional cell biology. The evidence that cells cultivated on cover slips differ dramatically from their three-dimensional natural counterparts, which depend on other cell types to support them, is extremely strong and should be addressed. We plan not only to develop tools such as the SPIM that allow us to experiment with more complex objects, but also to play an active role in the development of a revised approach to the life sciences.

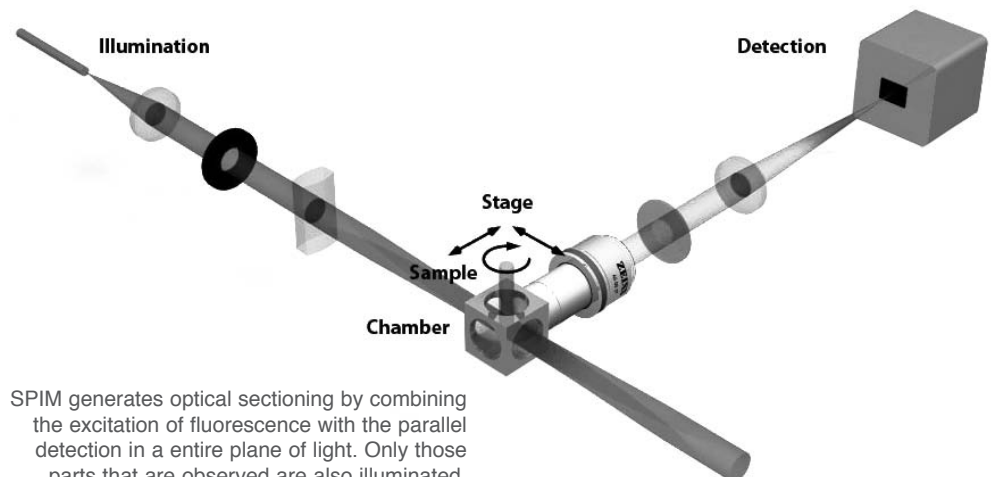
Selected references

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J. & Stelzer, E.H.K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305, 1007-1009

Rohrbach, A., Tischer, C., Neumayer, D., Florin, E.-L. & Stelzer, E.H.K. (2004). Trapping and tracking a local probe with a photonic force microscope. *Rev. Sci. Instr.*, 75, 2197-2210

Colombelli, J., Grill, S.W. & Stelzer, E.H.K. (2004). UV diffraction limited nanosurgery of live biological tissues. *Rev. Sci. Instr.*, 75, 472-478

Grill, S.W., Howard, J., Schaffer, E., Stelzer, E.H. & Hyman, A.A. (2003). The distribution of active force generators controls mitotic spindle position. *Science*, 301, 518-521



SPIM generates optical sectioning by combining the excitation of fluorescence with the parallel detection in a entire plane of light. Only those parts that are observed are also illuminated.

Thomas Surrey

PhD 1995, Eberhard-Karls University, Tübingen, Germany.
Postdoctoral research at Princeton University, USA and EMBL.
Staff Scientist 2001-2002.
Team Leader since 2002.



Physical biology of molecular motors involved in intracellular organisation

Previous and current research

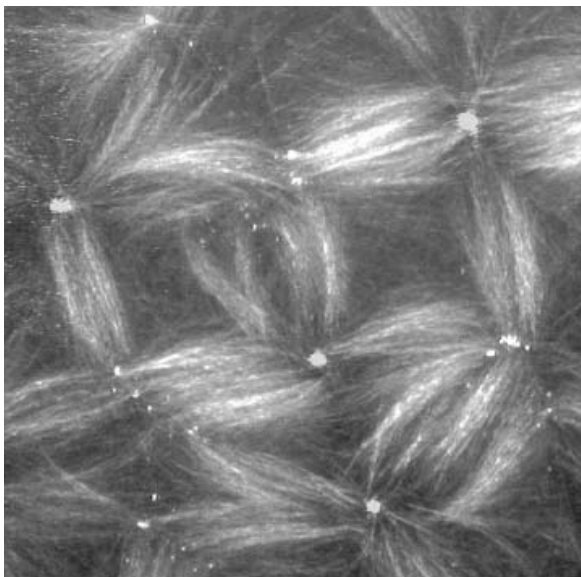
Motor proteins are key determinants for the spatial organisation of eukaryotic cells. They are thermodynamic non-equilibrium machines playing a crucial role for the dynamic nature of cellular order. In fact, they provide a paradigm for the concept of intracellular order depending on molecular dynamics. How exactly the collective behaviour of various motors with different kinetic properties drives the organisation of the cytoskeleton is not understood.

Presently, we are following several lines of research to determine how motors contribute to intracellular order. We work on: (1) motile properties of purified motors in vitro, single molecule and ensemble measurements; (2) the effect of molecular motors on microtubule dynamics in vitro; (3) self-organisation of motors and microtubules in vitro; (4) development of novel assays to measure kinetic properties of motors inside cells and in cell extracts; and (5) regulation of motors by kinases in vitro and in cell extracts.

The techniques we use include advanced light microscopy, biochemistry, cell biology and molecular biology.

Future projects and goals

Our goal is to understand how the biochemical and physical properties of microtubule-dependent motors determine their specific activity in a cell. We would like to understand how the biological “function” of a whole set of motors and microtubules for a given process (such as mitotic spindle assembly) is generated from their coordinated and regulated interactions. We are interested in searching for design principles underlying intracellular dynamics and organisation. Understanding the regulation of motor protein properties will be of crucial importance. Therefore, we will develop tools that will allow us monitor and to manipulate the spatio-temporal regulation of motor protein activities in cells and cell extracts using modern light microscopy techniques.



Network of microtubules and two kinds of motor proteins created by self-organisation in vitro. Microtubules are visualised by darkfield microscopy; one of the two motor species is visualised by fluorescence microscopy.

Selected references

- Bringmann, H., Skiniotis, G., Spilker, A., Kandels-Lewis, S., Vernos, I. & Surrey, T. (2004). A kinesin-like motor inhibits microtubule dynamic instability. *Science*, 303, 1519-1522
- Surrey, T., Nédélec, F., Leibler, S. & Karsenti, E. (2001). Physical properties determining self-organisation of motors and microtubules. *Science*, 292, 1167-1171

Developmental Biology Unit

Developmental Biology is concerned with how the body patterns of multicellular animals are designed and constructed. The life of a multicellular animal starts from a single haploid cell, the egg, which shortly after fertilisation begins to divide to make the embryo. Developmental Biologists wish to understand how the genetic information is used to make the many different types of cells in the embryo and organize them in a coherent body plan.

Communication between cells is a central feature of this process. Exchange of information between cells is needed to coordinate cell fate choice, control growth and direct morphogenesis. One focus of research in the Unit concerns the nature of these signals. What are they? How do cells receive and interpret them? How do signals elicit short-term changes in cell behaviour and long-term changes in the program of gene expression that makes cells different from one another? Work in *Drosophila*, *C. elegans* and frogs, fish and mice has shown that the signalling molecules and the molecular mechanisms by which the signals are transduced have been conserved in evolution. Several groups study how cells communicate. The approaches include signal transduction biochemistry, genetic analysis of signal transduction in mice, fish and flies; gene discovery and analysis of gene function in fish and flies; and studies of how signals affect cell fate specification, cell migration and morphogenesis in vivo in mice, fish and flies.

A concern with cellular mechanisms is an increasingly important focus of the Developmental Biology Unit. What do individual cells do and how do they work together? Research in this area includes studying how asymmetries are generated at the single-cell level and how such asymmetries control gene expression at the subcellular level. Another focus of interest is how groups of cells coordinate their movement. Understanding the cellular and molecular basis for these behaviours will help us to understand the morphogenetic processes that shape the body during embryonic development.

Stephen Cohen

PhD 1983, Princeton University, USA. Postdoctoral research at the Whitehead Institute, MIT, Boston, USA, and Max Planck Institute for Developmental Biology, Tübingen, Germany. Assistant Professor, Howard Hughes Medical Institute, Baylor College of Medicine, Houston, USA. Group Leader at EMBL since 1993. Unit Coordinator since 1996.



microRNAs growth and metabolism

Previous and current research

Gradients of secreted signalling proteins organize spatial pattern and control tissue growth during animal development. The signals are produced by small groups of cells at defined positions. Once secreted, these proteins function as morphogens to instruct cells about their fate as a function of the local concentration of the signalling protein, hence they can be said to convey positional information. We are taking molecular and genetic approaches to understanding how these morphogen gradients are established and how they control tissue growth during development. Genetic screens have begun to identify new growth regulators that are controlled by intercellular signalling.

Growth of tissues and organs during animal development involves careful coordination of the rates of cell proliferation and cell death. The connections between control of cell proliferation and apoptosis in normal development and in cancer are not yet well understood. Among the new genes we've identified are a protein kinase that controls cell survival in response to proliferative cues and a microRNA. microRNAs are a class of short 21-23 nucleotide RNA molecules implicated in the control of gene expression. To date few miRNAs have been assigned functions, but those that are understood regulate post-transcriptional gene expression. With Alex Stark and Rob Russell, we have developed bioinformatic tools to identify the target mRNAs regulated by miRNAs and find that the average miRNA regulates hundreds of genes. Hundreds of miRNAs have been identified leading us to estimate that miRNAs regulate at least a third of all genes. Combining target prediction with experimental analysis of miRNA expression has begun open the door to a broader understanding of the roles that miRNAs play in evolution and development.

Growth during embryonic life and metabolism are closely linked. Recent work has identified a number of genes implicated in control of metabolism as well as tissue growth.

Future projects and goals

Our goal is to understand the cellular and molecular mechanisms controlling growth and metabolism in animal development.

Selected references

- Brennecke, J., Stark, A., Russell, R.B. & Cohen, S.M. (2005). Principles of microRNA-target recognition. *PLoS Biol.*, 3, e85
- Teleman, A.A., Chen, Y.-W. & Cohen, S.M. (2005). 4E-BP functions as a metabolic brake used under stress conditions but not during normal growth. *Genes Dev.*, in press
- Stark, A., Brennecke, J., Russell, R.B. & Cohen, S.M. (2003). Identification of *Drosophila* microRNA targets. *PLoS Biol.*, 1, E60
- Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. & Cohen, S.M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell*, 113, 25-36



Detlev Arendt

PhD 1999, University of Freiburg, Germany.
Postdoctoral research at EMBL.
Team Leader since 2002.

Evolution of gastrulation and central nervous systems in Bilateria

Previous and current research

We are intrigued by one of the remaining great mysteries in animal evolution: how did our brain come into existence? What did it look like at first and how did it function? We are especially interested in the brain of an extinct animal known as Urbilateria, the last common ancestor of humans, flies and most other “higher” animals that live today. We know that these ancient creatures lived some 600 million years ago along the coastline of the ocean.

We have therefore chosen to work on a “living fossil”, the marine ragworm *Platynereis dumerilii* (Polychaeta, Annelida) that we keep in laboratory culture. This species exhibits many ancient features in its lifestyle, anatomy and development, and appears to be a good approximation to the Urbilaterian nervous system. In bioinformatics comparisons we found that *Platynereis* also shows an ancestral gene inventory and ancestral gene structure.

We combine morphological and molecular approaches in a novel evo-devo approach, the molecular comparison of cell types. Animal brains are made up of different sorts of sensory neurons, motor- and interneurons. Each type displays a characteristic “molecular fingerprint”, a unique combination of specifying transcription factors and downstream effector genes such as receptors, transmitters or neuropeptides. The comparison of molecular fingerprints allows the tracing of cell types through animal evolution. For example, in the *Platynereis* brain we have characterised a special type of photoreceptor cell, a “ciliary photoreceptor” that by molecular fingerprint comparison relates to the rods and cones, the visual photoreceptors of the vertebrate retina. This has led to the fascinating hypothesis that the vertebrate eye has evolved from within the Urbilaterian brain.

Selected references

- Arendt, D., Tessmar-Raible, K., Snyman, H., Dorrestijn, A.W. & Wittbrodt, J. (2004). Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science*, 306, 869-871
- Raible, F. & Arendt, D. (2004). Metazoan evolution: some animals are more equal than others. *Curr. Biol.*, 14, R106-R108
- Tessmar-Raible, K. & Arendt, D. (2003). Emerging systems: between vertebrates and arthropods, the Lophotrochozoa. *Curr. Opin. Genet. Dev.*, 13, 331-340
- Arendt, D. (2003). Evolution of eyes and photoreceptor cell types. *Int. J. Dev. Biol.*, 47, 563-571

Future projects and goals

Besides ciliary photoreceptors, the *Platynereis* brain harbours several types of sensory-neurosecretory cells. The ongoing molecular characterisation of these cell types again reveals striking parallels to vertebrate cell types, mostly situated in the hypothalamus. In addition, head sensory organs innervate the brain, such as three pairs of eyes and three pairs of chemo-/mechanosensory organs. Future projects will aim at the further molecular and functional characterisation of the *Platynereis* brain and associated sensory organs.

Unexpectedly, the clear picture is emerging that the *Platynereis* brain harbours many cell types so far known only for the vertebrates, but in a much more simple, very different overall arrangement. Our goal is to elucidate the functioning of these cell types in the ancient marine environment in order to gain insight into the evolutionary origins of the vertebrate brain.



Platynereis dumerilii (Polychaeta, Annelida, Lophotrochozoa)

Anne Ephrussi

PhD 1985, Massachusetts Institute of Technology, Cambridge, USA.
Postdoctoral research at Harvard University and the
Whitehead Institute, Cambridge, USA.
Group Leader at EMBL since 1992.



Cell polarity and RNA localisation

Previous and current research

Polarity is a central feature of most eukaryotic cells and is the basis for asymmetric cell division, specialised cell functions and basic developmental processes. Polarisation involves the asymmetric distribution of cytoskeletal structures, organelles and molecules within the cell. Some polarising proteins, such as Par-1 kinase, are highly conserved, suggesting that common mechanisms underlie cell polarisation, from yeast to vertebrates.

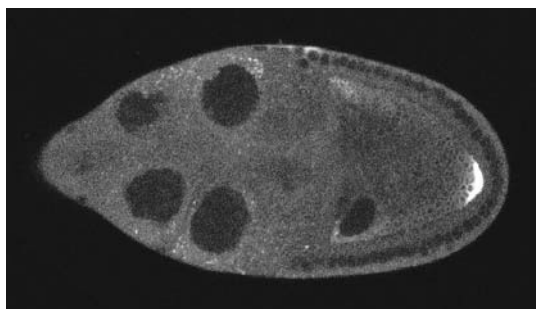
Establishment of the body axes of the *Drosophila* embryo relies on asymmetric localisation of cell fate determinants within the oocyte, before fertilisation. *oskar* is localised as an mRNA at the posterior pole of the oocyte. *Oskar* directs abdomen and germline formation, hence tight restriction of its activity to the posterior is critical. This is achieved by mRNA localisation-dependent translation: *oskar* translation is repressed during transport and activated once the mRNA reaches the posterior pole. Assembly of the *oskar* mRNA localisation complex begins in the nucleus with the splicing-dependent deposition of the Exon Junction Complex of proteins at the first exon-exon junction of *oskar* RNA. Co-assembled with the EJC and proteins bound to its 3'UTR, *oskar* mRNA is transported by a mechanism involving kinesin heavy chain to the posterior of the oocyte, where it is translated and anchored.

As oocyte polarity is critical for correct *oskar* mRNA localisation, this system is ideally suited to genetic, molecular and cell biological investigation of the processes of cell polarisation, mRNA localisation and translational control. We make use of the polarised *Drosophila* oocyte as a model to study (1) the mechanisms underlying establishment and maintenance of cell polarity, and in particular the role of Par-1 kinase in cell polarisation, and (2) assembly of the *oskar* mRNP complex and the mechanisms of *oskar* mRNA localisation, translational regulation and post-translational control.

Future projects and goals

Using genetics, proteomics, biochemistry, and molecular and cell biological approaches, we are investigating:

- the mechanisms underlying determination and polarisation of the *Drosophila* oocyte;
- Par-1 kinase, its targets and the mechanisms by which it exerts its polarising functions;
- the role of the cytoskeleton and motors in *oskar* mRNA localisation;
- the architecture of the *oskar* mRNA localisation complex: cis acting RNA elements, the EJC and other interacting proteins, and how they assemble to form a functional localisation mRNP;
- the mechanisms coupling *oskar* mRNA localisation and translational control;
- *Oskar* protein and its assembly of the polar granules, the germline granules of *Drosophila*.



Our goal is to understand how oocyte polarity is established and translated into a correctly patterned embryo.

In a *Drosophila* egg chamber *oskar* mRNA, Staufen protein and a microtubule polarity marker colocalise at the posterior of the oocyte.

Selected references

- Vaccari, T., Rabouille, C. & Ephrussi, A. (2005). The *Drosophila* PAR-1 spacer domain is required for lateral membrane association and for polarisation of follicular epithelial cells. *Curr. Biol.*, 15, 255-261
- Riechmann, V. & Ephrussi, A. (2004). Par-1 regulates bicoid mRNA localisation by phosphorylating Exuperantia. *Development*, 131, 5897-5907
- Yano, T., De Quinto, S.L., Matsui, Y., Shevchenko, A., Shevchenko, A. & Ephrussi, A. (2004). Hrp48, a *Drosophila* hnRNP/B homologue, binds and regulates translation of *oskar* mRNA. *Dev. Cell*, 6, 637-648
- Hachet, O. & Ephrussi, A. (2004). Splicing of *oskar* RNA in the nucleus is coupled to its cytoplasmic localisation. *Nature*, 428, 959-963
- Yoshida, S., Muller, H.A., Wodarz, A. & Ephrussi, A. (2004). PKA-R1 spatially restricts *Oskar* expression for *Drosophila* embryonic patterning. *Development*, 131, 1401-1410



Eileen Furlong

PhD 1996, University College Dublin, Ireland.
Postdoctoral research at Stanford University, USA.
Group Leader at EMBL since 2002.
Joint appointment with the Gene Expression Unit.

Global regulatory networks required for tissue development

Previous and current research

The progression of an undifferentiated mesodermal cell towards a mature muscle requires the activity of a number of highly conserved transcription factors. Genetic studies have clearly shown that these transcription factors are essential to regulate different development events, e.g. myoblast specification, myoblast fusion and terminal muscle differentiation (see figure). However, despite the essential and conserved role of these transcription factors, very few of their direct transcriptional target genes are known.

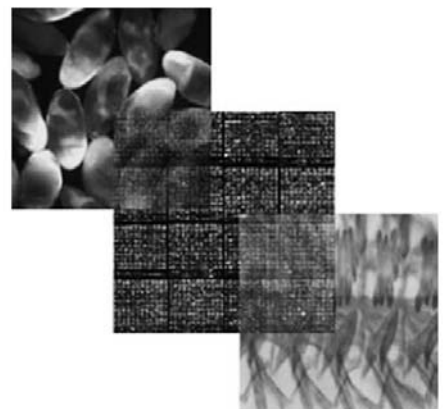
With the development of new techniques like gene expression profiling and Chromatin ImmunoPrecipitation followed by microarray analysis (ChIP on chip), it is now possible to study transcriptional regulation on a whole-genome scale. Our group uses and develops functional genomic tools in *Drosophila* to study the regulation of muscle development at a global level. Of particular interest is the regulation of myoblast specification and myoblast fusion. During development, the mesoderm becomes progressively subdivided through the action of a number of transcription factors into three major muscle types: the somatic muscle (body wall muscle), heart muscle and the visceral muscle (gut muscle). We are currently using genomic approaches to identify new genes that are involved in the specification and development of these three different muscle types. A large number of novel muscle genes have already been identified, many of which have conserved vertebrate homologues. The current challenge is to gain a detailed understanding of the function of these new genes in myoblast specification.

A second major interest of the lab is the study of myoblast fusion. During the development of skeletal muscle, specified muscle cells undergo an additional step of myoblast fusion before differentiating into a multinucleated myotube. Very little is known about the regulation of myoblast fusion and several key questions remain to be addressed. For example, the size of the resulting muscle depends on the number of fusion events that occurs. How is this counting mechanism of successive fusion events achieved? What controls the coordination of myoblast fusion with myotube differentiation? Previous work in the lab has identified a number of very interesting new genes, which are likely to be essential for myoblast fusion. In order to gain a comprehensive insight into the molecular regulation of myoblast fusion, we are currently using two genomic approaches to identify novel myoblast fusion genes, on a genome-wide scale.

Future projects and goals

In order to understand a process as complex as tissue development we need to place these targets genes into context. Using a combination of *Drosophila* genetics, genomics and bioinformatics approaches we are constructing a transcriptional network that will allow target genes to be classified into groups of genes or functional modules. Both the topology and dynamics of this network will be used to decipher the logic of how these integrated circuits regulate specific aspects of tissue development.

Identifying the direct targets of a transcription factor is a crucial step to understanding its function. In collaboration with a number of other groups we are beginning to examine the level of conservation and divergence of these transcription factors target genes. This will yield new insights into how the “muscle regulatory circuitry” evolved.



Selected references

- Furlong, E.E. (2004). Integrating transcriptional and signalling networks during muscle development. *Curr. Opin. Genet. Dev.*, 14, 343-350
- Artero, R., Furlong, E.E., Beckett, K., Scott, M.P. & Baylies, M. (2003). Notch and Ras signalling pathway effector genes expressed in fusion competent and founder cells during *Drosophila* myogenesis. *Development*, 130, 6257-6272
- Arbeitman, M.N., Furlong, E.E., Imam, F., Johnson, E., Null, B.H., Baker, B.S., Krasnow, M.A., Scott, M.P., Davis, R.W. & White, K.P. (2002). Gene expression during the life cycle of *Drosophila melanogaster*. *Science*, 297, 2270-2275
- Furlong, E.E., Andersen, E.C., Null, B., White, K.P. & Scott, M.P. (2001). Patterns of gene expression during *Drosophila* mesoderm development. *Science*, 293, 1629-1633
- Furlong, E.E., Proffitt, D. & Scott, M.P. (2001). Automated sorting of live transgenic embryos. *Nat. Biotechnol.*, 19, 153-156

Carl Neumann

PhD 1997, EMBL.
Postdoctoral research at the Max Planck Institute for
Developmental Biology, Tübingen, Germany.
Group Leader at EMBL since 2001.



The role of cell-cell signalling in patterning the zebrafish retina and paired fins

Previous and current research

Signalling among cells is used to generate pattern during the development of multicellular animals. Interaction between distinctly specified cells generates new cell types, and sequential rounds of cell-cell interactions culminate in the generation of precise numbers of differentiated cells in the correct locations. We are studying this process in two organs of the zebrafish, *Danio rerio*: the paired limbs and the visual system.

The vertebrate retina develops from an evagination of the neural tube. It is initially a non-neuronal, unpatterned epithelium, which develops into a highly organised three-dimensional array of neurons. Neurogenesis in the retina proceeds in a wave from the centre to the periphery, and this wave is driven by the signalling activity of the secreted Sonic Hedgehog protein. Hedgehog expression is induced in the first neurons that differentiate in the retina. These cells then secrete Hedgehog protein, which travels to nearby uncommitted cells, and there induces neuronal differentiation and another round of Hedgehog expression, thereby propagating the neurogenic wave.

Like the limbs of land vertebrates, the paired fins of zebrafish develop from buds that arise from the lateral plate mesoderm. The development of vertebrate limb buds is triggered by a cascade of genes including members of the Fgf and Wnt families, and these signalling events lead to the activation of the transcription factor Tbx5 in the lateral plate mesoderm. Tbx5, in turn, is required for the activation of Fgf10, which relays the limb inducing signal to the overlying ectoderm. We have found that the zebrafish Fgf24 gene, which belongs to the Fgf8/17/18 subfamily of Fgf ligands, acts downstream of Tbx5 to activate Fgf10 expression in the lateral plate mesoderm. This finding reveals an additional level of Fgf signalling in vertebrate limb initiation. In addition, we have also found that Fgf24 controls the migration of mesodermal precursor cells to the limb primordium.

Future projects and goals

Currently, we are interested in determining how the Sonic Hedgehog signal is interpreted in the retina to control neurogenesis. We are also analysing new mutants affecting both the development of the retina or of the paired fins to gain new insight into the genetic networks controlling the development of these structures.

It is our goal to understand the molecular mechanisms by which cell-cell interactions are used to generate pattern in vertebrates.

Selected references

- Shkumatava, A. & Neumann, C.J. (2005). Shh directs cell-cycle exit by activating p57Kip2 in the zebrafish retina. *EMBO Rep.*, 6, 563-569
- Shkumatava, A., Fischer, S., Muller, F., Strahle, U. & Neumann, C.J. (2004). Sonic hedgehog, secreted by amacrine cells, acts as a short-range signal to direct differentiation and lamination in the zebrafish retina. *Development*, 131, 3849-3858
- Fischer, S., Draper, B.W. & Neumann, C.J. (2003). The zebrafish fgf24 mutant identifies an additional level of Fgf signalling involved in vertebrate forelimb initiation. *Development*, 130, 3515-3524
- Neumann, C.J. (2001). Pattern formation in the zebrafish retina. *Sem. Cell Dev. Biol.*, 12, 485-490
- Neumann, C.J. & Nüsslein-Volhard, C. (2000). Patterning of the zebrafish retina by a wave of sonic hedgehog activity. *Science*, 289, 2137-2139



Pernille Rørth

PhD 1993, University of Copenhagen, Denmark.
Staff Associate at the Carnegie Institution of Washington, USA.
Group Leader at EMBL since 1998. Senior Scientist since 2003.
Joint appointment with the Cell Biology and Biophysics Unit.

Regulation of cell migration

Previous and current research

During animal development, many cells migrate from one place to another in order to perform their biological functions. Cell migration is a complex process that involves dynamic interactions between migrating cells and the tissue through which they migrate. In order to migrate, cells change their shape and adhesion properties and become invasive and motile. They must also read guidance cues provided by the target tissue that tell them where to go and when to stop. To analyse these events at the molecular and cellular level, we are studying a specific migration event in *Drosophila*: the migration of border cells in the ovary. We use genetics to identify genes important for cell migration *in vivo* and investigate the molecular action of these genes in order to understand how cell migration is controlled.

We have identified the guidance cues that direct border cell migration and the receptors that interpret this information. The receptors are two receptor tyrosine kinases, PVR and EGFR. We are now investigating the signalling pathway downstream of the receptors responsible for guidance, as well as how this signal is localised properly within the cell and its relationship to cell polarisation. We are also analysing how border cells control their adhesion to, and traction on, the substratum as well as the cellular mechanics of migration. In parallel with this, we are carrying out new genetic screens to identify other critical players in the control of migration.

A number of transcription factors, including *Slbo*, are specifically required in border cells for initiation of migration. To investigate how cells become migratory and the role of specific transcription factors, we are using DNA-microarrays. We investigate gene expression regulated by *Slbo* as well as more generally gene expression changes associated with initiation of cell migration. The MAL/SRF transcription factor complex has a different and intriguing role in migrating cells. MAL/SRF target genes and the molecular mechanism of MAL regulation are being explored.

Selected references

Somogyi, K. & Rørth, P. (2004). Evidence for tension-based regulation of *Drosophila* MAL and SRF during invasive cell migration. *Dev. Cell*, 7, 85-93

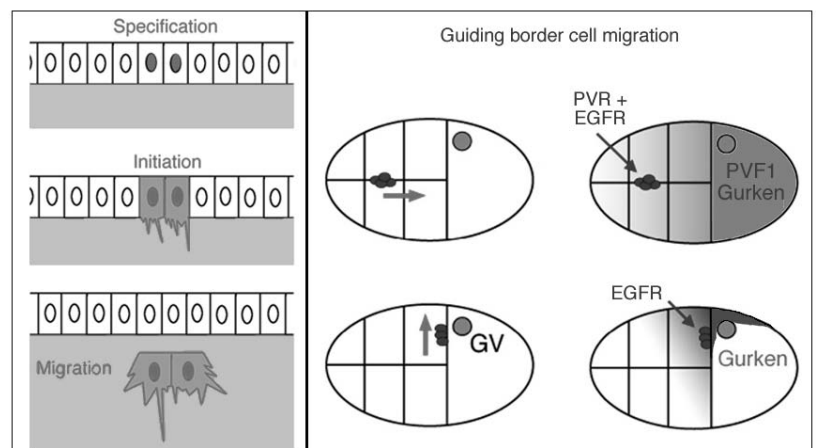
Fulga, T.A. & Rørth, P. (2002). Invasive cell migration is initiated by guided growth of long cellular extensions. *Nat. Cell Biol.*, 4, 715-719

Duchek, P., Somogyi, K., Jekely, G., Beccari, S. & Rørth, P. (2001). Guidance of cell migration by the *Drosophila* PDGF/VEGF receptor. *Cell*, 107, 17-26

Duchek, P. & Rørth, P. (2001). Guidance of cell migration by EGF receptor signalling during *Drosophila* oogenesis. *Science*, 291, 131-133

Future projects and goals

Although our emphasis is presently on the migration of border cells, we are also exploring other examples of cell migration during development. As part of our analysis, we have initiated real-time imaging of migrating cells *in vivo* and plan to look at real-time dynamics of key molecules. We are interested in further exploring the relationship between signalling and cytoskeletal forces.



Mathias Treier

PhD 2004, University of Heidelberg, Germany.
 Postdoctoral research at the University of California, San Diego, USA.
 Group Leader at EMBL since 2000.
 Joint appointment with EMBL-Monterotondo.



Mammalian organogenesis and physiology

Previous and current research

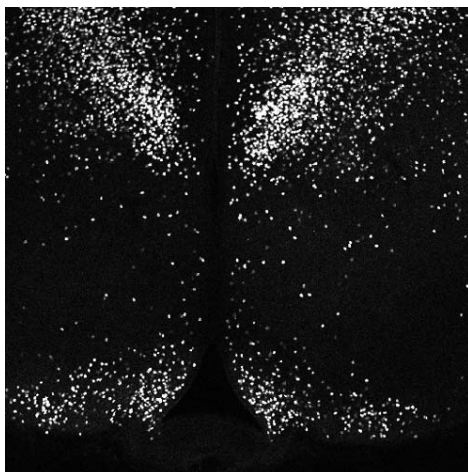
The development of specific cell types during mammalian organogenesis has been intensively studied over the last decade. The challenge now is to understand how these different cell types function in the context of a whole organ, and the ultimate goal is to determine how all the different organs function in a concerted action to create a whole mammalian organism. We are employing mouse genetics to study several steps of mammalian organogenesis with special focus on the hypothalamic-pituitary axis and the kidney. The hypothalamic-pituitary (HP)-axis serves critical homeostatic functions by regulating key peripheral endocrine organs i.e. thyroid gland, adrenal gland and kidney. In addition, the HP-axis plays a central role in the sexual maturation of the reproductive organs.

In particular, we are interested in how the HP-axis regulates mammalian energy homeostasis. Within the DIABESITY FP-6 framework consortium we are studying the development and function of hypothalamic circuitries that are implicated in this process (www.eurodiabetesity.org).

Another major research direction in the lab is concerned with the understanding of organ growth control. Hardly anything is known about why an organ stops growing after it has reached a certain size. We are mainly modelling two human diseases in the mouse to gain further insights into this problem. First, we are investigating the molecular mechanisms that underlie the development of Multiple Endocrine Neoplasia type 1 (MEN1) that leads to a general benign overgrowth of endocrine organs i.e. pituitary gland and pancreas. Second we are interested in the abnormal growth regulation during the development of Polycystic Kidney Disease (PKD). PKD is the most common genetic, life-threatening disease affecting an estimated 12.5 million people worldwide – regardless of sex, age, race or ethnic origin (www.pkd-cure.org).

Future projects and goals

We have established a series of mouse knockout strains for different transcription factors that serve us now as model systems to gain novel insights into the above described problems. Furthermore, we have been able to identify putative stem cell niches in our organs of interest. This will allow us in the future to genetically manipulate the proliferation state of these stem cell pools in the adult organism to devise strategies to delay degenerative processes within these organs.



Genetic labeling of hypothalamic neuronal populations with H2bGFP.

Selected references

- Müller, T., Anlag, K., Wildner, H., Britsch, S., Treier, M. & Birchmeier, C. (2005). The bHLH factor Olig3 coordinates the specification of dorsal neurons in the spinal cord. *Genes Dev.*, 19, 733-743
- Schmidt, D., Ovitt, C.E., Anlag, K., Fehsenfeld, S., Gredsted, L., Treier, A.C. & Treier, M. (2004). The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. *Development*, 131, 933-942
- Treier, M., O'Connell, S., Gleiberman, A., Price, J., Szeto, D.P., Burgess, R., Chuang, P.T., McMahon, A.P. & Rosenfeld, M.G. (2001). Hedgehog signalling is required for pituitary gland development. *Development*, 128, 377-386



Jochen Wittbrodt

PhD 1991, University of Munich, Germany.

Postdoctoral research at the Biocenter, University of Basel, Switzerland.

Junior Group Leader at the Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.

Group Leader at EMBL since 1999.

Brain patterning and eye development

Previous and current research

The vertebrate eye is composed of neuroectodermal (optic cup) and surface ectodermal (lens, cornea) derivatives and it emerges from an epithelial Anlage by inductive interactions beginning at late gastrula stages. Under the influence of midline signalling during neurulation the single retina Anlage is split into two retinal primordia localised in the lateral wall of the prosencephalon.

Subsequent evagination of the primordia results in the formation of optic vesicles that differentiate to the seven cell types of the neural retina, the retinal pigmented epithelium and the optic stalk respectively. In anamniotes (fish, amphibia), the ciliary margin of the neural retina contains a stem cell population that gives rise to all retinal cell types and facilitates lifelong growth of the eye.

We investigate vertebrate (medaka, zebrafish) eye development following three complementary experimental strategies: functional studies involving large scale mutagenesis screens and mutant analysis; functional analyses involving gain- and loss-of-function analysis by ectopic expressing of transgenes in the developing eye or by blocking gene function using the morpholino knock-down technique; and in vivo analysis of retinal morphogenesis and differentiation using 4D microscopy on transgenic lines stably expressing green fluorescent protein (GFP) in different substructures of the retina.

Future projects and goals

We slightly shifted our focus from early events of retinal development towards aspects of retinal differentiation, regeneration and retino-tectal projection. Here we plan to take advantage of the situation in fish that exhibit life-long growth of the retina and correspondingly, the optic tectum. This requires a close control of the balance between proliferation and differentiation. Addressing and understanding that in the context of fish will shed light onto the situation in amniotes, where retinal stem cells are not found in distinct domains. The maintenance of the topographic retino-tectal projection of a growing retina and tectum requires re-connectivity in the optic tectum. We plan to set up subtle genetic screens to address this question that again is of major bio-medical relevance.

Mutants and many other tools established in the lab are currently analysed with the aim of understanding the path that neural progenitors cells take towards terminal differentiation under conditions of growth and regeneration. Along the same line we plan to take advantage of the system to focus on the genetic and molecular basis of axonal re-connectivity as found during retinal growth and projection to the growing optic tectum in fish in vivo.

Selected references

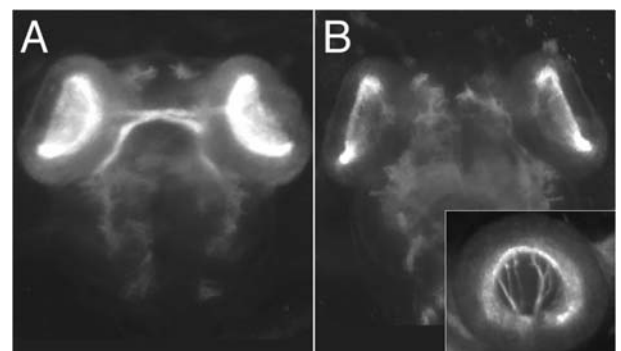
Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A.W. & Wittbrodt, J. (2004). Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science*, 306, 869-871

Del Bene, F., Tessmar-Raible, K. & Wittbrodt, J. (2004). Direct interaction of *geminin* and *Six3* in eye development. *Nature*, 427, 745-749

Loosli, F., Staub, W., Finger-Baier, K.C., Ober, E.A., Verkade, H., Wittbrodt, J. & Baier, H. (2003). Loss of eyes in zebrafish caused by mutation of *chokh/rx3*. *EMBO Rep.*, 4, 894-899

Wittbrodt, J., Shima, A. & Scharl, M. (2002). Medaka – a model organism from the far East. *Nat. Rev. Genet.*, 3, 53-64

Dorsal view of fish lines (medaka) stably expressing GFP in the eyes to visualise all retinal ganglion cells and their axonal projection (A) or only the newly born retinal ganglion cells (B and inset). Large pigments cells show auto-fluorescence.



Gene Expression Unit

The “central dogma” of molecular biology states that “DNA makes RNA makes protein”. This is a summary of the flow of information from the genetic material (DNA) to the effector molecules (RNA and protein) responsible for the implementation of the genetic blueprint. Each step of the pathway of gene expression is complex and, furthermore, subject to regulation. The goals of this Unit are to study the molecular details of the mechanisms of gene expression and its control in eukaryotes. The approaches employed are biochemical, genetic, molecular and cell biological. This powerful combination enables the dissection of even extremely complex processes on the expression pathway. Within the Unit, different groups study gene expression at different levels, and current focus is on the mechanisms and regulation of pre- and post-transcriptional steps of gene expression and on the structure and formation of the nucleus, where many steps of gene expression take place.

Genes are packaged into chromatin, and the template for gene expression is a complex of DNA with many proteins. We study how the composition and structure of chromatin affects gene expression in a global fashion and helps to provide stable patterns of gene expression. In eukaryotes, essentially all RNA species are transcribed in the form of precursor molecules. An essential step in eukaryotic gene expression is therefore the processing of the primary transcripts to their mature form. One aspect of this is the removal of intervening sequences by the process of splicing. Of particular interest is the regulation of splicing such that a single primary transcript can be processed to give rise to mRNAs encoding functionally distinct proteins.

Following processing, the mature RNA species have to be transported from the nucleus to the cytoplasm. This occurs via active processes involving mediators of transport that are specific to subsets of the RNAs produced. These transport processes are under active study in the Unit. Both pre-mRNA processing and nucleocytoplasmic trafficking take place in a structurally complex environment within the nucleus and, in the latter case, in traversing the pore complexes embedded in the nuclear envelope. The principles and mechanisms of assembly of these structures are also topics of research in the Unit.

A final control point, frequently used in eukaryotic cells, is the regulation of gene expression in the cytoplasm. This means regulation either at the level of translation or of messenger RNA stability. Thus the Unit is equipped for the study of gene expression at multiple levels. The importance of this in eukaryotes is that the expression of many genes is controlled at more than one step. In this way, in order to study gene regulation in its entirety, many approaches must be utilised in parallel.



Elisa Izaurralde

PhD 1990, University of Geneva, Switzerland.
Postdoctoral research at EMBL.
Research Group Leader at the University of Geneva, 1996-1999.
Group Leader at EMBL since 1999.

Post-transcriptional regulation of messenger RNA

Previous and current research

The main focus of our research is the elucidation of the molecular mechanisms that regulate gene expression at the post-transcriptional level, with a particular emphasis on the mechanisms of mRNA export, mRNA surveillance and turnover, and RNA silencing. Using a combination of biochemical and genomic approaches together with functional assays in cultured cells we have first focused on the mechanisms underlying the export of mRNA to the cytoplasm. Proteins involved in this process have also been implicated in several other post-transcriptional regulatory circuits. These include the nonsense-mediated mRNA decay (NMD) pathway, a surveillance mechanism that degrades mRNAs with premature translation termination codons, as well as mRNA localisation in the cytoplasm and translational efficiency. Furthermore, recent evidence suggests that the NMD pathway intersects the post-transcriptional RNA silencing pathway (RNAi). This led us to extend our research activities beyond transport to the investigation of additional post-transcriptional pathways: mRNA surveillance and decay, and RNA interference.

Future project and goals

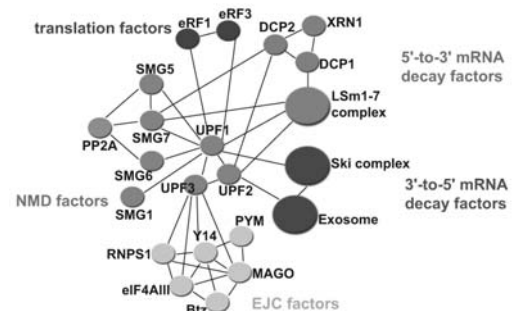
Proteins involved in post-transcriptional gene regulation have been identified biochemically or by genetic screens mainly in *S. cerevisiae*. In order to identify additional components of these pathways in metazoa, we plan to carry RNAi screens in *Drosophila* cells in collaboration with the group of Dr. Michael Boutros (ZMBH, Heidelberg). We expect that, together with biochemical approaches, these screens will yield information on the protein interaction networks involved in post-transcriptional pathways at the genomic level. Our goal is to understand how these networks are interconnected with each other and with different cellular processes.

Having a list of the proteins involved in post-transcriptional regulation and a map of their interactions is not enough to fully understand the molecular mechanisms underlying these processes. We therefore plan to investigate in parallel the function of the individual components of the network.

To determine the contribution of post-transcriptional processes to gene expression on a global scale and to identify endogenous transcripts regulated by these pathways, we are analysing gene expression profiles in cells in which these pathways are inhibited by depletion of essential components. The identification of endogenous targets enables us to obtain information on the cellular processes that are regulated post-transcriptionally.

Our future goals are: 1) to build a comprehensive protein interaction map for post-transcriptional processes; 2) to determine the contribution of these processes to gene expression on a global scale and to identify endogenous targets; and 3) to understand how the regulation of the endogenous targets leads to the complex phenotypes observed at the cellular level.

Protein interaction network involved in NMD in mammals. Proteins for which experimental evidence for a role in NMD is available are shown. Large nodes represent multimeric protein complexes (i.e. the exosome, the Ski complex and the LSM1-7 complex). Note that links do not necessarily represent direct interactions. The exon junction complex (EJC) proteins are linked to this network in mammals, but not in *Drosophila*.



Selected references

- Unterholzner, L. & Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell*, 16, 587-596
- Gatfield, D. & Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature*, 429, 575-578
- Palacios, I.M., Gatfield, D., St Johnston, D. & Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localisation and nonsense-mediated mRNA decay. *Nature*, 427, 753-757
- Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P. & Izaurralde, E. (2003). Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. *EMBO J.*, 22, 3960-3970

Asifa Akhtar

PhD 1998, Imperial Cancer Research Fund, London, UK.
Postdoctoral research at EMBL and
Adolf Butenandt Institut, Munich, Germany.
Group Leader at EMBL since 2001.



Mechanisms of transcription regulation through chromatin

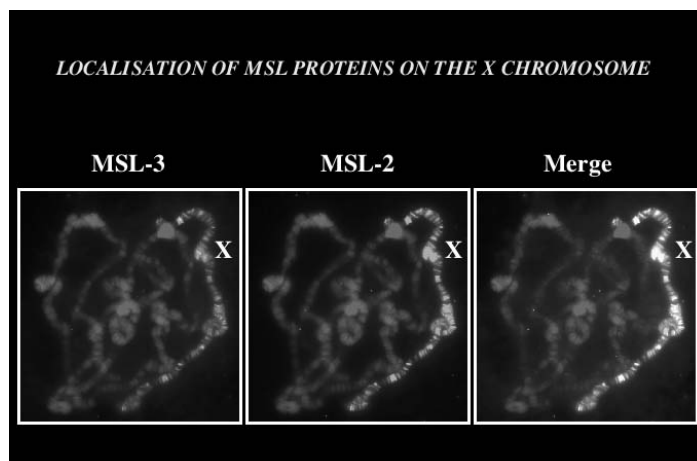
Previous and current research

DNA tightly packed together with histones into nucleosomes is not easily accessible to the enzymes that use it as a template for transcription or replication. Consequently, remodelling of chromatin structure may play an essential role in the regulation of gene expression. Structural changes in chromatin may also form the basis for dosage compensation mechanisms that have evolved to equalize levels of X-linked gene products between males and females. In humans, one of the two X chromosomes in females is randomly inactivated by condensation of the chromosome into a Barr body, a process known as X-inactivation. In contrast, in *Drosophila* this is achieved by a two fold hyper-transcription of the genes on the male X chromosome. Genetic studies have identified a number of factors that are important for dosage compensation in *Drosophila*. This includes five proteins (MSL1, MSL2, MSL3, MLE, MOF) and two non-coding RNAs (roX1 and roX2). The hyperactive X is also specifically hyper-acetylated at histone H4. This acetylation is achieved by the MOF histone acetyl transferase. We have recently shown that localisation of MOF to the X chromosome is RNase sensitive, suggesting a novel role for the non-coding RNAs in bridging the protein/chromatin interactions.

Future projects and goals

It is largely unknown how chromatin modifying factors are targeted to chromatin and it is generally thought to be mediated via a complex network of protein-protein interactions. The involvement of non-coding RNAs as potential targeting molecules adds another level of complexity. We are planning to study the role of non-coding RNAs in chromatin regulation by using evolutionary conserved process of *Drosophila* dosage compensation as a model system. We will investigate how these interactions influence transcription activation of the X-linked genes.

Furthermore, it is intriguing that there is a remarkable evolutionary conservation of all the known *Drosophila* dosage compensation complex members in mammals, even though dosage compensation is brought about by a different means, X chromosome inactivation. A separate aim of the lab is to elucidate the function of the mammalian MSL complex.



Immunostaining of polytene chromosomes from salivary glands of male *Drosophila* using antibodies directed against members of the dosage compensation complex (DCC). The figure shows that MSL-3 and MSL-2 co-localise specifically on hundreds of sites on the male X chromosome. All the chromosomes are also stained with Hoechst to show staining of DNA. The position of the X chromosome is indicated by X.

Selected references

- Taipale, M., Rea, S., Richter, K., Vilar, A., Lichter, P., Imhof, A. & Akhtar, A. (2005). hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol. Cell. Biol.*, 25, 6798-6810
- Buscaino, A., Kocher, T., Kind, J.H., Holz, H., Taipale, M., Wagner, K., Wilm, M. & Akhtar, A. (2003). MOF-regulated acetylation of MSL-3 in the *Drosophila* dosage compensation complex. *Mol. Cell*, 11, 1265-1277
- Akhtar, A. (2003). Dosage compensation: an intertwined world of RNA and chromatin remodelling. *Curr. Opin. Genet. Dev.*, 13, 161-169
- Akhtar, A., Zink, D. & Becker, P.B. (2000). Chromodomains are protein-RNA interaction modules. *Nature*, 407, 405-409
- Akhtar, A. & Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell*, 5, 367-375



Jan Ellenberg

PhD 1998, postdoctoral research at the Cell Biology and Metabolism Branch, NICHD, NIH, Bethesda, USA.
Group Leader at EMBL since 1999.
Joint appointment with the Cell Biology and Biophysics Unit.

Selected references

Lénart, P., Bacher, C.P., Daigle, N., Hand, A.R., Eils, R., Terasaki, M. & Ellenberg, J. (2005). A contractile nuclear actin network drives chromosome congression in oocytes. *Nature*, in press

Rabut, G., Doye, V. & Ellenberg, J. (2004). Mapping the dynamic organisation of the nuclear pore complex inside single living cells. *Nat. Cell Biol.*, 6, 1114-1121

Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R. & Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell*, 112, 751-764

Beaudouin, J., Gerlich, D., Daigle, N., Eils, R. & Ellenberg, J. (2002). Nuclear envelope breakdown proceeds by microtubule-induced tearing of the lamina. *Cell*, 108, 83-96

Daigle, N., Beaudouin, J., Hartnell, L., Imreh, G., Hallberg, E., Lippincott-Schwartz, J. & Ellenberg, J. (2001). Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells. *J. Cell Biol.*, 154, 71-84

Functional dynamics of nuclear structure during cell cycle

Previous and current research

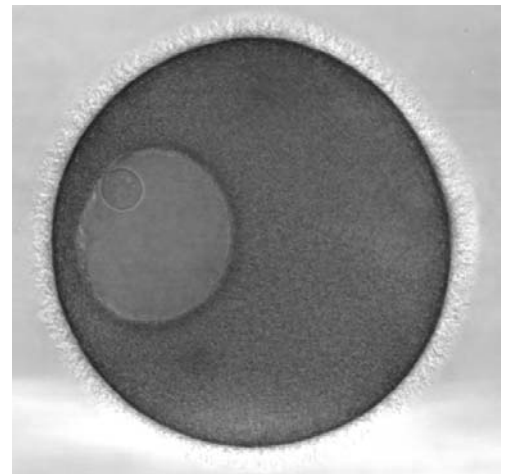
The genome of eukaryotic cells is compartmentalised inside the nucleus, delimited by the nuclear envelope (NE). The NE consists of two concentric membranes continuous with the ER. In metazoans, the nuclear lamina meshwork of intermediate filaments stabilises the NE. The double membrane is perforated by nuclear pore complexes (NPCs), which allow selective traffic between nucleus and cytoplasm. Attached to the nuclear periphery are heterochromatic regions of chromosomes. Many molecular interactions between these four major structural components of the nucleus have been identified. In M-phase, most metazoan cells dismantle the highly ordered topology of the NE resulting in a highly regulated switch in the confinement of the genome. Nuclear membranes that surround chromatin in interphase are replaced by cytoplasmic spindle microtubules, which segregate the condensed chromosomes in an “open” division. After chromosome segregation the nucleus rapidly reassembles.

The overall aim of our research is to elucidate the mechanisms underlying cell cycle remodelling of the nucleus in live cells. Breakdown and reassembly of the nucleus and the formation and correct movement of compact chromosomes are essential but poorly understood processes. To study them, we are assaying fluorescently tagged nuclear structural proteins and their regulators. 4D confocal microscopy is used to directly observe structural dynamics while laser photobleaching/photoactivation methods serve to analyse binding interactions and diffusion of proteins. Because of the complexity of kinetics and geometry in live cells, we then use computer simulations to extract biophysical parameters and build predictive models.

In the past, we showed that NPCs and lamins form a stably interlinked protein network in interphase. We could define the ER as the reservoir and means of partitioning for nuclear membrane proteins in mitosis and found that nuclear breakdown is facilitated by microtubule mediated tearing of the nuclear lamina. During meiotic maturation of oocytes, we demonstrated that partial NPC disassembly is the earliest event of nuclear disassembly and that subsequent long range chromosome motion is driven by actin. In addition, we have analysed mitotic chromosome dynamics and showed that their overall arrangement is transmitted through mitosis.

Future projects and goals

The objective of our future work is to gain further mechanistic insight into nuclear remodelling in live cells. In particular, we are focusing on the mechanism of nuclear growth in interphase, nuclear disassembly and reformation as well as chromosome condensation and positioning in somatic cells and microtubule independent chromosome motion in oocytes. In addition we are using live cell microscopy to identify novel genes that function in the above processes by RNAi screening.



An oocyte from starfish (*Asterina miniata*) before entering meiotic cell division.

Matthias W. Hentze

MD 1984, University of Münster, Germany.
 Postdoctoral research at the NIH, Bethesda, USA.
 Group Leader at EMBL since 1989. Associate Director since 2005.
 Co-Director of the EMBL/University of Heidelberg Molecular Medicine Partnership Unit.



Cytoplasmic gene regulation and molecular medicine

Previous and current research

Important steps in the control of gene expression are executed in the cytoplasm: the regulation of mRNA translation and stability. We are elucidating these regulatory mechanisms, including the mode of function of miRNAs, using mostly biochemical approaches and mammalian, yeast and *Drosophila* model systems. Within the Molecular Medicine Partnership Unit, we are investigating the mammalian post-transcriptional mRNA quality control mechanism “nonsense-mediated decay” (NMD) and its importance in genetic diseases.

Our second major interest is the “systems biology” of mammalian iron metabolism. This work includes the system-wide exploration of the functions of the IRE/IRP regulatory network as well as studies of the molecular basis of genetic and non-genetic diseases of human iron metabolism. This work employs conditional knockout mouse strains for IRP1 and IRP2 and mouse models of iron metabolism diseases. We also use a unique DNA microarray platform (the IronChip) that we have developed.

Future projects and goals

- To uncover the basic mechanisms underlying protein synthesis and its regulation by RNA-binding proteins and miRNAs in cell metabolism, differentiation and development.
- To understand the molecular mechanisms and regulatory circuits to maintain physiological iron homeostasis and its connections to the immune system.
- To contribute to the elucidation of the molecular pathophysiology of common iron overload (haemochromatosis), iron deficiency (anaemia) and iron management (anaemia, Parkinson's disease) disorders.

Selected references

- Beckmann, K., Grskovic, M., Gebauer, F. & Hentze, M.W. (2005). A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in *Drosophila*. *Cell*, in press
- Gebauer, F. & Hentze, M.W. (2004). Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell Biol.*, 5, 827-835
- Thoma, C., Bergamini, G., Galy, B., Hundsdoerfer, P. & Hentze, M.W. (2004). Enhancement of IRES-mediated translation of the c-myc and BiP mRNAs by the poly(A) tail is independent of intact eIF4G and PABP. *Mol. Cell*, 15, 925-935
- Holbrook, J.A., Neu-Yilik, G., Hentze, M.W. & Kulozik, A.E. (2004). Nonsense-mediated decay approaches the clinic. *Nat. Genet.*, 36, 801-808
- Hentze, M.W., Muckenthaler, M.U. & Andrews, N.C. (2004). Balancing acts: molecular control of mammalian iron metabolism. *Cell*, 117, 285-297
- Ostareck, D.H., Ostareck-Lederer, A., Shatsky, I.N. & Hentze, M.W. (2001). Lipoxigenase mRNA silencing in erythroid differentiation: The 3'UTR regulatory complex controls 60S ribosomal subunit joining. *Cell*, 104, 281-290



Andreas Ladurner

PhD 1998, Cambridge University, UK.

Postdoctoral research at the Howard Hughes Medical Institute, University of California, Berkeley, USA.
Group Leader at EMBL since 2003.

Joint appointment with the Structural and Computational Biology Unit.

Chromatin plasticity

Previous and current research

Chromatin packages our DNA into a compact, highly dynamic structure. Distinct post-translational modifications, RNA molecules and specialised histones all affect the behaviour of the fibre. Our lab studies this “epigenetic” level of biological control.

We focus our attention on two important chromosomal structures, the human centromere and the inactive mammalian X chromosome. The first is key to genetic stability, as accurate chromosome segregation is vital. The second provides the molecular basis for gene dosage compensation between males and females (females have two rather than one X chromosome, so need to shut down gene expression on the second chromosome). We address how genes are silenced and how silencing is maintained, determine the role of specialised histones and identify cellular metabolites that regulate chromatin structure. We use a highly interdisciplinary approach to answer fundamental biological questions and focus on tackling novel molecular mechanisms for chromatin plasticity.

In particular, we have generated high-resolution human genomic microarrays to study centromeres. All living entities are propagated by cell division and the proteins that make up the centromere are essential in this complex event. But we do not know how these proteins faithfully decide to locate to a single chromosomal location. The arrays are allowing us to address this central question of centromere identity. Secondly, biochemical approaches are used to identify heterochromatic complexes in *S. pombe*. Proteins rarely act in isolation, forming larger molecular assemblies. It is often in the context of these complexes that proteins reveal their activity and let us understand how they may be recruited to specific places in our genome. We also study (potential) RNAi-related chromatin-targeting complexes in humans.

We are exploring connections between NAD metabolism and chromatin. We have shown that a histone variant which often marks the inactive X chromosome binds NAD metabolites through its macro domain. Our data suggest that chromosomes may be under the control of NAD metabolites, providing a link between metabolism and human chromatin structure.

Future project and goals

- Determination of the molecular connections between NAD metabolism, chromatin and gene silencing.
- Identification of the physiological role of macroH2A histones and mammalian Sir2 deacetylases, which have been linked to obesity, cancer and longevity.
- Identification of how heterochromatin effectors are targeted to both fission yeast and human centromeres, including the role of RNAi in transcriptional gene silencing.
- Exploration of the fundamental structure of human centromeres, their organisation and what targets them to these unique genomic locations.



Female cats often have a marbled coat (left), a result of one of the X chromosomes being silenced. The histone macroH2A is enriched on this chromosome. We have recently shown that this histone (right) binds NAD metabolites, potentially linking metabolism with the regulation of chromatin structure.

Selected references

Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K. & Ladurner, A.G. (2005). Splicing regulates NAD metabolite binding to histone macroH2A. *Nat. Struct. Mol. Biol.*, 12, 624-625

Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M. & Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. *EMBO J.*, 24, 1911-1920

Lejeune, E. & Ladurner, A.G. (2005). Hitting transcription in all the right places. *Nat. Struct. Mol. Biol.*, 12, 390-392

Ladurner, A.G. (2003). Inactivating chromosomes: a macro domain that minimizes transcription. *Mol. Cell*, 12, 1-3

Jürg Müller

PhD 1991, University of Zürich, Switzerland.
 Postdoctoral research at the MRC Laboratory of Molecular Biology, Cambridge, UK.
 Junior Group Leader at the Max Planck Institute for Developmental Biology, Tübingen, Germany.
 Group Leader at EMBL since 2001.



Mechanisms of transcriptional regulation in development

Previous and current research

Our laboratory studies the molecular mechanisms by which trithorax group (trxG) and Polycomb group (PcG) proteins maintain transcriptional ON and OFF states of target genes. PcG and trxG proteins are conserved in both animals and plants where they are essential for diverse developmentally regulated processes, ranging from the maintenance HOX gene expression patterns in animals, X-chromosome inactivation in mammals to the control of seed development and flowering time in plants. We use *Drosophila* as a model system and much of our work has been focused on understanding the mechanisms by which PcG and trxG regulators maintain HOX gene expression patterns. PcG and trxG proteins control gene expression at the level of chromatin in ways that are still poorly understood. Most PcG proteins do not bind to DNA directly but they do bind to chromatin; one major aim of our work has therefore been to understand how PcG and trxG proteins are specifically targeted to the genes that they regulate. Another focus has been the functional characterisation of PcG protein complexes using in vitro and in vivo approaches with the aim of elucidating the molecular mechanisms by which these protein complexes modify chromatin and repress transcription of target genes.

Future projects and goals

We will continue to use HOX gene regulation in *Drosophila* as a framework to gain mechanistic insights into PcG and trxG protein function. Biochemical purification and forward genetic screening strategies to identify novel components of the PcG/trxG system, combined with in-depth in vitro and in vivo analyses of newly identified proteins, are central to this part of our work. A second focus will be to identify novel target genes of the PcG/trxG system using functional genomic tools.

Our long-term goal is to understand how transcriptional ON and OFF states controlled by the PcG/trxG system are propagated through replication and cell division.

Selected references

- Nekrasov, M., Wild, B. & Müller, J. (2005). Nucleosome binding and histone methyltransferase activity of *Drosophila* PRC2. *EMBO Rep.*, 6, 348-353
- Sengupta, A.K., Kuhrs, A. & Müller, J. (2004). General transcriptional silencing by a Polycomb response element in *Drosophila*. *Development*, 131, 1959-1965
- Klymenko, T. & Müller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.*, 5, 373-377
- Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E. & Simon, J.A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell*, 111, 197-208



Carsten Schultz

PhD 1989, University of Bremen, Germany.

Postdoctoral research at the University of California, San Diego, USA.

Habilitation 1997, Organic Chemistry, University of Bremen, Germany.

Group Leader MPI for Molecular Physiology, Dortmund, Germany. Group Leader at EMBL since 2001.

Bio-organic chemistry of signalling molecules

Previous and current research

In the past, our research focused on finding novel ways to stimulate chloride and water secretion of epithelial cells due to their crucial role in the genetic disease cystic fibrosis (CF). Our compounds helped to investigate some of the underlying intracellular signalling pathways and provided drug candidates that effectively increased chloride secretion and reduced water uptake of nasal epithelia from CF patients *ex vivo*. We developed chemical methods to convert highly polar signalling molecules like cyclic nucleotides, inositol phosphates and phosphoinositides to membrane-permeant, bioactivatable derivatives. These compounds ("pro-drugs" or "Trojan horse compounds") help to deliver polar compounds to the cytosol. With the resulting tools we were able to elevate the concentration of a signalling molecule of interest inside living cells without disrupting the plasma membrane. This technology was successfully extended to drug candidates based on these signalling molecules.

The goal now is to visualise the signalling pathways involved in chloride secretion by fluorescent probes. By doing so, we hope to provide a more complete picture of the signalling network and to create novel ways to screen compounds that might be beneficial for CF patients. The function of the probes is based on fluorescent resonance energy transfer (FRET) and is suitable for ratio imaging and confocal microscopy. The three approaches we pursue are depicted in the illustration. They are currently combined in an approach called Multiparameter Imaging, where 5-6 cellular events are monitored simultaneously. This work is funded by the "Molecular Imaging" IP in FP6 and is a joint effort with Dorus Gadella (Amsterdam).

Small molecule fluorescent FRET probes are prepared to study intracellular enzyme activities with a focus on phospholipases. Recently we prepared a very effective probe to monitor phospholipase A2 activity in cells and small organisms. Studies on the development of fish embryos were performed in collaboration with Jochen Wittbrodt (EMBL) and are funded by the DFG and the EU.

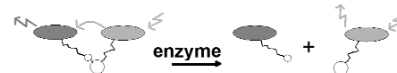
To study the interaction of phospholipids with lipid-binding proteins we are preparing fluorescently labeled phospholipids that serve as FRET partners for fluorescent fusion proteins. The labeled lipids are membrane-permeant and allow lipid translocation events to be studied.

A third approach is based on conformational changes of double fluorescently labeled proteins that lead to a change in FRET. With these reporter probes we monitor several phosphorylation/dephosphorylation events as well as transcription factor activity in living cells. These projects are in collaboration with the groups of Conti, Sattler and Gannon (EMBL) and Gautel (London).

Future projects and goals

In 2005 a new collaboration started to focus on visualising protein metabolism in living cells. This research is performed together with Philippe Pierre (Marseille), Manuel Santos (Aveiro, Portugal) and Michel Desjardins (Montreal) and is funded by the HFSP.

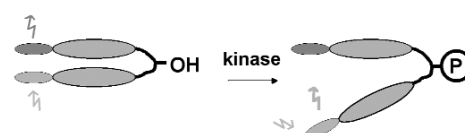
small molecules
as enzyme substrates



small molecule-
protein interaction



protein-protein
interaction,
conformational
changes



Selected references

Schultz, C., Schleifenbaum, A., Goedhart, J. & Gadella, T.W.J. (2005). Multiparameter imaging for the analysis of intracellular signalling. *Chembiochem.*, in press

Schleifenbaum, A., Stier, G., Gasch, A., Sattler, M. & Schultz, C. (2004). A genetically encoded FRET probe for PKC activity based on pleckstrin. *J. Am. Chem. Soc.*, 126, 11786-11787

Schultz, C. (2003). Prodrugs of biologically active phosphate esters. *Bioorg. Med. Chem.*, 11, 885-898

Lars M. Steinmetz

PhD 2001, Stanford University, USA.
Postdoctoral research at the Stanford
Genome Technology Center, USA.
Group Leader at EMBL since 2003.



Functional genomics of complex traits and pathways

Previous and current research

Little is known about the genetic factors underlying complex traits in humans and in other organisms, yet most natural phenotypes and diseases belong to this class. We use and develop functional genomic approaches and high-throughput methods to identify the genetic variants that underlie complex traits and pathways in yeast, and apply our findings to humans.

Using oligonucleotide microarrays we can detect and genotype single nucleotide polymorphisms at high-density over the entire yeast genome. We have used this technology to map at high-resolution the complex, quantitative trait loci (QTL) for virulence traits of high-temperature resistant pathogenic *Saccharomyces cerevisiae*. We have developed a new technology, Reciprocal Hemizyosity Scanning (RHS), that allows the contribution to phenotype to be determined for all alleles from the genomes of two independent strains. Using this technology, we will map and quantitate the contribution of all phenotypically relevant alleles for any complex trait in a single tube assay. These technologies will allow us to explore genetic contributions to quantitative traits under a variety of environmental conditions and to investigate ecology at the genome level.

We use the recently completed collection of yeast single gene deletion strains to characterise gene function on a genomic scale. In each strain, one gene is deleted and replaced with a cassette containing a kanamycin resistance gene for selection purposes and two molecular barcode tags. The tags allow strain detection with a high-density oligonucleotide microarray, containing sequences complementary to each tag. The entire collection of deletions strains can be grown as a pool in a single culture under different environmental conditions and provides an opportunity to quantitate the contribution to fitness for each gene in the genome. We have used this collection to identify genes involved in mitochondrial function. By identifying human orthologs, we determined new candidate genes for human putative-mitochondrial disorders. We are expanding our understanding of the proteins involved in mitochondria using a combination of computational and functional genomic methods in yeast, including whole organelle proteome and transcriptome analyses. These approaches will allow the study of the mitochondrial organelle at a system level and speed the discovery of disease-related genes in humans.

Future projects and goals

We are interested in continuing to explore ways to unravel the genetic complexity of quantitative traits in yeast and to bridge the gap between model systems biology and medicine. In particular we will focus on applying our yeast mitochondrial studies to human proteins and testing mitochondrial candidate genes in patients. We also continue to explore ways to apply our functional genomic dataset to fundamental questions in evolutionary biology. Most recently, we have analysed the genome sequence of a clinical isolate of yeast and found lower evolutionary rates than in the laboratory strain of yeast, providing evidence that the laboratory strain may have evolved faster than some of its wild counterparts.

Selected references

- Gu, Z., David, L., Petrov, D., Jones, T., Davis, R.W. & Steinmetz, L.M. (2005). Elevated evolutionary rates in the laboratory strain of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, 102, 1092-1097
- Prokisch, H., Scharfe, C., Camp DG, 2nd, Xiao, W., David, L., Andreoli, C., Monroe, M.E., Moore, R.J., Gritsenko, M.A., Kozany, C., et al. (2004). Integrative analysis of the mitochondrial proteome in yeast. *PLoS Biol.*, 2, e160
- Steinmetz, L.M. & Davis, R.W. (2004). Maximizing the potential of functional genomics. *Nat. Rev. Genet.*, 5, 190-201
- Steinmetz, L.M., Scharfe, C., Deutschbauer, A.M., Mokranjac, D., Herman, Z.S., Jones, T., Chu, A.M., Giaever, G., Prokisch, H., Oefner, P.J. & Davis, R.W. (2002). Systematic screen for human disease genes in yeast. *Nat. Genet.*, 31, 400-404
- Steinmetz, L.M., Sinha, H., Richards, D.R., Spiegelman, J.I., Oefner, P.J., McCusker, J.H. & Davis, R.W. (2002). Dissecting the architecture of a quantitative trait locus in yeast. *Nature*, 416, 326-330



Matthias Wilm

MD 1991, University of Münster, Germany.

PhD 1996, Physics, EMBL and the University of Münster, Germany.

Postdoctoral research at the University in Düsseldorf, Germany and at EMBL.

Group Leader at EMBL since 1998. Joint appointment with the Gene Expression Unit.

Molecular analysis by mass spectrometry in biological research

Previous and current research

The ability to identify proteins with a very high sensitivity scale with mass spectrometry has profoundly influenced biological research. The direct exploration of the biological systems on the protein level has become a reality. The ongoing challenge in proteomic research is to develop appropriate techniques and to use them in a beneficial way.

Our past research was influenced by the efforts to reveal the function of proteins within their biological context (functional proteomics). In 1999 Bertrand Séraphin, at that time Group Leader in the Gene Expression Unit, published with our support the “Tandem Affinity Purification” (TAP) technique, a high-quality method to extract functional protein complexes from total cell lysates of yeast cells. The ability to purify very specifically non-covalent protein complexes is a key technology for assigning a functional context to proteins. This is because protein complexes represent a major principle of order within a cell. Proteins evolved in a way that they have affinity for other proteins that are needed to fulfill their function. We focused our efforts on characterising as many proteins as possible coming from TAP based affinity purifications. In nearly all cases, we revealed proteins involved in the targeted biological pathways. This project served as the starting point for a spin-off company (Cellzome, <http://www.cellzome.com>). The work of this company led to the first characterisation of an organism-wide network of functionally related protein complexes.

It was not easily possible to use the TAP technique with cells from higher eukaryotic organisms. The original technique made use of yeast-dependent genetic techniques (homologue recombination in haploid strains). In 2002 we introduced the iTAP technique, a modification that demonstrated how TAP can be successfully applied to cells from higher eukaryotic organisms. The key was to combine it with RNAi based knock-downs of the endogenous protein. We were able to extend the principle to describe an organism as a dynamic network of interconnected protein complexes to higher organisms including human cells on an experimental level.

Future projects and goals

In the coming years we will focus our activity more towards quantitative proteomics. The activity of a biological system is not only defined by the identity of the proteins contributing to it, but by their quantity and the degree of their modifications. We will concentrate on revealing protein sequence, the state of the secondary modifications and their quantities in proteomic samples.

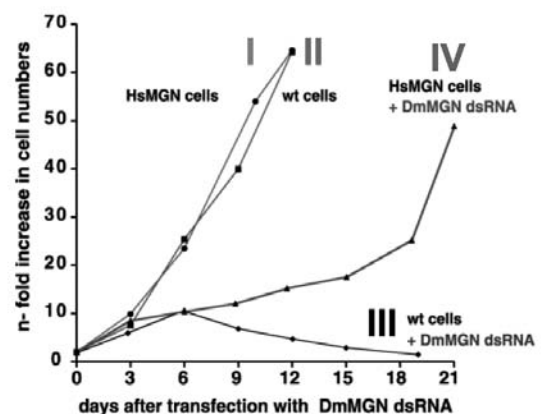
Selected references

Forler, D., Kocher, T., Rode, M., Gentzel, M., Izaurralde, E. & Wilm, M. (2003). An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat. Biotechnol.*, 21, 89-92

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. (2002). Functional organisation of the yeast proteome by systematic analysis of protein complexes. *Nature*, 415, 141-147

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. & Séraphin, B. (1999). A generic protein purification method for protein complex characterisation and proteome exploration. *Nat. Biotechnol.*, 17, 1030-1032

Growth curves of cells when applying the iTAP technique to an essential protein in *Drosophila*. Curve IV shows that the iTAP method favours the selection of cells that allow the specific purification of complexes containing the tagged form of the protein.



Matthias W. Hentze (see page 33 for biography), **Andreas E. Kulozik**

MD 1984, University of Münster, Germany.
PhD (CNAO) 1987, MRC Molecular Haematology Unit, Oxford University, UK.
Head of the Department of Pediatric Oncology, Hematology
and Immunology, University of Heidelberg since 2001.
Co-Director of the EMBL/University of Heidelberg Molecular Medicine Partnership Unit since 2002.



EMBL/University of Heidelberg Molecular Medicine Partnership Unit: Diseases of mRNA metabolism

Previous and current research

To ensure the accuracy of gene expression, cells have developed a variety of control mechanism for the quality of gene expression, which identify and eliminate aberrant messenger RNAs. Interestingly, termination codons upstream of the normal stop codon destabilise the affected transcript. This underlying molecular process, referred to as nonsense-mediated mRNA decay, has been the focus of our scientific interests for a number of years. The biological and medical implications of NMD are highlighted by nonsense mutations in the β -globin gene which lead to β -thalassemia. If the nonsense mutation is located at a NMD competent position, the mRNA is degraded and heterozygotes remain healthy. When the nonsense mutation is located at a position that does not activate the NMD pathway, the mRNA is stable and translated into a C-terminally truncated protein that acts in a dominant negative fashion. In these cases, heterozygotes are symptomatic and the thalassemia is inherited in a dominant mode. The effect of NMD in human diseases is apparent in many more inherited and acquired genetic disorders and has recently become a target for therapeutic interventions. Our current research is directed towards the molecular mechanism of NMD and how NMD interconnects with other processes of RNA processing to regulate gene expression on a post-transcriptional level. An understanding of such regulatory networks will be important to further develop the biological and medical dimension of NMD.

In a second project, our group is characterising the molecular architecture of the prothrombin 3' end formation signal that represents a hotspot for thrombophilia-causing mutations. A common mutation (G>A) at position 20210 of the prothrombin gene (blood coagulation factor II, F2) affects about 2% of the population of northwest Europe and increases the risk of thrombosis in heterozygous carriers about 3-5 fold. We have previously shown that the normal (unmutated) prothrombin mRNA is inefficiently processed at the 3' end. The 20210 G>A mutation causes an upregulation of 3' end processing efficiency, which represented a novel principle for a gain-of-function mutation involved in the pathogenesis of a hereditary disorder. We have subsequently shown that the physiological 3' end formation efficiency is balanced by a characteristic architecture of the prothrombin mRNA processing signal with stimulatory and inhibitory cis-acting elements.

Future projects and goals

The mechanistic work in the NMD project is directed towards an understanding of the functional network of nuclear and cytoplasmic proteins that assemble on the mRNA and either trigger the mRNA decay pathway or specify a stable mRNA. This work involves the analysis of the functional hierarchy of the proteins, the assembly of NMD-activating mRNPs and the characterisation of potential regulatory entry points into the NMD pathway.

In a more applied approach of NMD research, we currently analyse how far NMD efficiency differs between individuals and how far it can be modulated by external stimuli. This part of the project will generate information about the physiological role of NMD as a regulated process and about its role as a modifying factor in human genetic diseases.

In the prothrombin project, we aim to establish a mechanistic understanding of 3' end mRNA processing of this key transcript. We currently analyse the proteins that bind to the characteristic inhibitory and stimulatory elements in the prothrombin 3' end formation signal. We aim to understand the physiological regulation of the post-transcriptional processing of this unusual mRNA and how this can be disturbed in thrombophilia or under conditions of cellular stress.

Selected references

- Danckwardt, S., Gehring, N.H., Neu-Yilik, G., Hundsdoerfer, P., Pförsich, M., Hentze, M.W. & Kulozik, A.E. (2004). The prothrombin 3' end formation signal reveals a unique architecture that is sensitive to thrombophilic gain-of-function mutations. *Blood*, 104, 428-435
- Holbrook, J., Neu-Yilik, G., Hentze, M.W. & Kulozik, A.E. (2004). Nonsense mediated decay approaches the clinic. *Nat. Genetics*, 36, 801-809
- Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W. & Kulozik, A.E. (2003). Y14 and hUpf3b form an NMD-activating complex. *Mol. Cell*, 11, 939-949
- Gehring, N.H., Frede, U., Neu-Yilik, G., Hundsdoerfer, P., Hentze, M.W. & Kulozik, A.E. (2001). Increased efficiency of mRNA 3' end formation: a novel genetic mechanism contributing to hereditary thrombophilia. *Nat. Genetics*, 28, 389-392

Structural and Computational Biology Unit

Molecular information is required for the understanding of any higher organisation in a cell or organism. The opposite is also true: the biology of live molecules is ultimately understood only in the cellular context. Molecular biology is undergoing a very rapid expansion during the first decade of the 21st century. The various causes of this new explosion include the current systematic exploration of bacterial and eukaryotic genomes, development of rapid screening methods for gene expression patterns and related information and other technical improvements, such as applications of mass spectrometric analysis of large complexes.

Structural and Computational biology is an integral part of this development. On the one hand, it will provide three-dimensional information on RNAs and proteins that will constrain and solidify research hypotheses and make their predictions more reliable. Knowledge of gene structures (sequences) will eventually be complemented by structural principles of macromolecules: galleries of possible topologies, folding rules, modes of interaction and molecular adaptation to different functions and temperatures. This will be an enormous task for decades to come. The experimental technology used by structural biology develops rapidly and provides tools that can be employed in a flexible and creative manner to study molecular processes within and between cells.

On the other hand, computational approaches will become more of an indispensable tool in any molecular biology laboratory. From sequence analysis to the modelling of complex biological systems, more and more tools are being developed that will contribute to the understanding of how a cell or organism operates. The Structural and Computational Biology Unit includes groups using classical methods of structural analysis such as X-ray crystallography, protein NMR and electron microscopy as well as groups specialising in theoretical approaches. More recently we have introduced 3D tomography, which, in the future and in combination with other structural techniques, could allow us to obtain images of a cell at atomic resolution. There is a continuing interplay between these different methodologies, reflecting the belief that a combination of structural and functional studies is the most rewarding route to an understanding of the molecular basis of biological function.

Researchers in the Unit have a solid background in structural and computational methods and in the principles underlying protein and nucleic acid structure; this is consolidated through weekly informal seminars and journal clubs. The experimental approaches are dependent upon the availability of the appropriately prepared material, and the molecular biology methods required for this have also been integrated into the research groups. Theoretical analysis of genomic data, protein expression data, gene networks and macromolecular structures on all levels is an important component of our research. We find that close contact between the experimental and computational groups is a very valuable intellectual resource. For instance, studies within the groups engaged with sequence comparison and genomics have been fundamental to the initiation of new projects, particularly in the NMR laboratory.

The Unit is very well equipped for experimental and computational work. Experimental facilities include area detectors for the collection of X-ray diffraction data, a 500 MHz and 600 MHz NMR spectrometer, transmission electron microscopes and scanning microdensitometers, as well as facilities for electron cryo-microscopy, cryo-3D tomography crystallography, CD and fluorescence spectroscopy, analytical ultracentrifugation and for large scale growth of prokaryotes and eukaryotes. The central computing is organised around a UNIX cluster. The whole computing environment is conveniently networked and excellent facilities for high performance computing and computer graphics are available.

Luis Serrano

PhD 1985, Universidad Autónoma, Madrid, Spain.
Postdoctoral research at Universidad Autónoma, Madrid, Spain,
and at the MRC Unit for Protein Engineering, Cambridge, UK.
Group Leader at EMBL since 1992. Unit Coordinator since 2001.



Synthetic biology: designing biological systems from proteins to cells

Previous and current research

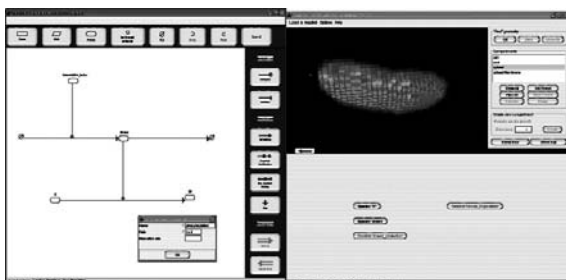
Our group is interested in engineering and designing biological systems. These can range from a protein to a cell, though the characteristic they have in common is that the total is more than the sum of the parts. We believe that the combination of theoretical and experimental approaches is the most fruitful way to address this complex goal; as a result, the group has strong computational expertise to develop software tools to simulate and engineer proteins or gene networks, and strong experimental expertise that validates our predictions as well as offering data and insight to our theoretical work.

Currently we are working on two main lines of research, one dealing with protein design and the other with gene network simulation. Recently the two have started to converge.

- On the protein design side we have developed software called FOLD-X that automatically designs proteins. We are using this software and structural information not only to modify proteins to introduce new properties but also to predict the interaction partners across the genome. In this way we can predict the binding partners of numerous proteins as well as estimating binding constants and Kon and Koff parameters, which are essential for the understanding from a systems biology approach how gene and protein networks operate. As a side project we have continued to investigate the relationship between sequence and amyloidogenesis, and we have established rules that determine the amyloido-geneicity of a particular protein sequence.
- On the gene network side we are building software called SmartCell that allows simulation of gene networks, taking space into consideration. To make the simulation tool accessible to wet labs, we have designed a user-friendly graphic interface that allows the scientists to draw reactions. In collaboration with other groups at EMBL, we have also started to integrate 3D detailed reconstruction of cells.

Future projects and goals

- To develop FOLD-X to predict protein nucleotide interactions, as well as using it for drug design.
- To automate the modelling and prediction of protein partners using structural information and apply it to in silico prediction of protein-protein and protein-DNA/RNA networks.
- To integrate tomograms of different cell types in SmartCell with links to databases, as well as to integrate other simulation approaches so that the user can choose the one more appropriate for his/her problem.
- To create tools to design and engineer networks in cells (synthetic biology).
- To eventually integrate FOLD-X and SmartCell to create a piece of software which allows users to design not only proteins but also networks.



Two pages of SmartCell GUI; left, the page used to draw the reactions; right, the page displaying the 3D voxelized shape of a cell where the user can localise the different reactions.

Selected references

- Isalan, M., Lemerle, C. & Serrano, L. (2005). Engineering gene networks to emulate Drosophila embryonic pattern formation. *PLoS Biol.*, 3, e64
- Fernandez-Escamilla, A.M., Rousseau, F., Schymkowitz, J. & Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.*, 22, 1302-1306
- Aloy, P., Böttcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.C., Bork, P., Superti-Furga, G., Serrano, L. & Russell, R.B. (2004). Structure-based assembly of protein complexes in yeast. *Science*, 303, 2026-2029
- Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M., Kaplan, A., Lemerle, C., Tomas-Oliveira, I. & Serrano, L. (2004). SmartCell: a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple gene networks. *Systems Biology*, 1, 129-139



Peer Bork

PhD 1990, University of Leipzig, Germany.
Habilitation 1995, Humboldt University, Berlin, Germany.
At EMBL since 1991.
Joint Unit Coordinator since 2001.

Predicting function in biological systems

Previous and current research

The main goal of this group is to predict function and gain insight into evolution by comparative analysis. The group currently predicts functionality at three different levels: genes and proteins; protein networks and cellular processes; and phenotypes.

The topics require both the development of methods and their applications. A selection of projects is given below.

- Prediction of structural and functional features of proteins.
- Prediction of pathways and cellular systems.
- Comparative analysis of genes, genomes and metagenomes.
- Evolution of genome content and networks.
- Impact of alternative splicing on function.

Future projects and goals

Our future projects include:

- modelling of biochemical pathways and networks;
- understanding molecular function in space and time.

Our long-term goals remain the conversion of molecular data into biological knowledge and the understanding of evolution to improve comparative methods. We want to achieve this by interacting closely with experimental groups whereby cycles of data generation and analysis as well as hypothesis generation and verification are required.

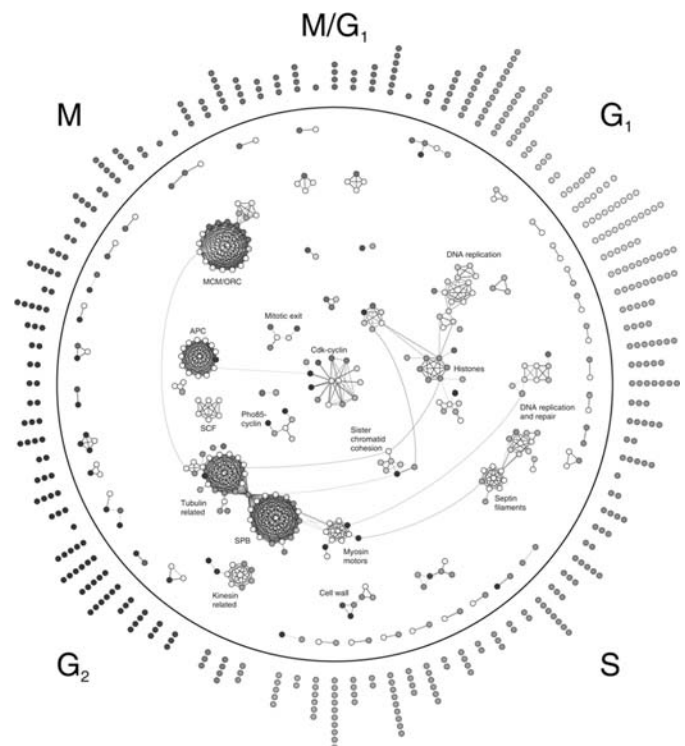
Selected references

Korbel, J.O., Doerks, T., Jensen, L.J., Perez-Iratxeta, C., Kaczanowski, S., Hooper, S.D., Andrade, M.A. & Bork, P. (2005). Systematic association of genes to phenotypes by genome and literature mining. *PLoS Biol.*, 3, e134

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter, J.C., Bork, P., Hugenholtz, P. & Rubin, E.M. (2005). Comparative metagenomics of microbial communities. *Science*, 308, 554-557

de Lichtenberg, U., Jensen, L.J., Brunak, S. & Bork, P. (2005). Dynamics of complex formation during the yeast cell cycle. *Science*, 307, 724-727

Temporary interaction networks and dynamic complex formation during yeast cell cycle. 600 cell cycle regulated proteins in yeast (shaded) dots as identified from microarray data interact with noncyclic scaffolding proteins (white).



Bettina Böttcher

PhD 1991, University of Stuttgart, Germany.
Postdoctoral research at the MRC Laboratory of Molecular Biology, Cambridge, UK.
Assistant Professor at Freiburg University, Germany.
Group Leader at EMBL since 1999.



Electron microscopy and image reconstruction of biological macromolecules

Previous and current research

Electron microscopy and image processing is one of the most powerful tools in the structural investigation of large biological complexes. Insights into the structural organisation of the complexes can be gained without the necessity of crystallising the complex. We exploit these methods to investigate a variety of different complexes within 3D-Repertoire (EU NOE). These investigations are aimed at developing a comprehensive picture of the stable complexes within a yeast cell and provide detailed structural information which can be used as a spatial scaffold on which protein protein interactions can be modelled. Many of the yeast complexes have very low abundance, which make structural investigations extremely challenging and leave electron microscopy often as the only option for obtaining structural information. Alternative to the purification of the native complexes, we build complexes from over expressed protein either by co-expressing subunits or mixing them after expression. These in vitro complexes can usually be purified in large quantities, enabling crystallography as well as detailed characterisation of interaction by different biophysical methods providing exact knowledge on shape, stoichiometry, secondary structure and binding constants. Besides providing a basis for structural modelling this information is also vital for simulating protein-protein interactions in Systems Biology.

Another fascinating aspect of our work is focused on the investigation of structure and dynamics of virus capsids by electron microscopy and image processing. These capsids are often regarded as rigid containers for a genome. However, this perception does not at all describe the versatile role of a capsid. Our research is aimed on demonstrating the structural dynamics of virus capsids by observing capsids in different states of their life cycle or actively challenging the capsid structure by introducing limited conformational stress. So far, we have observed a variety of conformational response mechanisms, which provide structural insight into underlying signalling mechanisms.

Future projects and goals



Our general goal is to understand the structural organisation of protein complexes and their interplay within the cell. Therefore the projects are either aimed at obtaining as detailed structural information of a single complex as possible or at elucidating the interaction between complexes in a native environment and their impact on cell morphology. Here our focus is mainly on membrane bound complexes, which often form super complexes that have a significant impact on membrane morphology.

Electron microscopy and image analysis provides a 3D- map of complex, which serves as a scaffold to which known atomic structures of smaller sub complexes can be fitted, generating a pseudo atomic model of the complete complex. Here, as an example, the ATPsynthase from chloroplasts is shown.

Selected references

- Venzke, D., Domgall, I., Kocher, T., Fethière, J., Fischer, S. & Böttcher, B. (2005). Elucidation of the stator organisation in the V-ATPase of *Neurospora crassa*. *J. Mol. Biol.*, 349, 659-669
- Fethière, J., Venzke, D., Diepholz, M., Seybert, A., Geerlof, A., Gentzel, M., Wilm, M. & Böttcher, B. (2004). Building the stator of the yeast vacuolar-ATPase: specific interaction between subunits E and G. *J. Biol. Chem.*, 279, 40670-40676
- Aloy, P., Böttcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.C., Bork, P., Superti-Furga, G., Serrano, L. & Russell, R.B. (2004). Structure-based assembly of protein complexes in yeast. *Science*, 303, 2026-2029
- Mellwig, C. & Böttcher, B. (2003). A unique resting position of the ATP-synthase from chloroplasts. *J. Biol. Chem.*, 278, 18544-18549



Elena Conti

PhD 1996, Imperial College, London, UK.

Postdoctoral research at the Rockefeller University, New York, USA.

Group Leader at EMBL since 1999.

Joint appointment with the Gene Expression Unit.

Selected references

Cook, A., Fernández, E., Lindner, D., Ebert, J., Schlenstedt, G. & Conti, E. (2005). The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding. *Mol. Cell*, 18, 355-367

Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E. & Conti, E. (2005). SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol. Cell*, 17, 537-547

Jinek, M., Rehwinkel, J., Lazarus, B.D., Izaurralde, E., Hanover, J.A. & Conti, E. (2004). The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin α . *Nat. Struct. Mol. Biol.*, 11, 1001-1007

Bono, F., Ebert, J., Unterholzner, L., Guttler, T., Izaurralde, E. & Conti, E. (2004). Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. *EMBO Rep.*, 5, 304-310

X-ray crystallographic studies of nucleo-cytoplasmic transport

Previous and current research

Our laboratory is interested in the molecular mechanisms that govern the transport of nuclear proteins and RNAs from their site of synthesis to their site of function, a process that allows communication to occur between cytosolic translation and nuclear replication and transcription in eukaryotes. We are also interested in the connections of nuclear transport to upstream and downstream processes. In the case of nuclear export, for example, there is increasing evidence that proteins involved in mRNA transport are linked to mRNA processing and surveillance mechanisms. To obtain molecular insights into these cellular processes, we use X-ray crystallography in combination with other biophysical and biochemical methods.

Nucleo-cytoplasmic transport relies on the recognition of targeting signals by a set of transport factors. We are working on transport factors of the karyopherin- β family (such as the protein Cse1) that mediate the bulk of nuclear import/export. In the case of mRNA, nuclear export is connected to a complex series of events so that only correct and properly spliced mRNAs are translated into proteins. A crucial role in this context is played by the exon-exon junction complex (EJC), a multiprotein complex deposited on mRNA upon splicing. In humans, EJC proteins such as Mago and Y14 function together with the SMG proteins to mediate nonsense-mediated mRNA decay (NMD). NMD is a quality control mechanism that recognizes mRNAs with premature stop codons and targets them for rapid degradation, thus avoiding the synthesis of truncated and potentially harmful protein products.

Future projects and goals

Our goal is to understand the mechanisms with which macromolecular complexes in the nucleus and the cytosol interact with each other, ensuring the continuous flow of information that is essential for eukaryotic life. In particular, we will study the mRNA decay machinery, both in terms of premature stop codon recognition (i.e. the exon junction complex) and in terms of RNA degradation (i.e. the exosome complex). We also plan to continue our studies on export mediators, in particular on the transport factors that export tRNAs and pre-micro RNAs, and more generally on the connection of nuclear transport to mitotic entry.



Structure of the nuclear export receptor Cse1 captured in the closed ring-like conformation it assumes in the cytosol in the absence of RanGTP.

Achilleas S. Frangakis

PhD 2001, Max Planck Institute of Biochemistry,
Martinsried, Germany.
Postdoctoral research at Caltech, Pasadena, California, USA.
Group Leader at EMBL since 2003.



Electron tomography

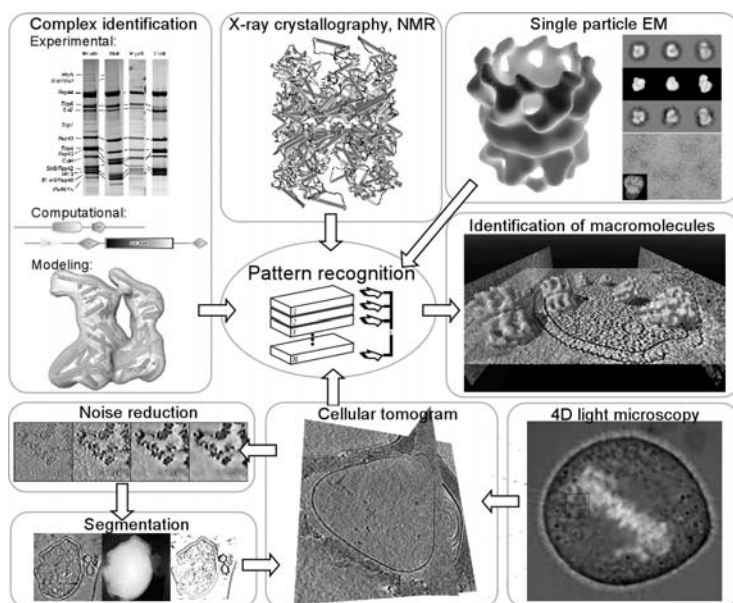
Previous and current research

The main focus of the group is to reveal the macromolecular organisation of living cells by means of cryo-electron tomography. Cryo-electron tomography is the only technique that can obtain molecular resolution images of intact cells in a quasi-native environment. The tomograms contain an imposing amount of information; they are essentially a three-dimensional map of the cellular proteome and depict the whole network of macromolecular interactions. Information mining algorithms exploit structural data from various techniques, identify distinct macromolecules and computationally fit atomic resolution structures in the cellular tomograms, thereby bridging the resolution gap.

A multitude of biological questions can be answered by electron tomography; visualisation of the cellular structure at molecular resolution is largely uncharted territory. The group works with a wide spectrum of specimens, including prokaryotic and eukaryotic cells, but also model systems. Prokaryotic cells are smaller and can therefore be easily penetrated by the electrons. Eukaryotic cells have compartments in which the protein density is lower, facilitating pattern recognition techniques. Model systems, on the other hand, are particularly helpful in improving computational algorithms and in providing solutions for cell systems that are too complex to be investigated by electron tomography.

Future projects and goals

Achieving these goals will enable us to visualise macromolecules in an unperturbed cellular environment and to chart the network of interactions underlying cellular functions. This aim of the group is to prove that a cell is not an envelope of freely diffusing enzymes and substrates, but rather a highly organised and coordinated machine.



The scheme visualises the ultimate goal of the group and the true power of cryo-electron tomography. We are using cryo-electron tomography in conjunction with pattern recognition techniques in order to match atomically resolved structures in the context of living cells. Practically, we integrate the information from X-ray crystallography, structural genomics and single-particle electron microscopy in order to computationally search for macromolecular complexes in the three-dimensional cryo-electron tomograms.

Selected references

- Medalia, O., Weber, I., Frangakis, A.S., Nicastro, D., Gerisch, G. & Baumeister, W. (2002). Macromolecular architecture in eukaryotic cells visualised by cryoelectron tomography. *Science*, 298, 1209-1213
- Frangakis, A.S., Bohm, J., Forster, F., Nickell, S., Nicastro, D., Typke, D., Hegerl, R. & Baumeister, W. (2002). Identification of macromolecular complexes in cryoelectron tomograms of phantom cells. *Proc. Natl. Acad. Sci. USA*, 99, 14153-14158
- Frangakis, A.S. & Hegerl, R. (2001). Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion. *J. Struct. Biol.*, 135, 239-250
- Bohm, J., Frangakis, A.S., Hegerl, R., Nickell, S., Typke, D. & Baumeister, W. (2000). Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms. *Proc. Natl. Acad. Sci. USA*, 97, 14245-14250



Toby Gibson

PhD 1984, Cambridge University, UK.
Postdoctoral research at the Laboratory of
Molecular Biology, Cambridge, UK.
Team Leader at EMBL since 1986.

Biological sequence analysis

Previous and current research

The group seeks to gain insight through the computational analysis of biological molecules, particularly at the sequence level. To this end, we deploy many sequence analysis methods and look to develop new tools as the need arises. Where possible, we contribute to multidisciplinary projects involving structural and experimental groups at EMBL.

We are probably best known for our involvement with the Clustal W and Clustal X programs that are widely used for multiple sequence alignment. We work closely with Julie Thompson (Strasbourg) and Des Higgins (Dublin) to maintain and develop these programs. We also maintain several public web servers at EMBL, including ELM, the protein linear motif resource; Phospho.ELM, a collection of reported phosphorylation sites; GlobPlot, a tool for exploring protein disorder; and Gene2EST, a BLAST server specialised for mapping ESTs to gene sequences.

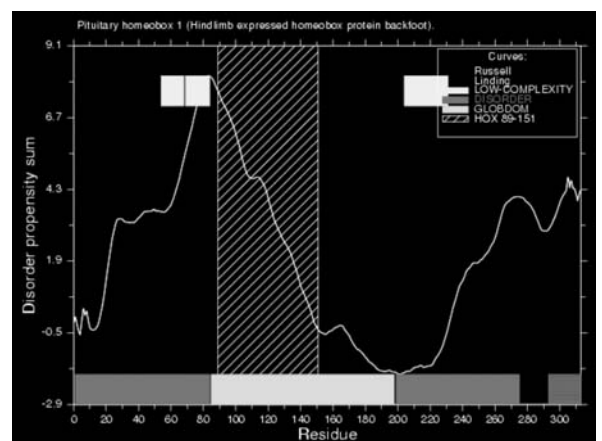
A major focus recently has been to develop and deploy tools for protein architecture analysis. Our group coordinated the EU-funded ELM consortium that has developed the new Eukaryotic Linear Motif resource to help users find functional sites in modular protein sequences. Short functional sites are used for the dynamic regulation of large cellular protein complexes and their characterisation is essential for understanding cell signalling. Such sites are most often in IUP (intrinsically unstructured protein segments) and we have developed tools such as GlobPlot to locate these segments as a prerequisite of ELM motif hunting.

Future projects and goals

We apply computers in molecular biology in the hope of gaining new biological insight that may inform experimental strategies. For example, we have proposed new vaccine targets after a combined phylogenetic analysis and proteome survey revealed that bacteria have acquired $\alpha 2$ -macroglobulin genes found in metazoa. We will continue to survey individual gene families in depth and will undertake proteome surveys when we have specific questions to answer. Molecular evolution is one of the group's interests, especially when it has practical applications.

With our collaborators, we will look to build up the protein architecture tools, especially the unique ELM resource, taking them to a new level of power and applicability. We will apply the tools in the investigation of modular protein function and deploy them in proteome analysis pipelines. Our links to experimental and structural groups should ensure that bioinformatics results feed into experimental analyses of signalling interactions and descriptions of the structures of modular proteins and their complexes, with one focus being regulatory chromatin proteins.

GlobPlot of the Ptx1 transcription factor. Ascending slope indicates disorder preference and descending slope a preference for order. The known Homeobox domain lies within the assigned globular preference while the remaining sequence is unstructured and is expected to contain multiple regulatory motifs interacting with other chromatin proteins.



Selected references

Budd, A., Blandin, S., Levashina, E.A. & Gibson, T.J. (2004). Bacterial $\alpha 2$ -macroglobulins: colonisation factors acquired by horizontal gene transfer from the metazoan genome? *Genome Biol.*, 5, R38

Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J. & Russell, R.B. (2003). Protein disorder prediction. Implications for structural proteomics. *Structure (Camb)*, 11, 1453-1459

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. & Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, 31, 3497-3500

Punertvoll, P., Linding, R., Gemund, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D.M., Ausiello, G., Brannetti, B., Costantini, A., et al., (2003). ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res.*, 31, 3625-3630

Andreas Hoenger

PhD 1993, M.E. Müller Institute, Biocenter, University of Basel, Switzerland.
 Postdoctoral research at the Scripps Research Institute, La Jolla, USA,
 and at the Institute of Cell Biology, ETH-Zürich, Switzerland.
 Group Leader at EMBL since 1998. Joint appointment with the Cell Biology and Biophysics Unit.



Microtubule interactions with molecular motors

Previous and current research

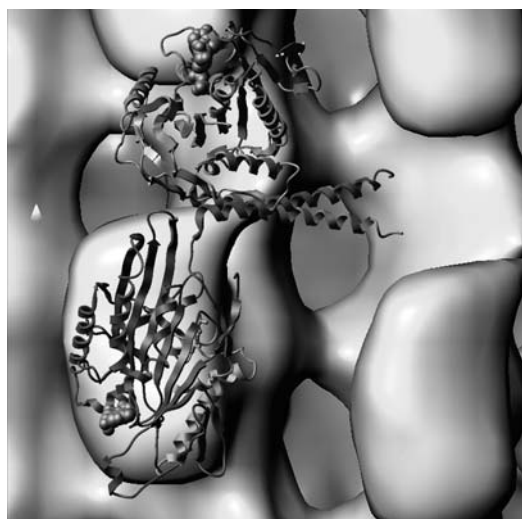
We are studying the interaction of kinesin-like motor proteins with microtubules by cryo-electron microscopy and helical or crystallographic image reconstruction methods. Microtubule-motor protein interactions play a crucial role in many cellular processes, such as chromosomal segregation during cell division, axonal transport and flagellar motion. We visualise motor movement in the electron microscope by inducing distinct conformational changes into motor domains and tubulin binding sites which occur during the energy converting cycles to gain insight into the molecular mechanisms involved in motor movement and directionality. To this end, we analyse the resulting three-dimensional maps of the microtubule-motor protein complexes by merging them with atomic resolution X-ray or NMR data from individual components.

Our main focus is on combining the power of the different methods for structure determination of macro-molecular assemblies used today, such as electron microscopy, X-ray crystallography and NMR spectroscopy to get a detailed understanding of the dynamic properties of the cytoskeleton. This includes the walking mechanisms of molecular motors as well as the dynamics and structural properties of the motor "tracks", such as microtubules.

Future projects and goals

Our immediate projects include a detailed analysis of the functional relevance of conformational changes which seem to occur upon motor protein binding, and their involvement in motor dynamics. We will study the origin of the different directionalities of minus-end and plus-end directed motors, in particular of kinesin and *ncd*, in detail. We will compare the "walking" mechanisms of double-headed kinesin-like motor proteins with those of single-headed ones. We will also push the resolution of cryo-electron microscopy-based 3D maps of tubulin-motor protein complexes by using different protofilament polymers, such as flattened tubulin sheets or zinc-induced sheets.

Our goal is to combine cryo-electron microscopy of supramolecular assemblies with atomic resolution data to study the mechanisms and dynamics of microtubule-motor protein interaction.



Docking of the kinesin dimer X-ray structure into the EM-derived envelope of the microtubule-kinesin complex.

Selected references

- Hoenger, A., Sack, S., Thormahlen, M., Marx, A., Müller, J., Gross, H. & Mandelkow, E. (1998). Image reconstructions of microtubules decorated with monomeric and dimeric kinesins: comparison with X-ray structure and implications for motility. *J. Cell Biol.*, 141, 419-430
- Hoenger, A. & Milligan, R.A. (1997). Motor domains of kinesin and *ncd* interact with microtubule protofilaments with the same binding geometry. *J. Mol. Biol.*, 265, 553-564
- Hoenger, A., Sablin, E.P., Vale, R.D., Fletterick, R.J. & Milligan, R.A. (1995). Three-dimensional structure of a tubulin-motor-protein complex. *Nature*, 376, 271-274



Rob Russell

PhD 1993, Oxford University, UK.
 Postdoctoral research at Imperial Cancer Research Fund, London, UK.
 Senior Investigator at SmithKline Beecham, London, UK.
 Group Leader at EMBL since 2001. Joint appointment with EMBL-EBI.

Structural bioinformatics

Previous and current research

Our group is interested in the relationships between protein three-dimensional (3D) structure, function and evolution. Current projects are focused in three main areas.

Protein interactions and complexes. Protein interaction networks are central to any understanding of cellular processes, and though many thousands are now known, few initiatives to uncover them pay much attention to one of the best sources of data available: complexes of known 3D structure. We thus study protein interactions by considering known 3D structures. We use 3D complexes to interrogate interactions identified by other methods (e.g. yeast two-hybrids) and to predict specific interactions within protein families. A major initiative in the group is related building as complete models as possible for all interacting proteins and complexes in a whole cell. This is particularly useful when combined with experimental methods like electron microscopy or interaction discovery.

Predicting function from structure. Thriving structural genomics projects, together with the growing pace of structural biology, now means that protein 3D structures can be known before function. These structures present fascinating challenges and require many new approaches to be most useful. We are currently developing methods for predicting function for structures by way of protein and small-molecules comparison. We also work on methods to predict regions in sequences most likely to be globular, which are aimed more at the target side of structural genomics (i.e. before structure determination begins).

Protein and DNA sequence motifs. A major current challenge in biology is to discover and understand short protein or nucleic acid stretches that mediate functional interactions. We currently search for new protein-peptide and microRNA target sequences in genomes using a variety of techniques. Both methods already make fascinating predictions of biological phenomena and provide a wealth of information for people working with such sequences experimentally.

Future projects and goals

- To develop methods for modelling 3D structures for complexes by combining binary interactions. In this we have defined a new generation of structure prediction.
- To test predicted protein-peptide interactions in concert with experimental groups.
- To develop methods for predicting small-molecules and peptide interaction sites on protein structures, and apply this to drug design.
- To build databases relating small-molecules to their interacting proteins.
- To foster new and existing collaborations with experimental groups in EMBL and elsewhere.

Selected references

Stark, A., Shkumatov, A. & Russell, R.B. (2004). Finding functional sites in structural genomics proteins. *Structure (Camb)*, 12, 1405-1412

Aloy, P., Böttcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.C., Bork, P., Superti-Furga, G., Serrano, L. & Russell, R.B. (2004). Structure-based assembly of protein complexes in yeast. *Science*, 303, 2026-2029

Stark, A., Brennecke, J., Russell, R.B. & Cohen, S.M. (2003). Identification of *Drosophila* microRNA targets. *PLoS Biol.*, 1, E60

Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J. & Russell, R.B. (2003). Protein disorder prediction. Implications for structural proteomics. *Structure (Camb)*, 11, 1453-1459

Michael Sattler

PhD 1995, University of Frankfurt, Germany.
Postdoctoral research at Abbott Laboratories, Abbott Park, USA.
Group Leader at EMBL since 1997.
Joint appointment with the Gene Expression Unit.



Biomolecular NMR spectroscopy

Previous and current research

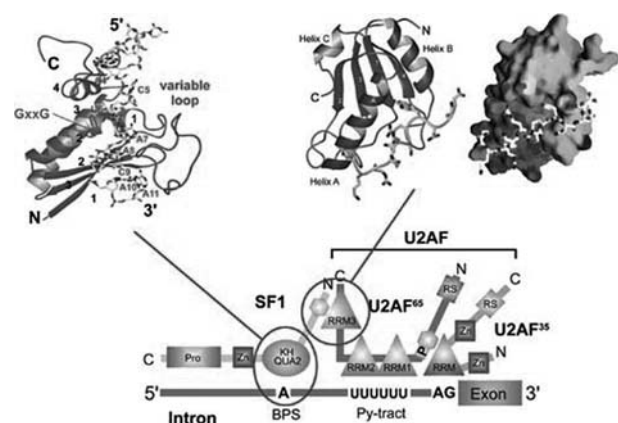
We study the structure and dynamics of biomolecules using nuclear magnetic resonance (NMR). NMR spectroscopy is a powerful technique not only for determining the three-dimensional structures but also for the characterisation of interactions and dynamics of biological molecules in solution. Recent advances in NMR methodology and instrumentation allow studying multi-domain proteins and complexes of 100 kDa molecular weight and higher in solution.

Recognition of proteins and nucleic acids forms the molecular basis for many biological processes. In combination with functional assays and mutational analyses the three-dimensional structure provides insight into biological function. NMR is well suited to study protein domains, which form the structural building blocks of larger proteins and typically consist of 50-150 amino acids. They usually interact with other proteins to mediate signal transduction or bind to nucleic acids to regulate gene expression. We are currently studying protein domains which are implicated in human disease and function in signal transduction or gene expression.

A main focus in the group is to understand the structural basis of protein-RNA interactions that are functionally important for various aspects of gene expression, such as (alternative) splicing or RNA interference. The spliceosome is a highly dynamic machinery, which consists of numerous protein/RNA assemblies. During the different steps that eventually lead to splicing of the pre-mRNA, these complexes are continuously rearranged. While this requires that the molecular interactions are dynamic, specific and tight complexes are still formed by the cooperative combination of multiple weak protein/protein and protein/RNA interactions. Protein/RNA recognition often involves mutually induced fit of the binding partners, which may be reflected in the conformational dynamics of these molecules. NMR is well suited to study such dynamic interactions in solution. Current projects focus on protein/protein and protein/RNA interactions that play important roles in various aspects of RNA metabolism.

Future projects and goals

We characterise the structures and molecular interactions of proteins in solution in order to elucidate the structural basis of their biological function. A main focus of our research lies in protein-protein and protein-RNA interactions, which play important roles in RNA metabolism. We are optimising NMR methods for studying larger protein-protein and protein-RNA complexes in order to characterise their cooperative interactions in solution.



Recognition of the 3' splice site in an pre-mRNA intron by Splicing Factor 1 (SF1) and the U2 auxiliary factor (U2AF). U2AF is a heterodimer of a small (35 kDa) and a large (65 kDa) subunit. Three-dimensional structures that we have determined in the past are shown. SF1 specifically recognizes the conserved branch point adenosine in the pre-mRNA intron. The N-terminus of SF1 interacts with the third RRM domain of U2AF65.

Selected references

- Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. (2004). Nucleic acid 3' end recognition by the Argonaute 2 PAZ domain. *Nat. Struct. Mol. Biol.*, 11, 576-577
- Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature*, 426, 465-469
- Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Kramer, A. & Sattler, M. (2003). Structural basis for the molecular recognition between human splicing factors U2AF(65) and SF1/mBBP. *Mol. Cell*, 11, 965-976
- Liu, Z., Luyten, I., Bottomley, M.J., Messias, A.C., Houngrinou-Molango, S., Sprangers, R., Zanier, K., Kramer, A. & Sattler, M. (2001). Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science*, 294, 1098-1102



Klaus Scheffzek

PhD, the Max Planck Institute for Medical Research, Heidelberg, Germany.
Postdoctoral research at the Max Planck Institute of
Molecular Physiology, Dortmund, Germany and EMBL.
Team Leader at EMBL since 1999.

Molecular medicine – structure function analysis of signalling proteins

Previous and current research

Defects in signalling pathways are often associated with the occurrence of severe diseases, with cancer being a very common example. We are interested in understanding the mechanisms of pathogenesis associated with cancer related diseases. In previous work we have characterised the regulation of Ras and the related Rho proteins. Ras is a GTP binding protein mutated in 30% of human tumours. It functions like a binary molecular switch cycling between GTP-bound ON- and GDP-bound OFF-states. Ras mediated GTP hydrolysis turns the switch off. This intrinsically slow process is enhanced by so called GTPase activating proteins (GAPs). Oncogenic Ras mutants are permanently activated and are not sensitive to GAPs. In earlier studies we have elucidated the chemical mechanism of GTPase activation and explained why oncogenic Ras mutants are not GAP sensitive.

Currently our major focus is on neurofibromatosis type 1 (NF1), a genetic disease with an incidence of 1 in 3,500 worldwide. NF1 patients have an increased risk of developing certain types of tumours and frequently show learning disabilities. The NF1 gene encodes a huge protein of about 2,800 amino acids, termed neurofibromin, and is responsible for the pathogenesis of the disease. Neurofibromin acts as a Ras specific GAP, and in some tumour types lacking the protein, Ras is indeed upregulated. In previous studies we have characterised the GTPase activating domain and investigated potential effects of patient mutations.

The GAP activity of neurofibromin resides in a segment, which represents only 10% of the protein. After ten years of research since the discovery of the NF1 gene, no other biochemical function of neurofibromin has been clearly defined. We are following a structural approach to find out about functions of the remaining 90% of the protein. The idea is to identify neurofibromin segments that can be expressed as soluble proteins, determine the structures of such fragments, and by comparison with known protein structures in databases or by bound ligands obtain ideas for functional/biochemical experiments. Our main technique is X-ray crystallography but other techniques are increasingly employed. Using a structure-based approach we have been recently successful in identifying a novel domain in neurofibromin that we are currently characterising. Our approach has also been successful with other proteins and is becoming increasingly important in the functional analyses of genomic sequences coding for hypothetical proteins of hitherto unknown functions.

Future projects and goals

A major goal is to arrive at a three-dimensional model of neurofibromin. This is a challenging task, given the size and the expected complexity of the molecule. Therefore we adopt a “divide and conquer” strategy, which will enable us to obtain structures of at least some segments. In summary, we want to understand neurofibromin function via structure analysis.

Other projects of the laboratory include signalling by eukaryotic and prokaryotic protein kinases, further aspects in structural neurobiology and regulation of viral transcription.

Selected references

Márquez, J.A., Smith, C.I., Petoukhov, M.V., Lo Surdo, P., Mattsson, P.T., Knekt, M., Westlund, A., Scheffzek, K., Saraste, M. & Svergun, D.I. (2003). Conformation of full-length Bruton tyrosine kinase (Btk) from synchrotron X-ray solution scattering. *EMBO J.*, 22, 4616-4624

Ahmadian, M.R., Kiel, C., Stege, P. & Scheffzek, K. (2003). Structural fingerprints of the Ras-GTPase activating proteins neurofibromin and p120GAP. *J. Mol. Biol.*, 329, 699-710

Márquez, J.A., Hasenbein, S., Koch, B., Fieulaine, S., Nessler, S., Russell, R.B., Hengstenberg, W. & Scheffzek, K. (2002). Structure of the full-length HPr kinase/phosphatase from *Staphylococcus xylosum* at 1.95 Å resolution: mimicking the product/substrate of the phospho transfer reactions. *Proc. Natl. Acad. Sci. USA*, 99, 3458-3463

Reinhard Schneider

PhD 1994, University of Heidelberg, Germany. Postdoctoral research at EMBL.
Co-founder and Chief Information Officer at LION bioscience AG, Germany.
Chief Executive Officer at LION bioscience Research Inc., Cambridge, MA, USA.
Team Leader at EMBL since 2004.



Data integration and knowledge management

Previous and current research

Today it is widely recognised that a comprehensive integration of data can be one of the key factors to improve productivity and efficiency in the biopharmaceutical research and development process. Successful data integration helps researchers to discover relationships that enable them to make better and faster decisions, thus considerably saving time and money. In a logical extension to these arguments one can apply the ideas and technologies used in industry in a basic research environment. An additional challenge in an academic environment is the even less structured “process chain” with completely new data types and fast changing requirements from the end-users.

Over the last 20 years, biological research has seen a very strong proliferation of data sources. Each research group and each new experimental technique generates a source of valuable data. The creation, use, integration and warehousing of biological data is central to large-scale efforts in understanding biological systems. These tasks pose significant challenges from the standpoint of data storage, indexing, retrieval and system scalability over disparate types of data.

The principal aim of the group is to capture and centralise the knowledge generated by the scientists in the several divisions, and to organise that knowledge in such a way that it is easily mined, browsed and navigated. By providing access to all scientists in the organisation, it will foster collaboration between researchers in different cross-functional groups. The intention is to first start with pilot projects in the Heidelberg facility and potentially extends the efforts to the experimental outstations.

The group is involved in the following areas:

- data schema design and technical implementation;
- metadata annotation with respect to experimental data;
- design and implementation of a scientific data portal;
- providing access to and further development of data-mining tools (e.g. text-mining).

Future projects and goals

Our goal is to develop a comprehensive knowledge platform for the life sciences. We will first focus on the biology-driven research areas, but will extend into chemistry-related fields preliminarily by collaborating with groups within EMBL. Other research areas will include advanced data-mining and visualisation techniques, data integration and knowledge management.

Selected references

Kremer, A., Schneider, R. & Terstappen, G.C. (2005). Proteomics data: storage, analysis and integration – a short overview. Bioscience Reports, in press

Ofran, Y., Punta, M., Schneider, R. & Rost, B. (2005). Beyond annotation transfer by homology: novel protein function prediction methods that can assist drug discovery. Drug Discovery Today, in press

Ramakrishnan, S., Caruso, A. & Schneider, R. (2002). Improving research productivity at a pharmaceutical company. LION bioscience White paper



Dietrich Suck

PhD 1971, the Max Planck Institute for Experimental Medicine, Göttingen, Germany.
Postdoctoral research at the MPI, Göttingen, Germany,
and Purdue University, Indiana, USA.
Group Leader at EMBL since 1982.

Structural basis of protein-nucleic acid interactions

Previous and current research

We study protein-nucleic acid interactions by a combination of crystallographic and biochemical techniques to understand the structural basis of their varying specificities, the formation of multi-component protein-nucleic acid complexes and the catalytic mechanisms involved in nucleic acid modification.

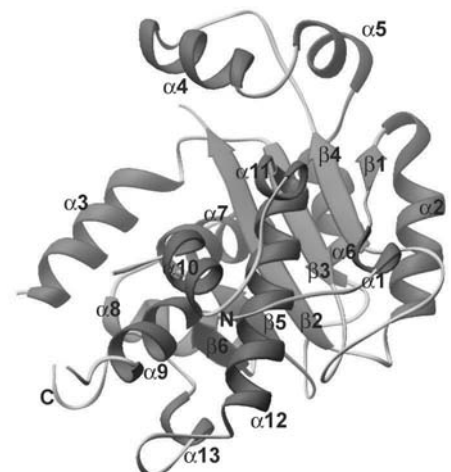
A continuing theme of our research has been the substrate recognition by structure-selective nucleases. The level of their structural specificity ranges from simple discrimination between single- and double-stranded substrates (nucleases P1 and S1), the recognition of groove geometry and flexibility (DNase I), to the recognition of specialised structures, such as flap DNA (T5 5'-nuclease) or Holliday junctions (HJ); (T4 endonuclease VII, Cre recombinase). Presently, we are focusing on HJ resolvases including phage T4 endonuclease VII, whose intrinsic conformational flexibility is thought to be of functional significance for its broad substrate specificity. Other resolvases we are studying (archaeal Hjc's, yeast Cce1), in contrast to Endo VII, do specifically recognise and cleave 4-way DNA junctions.

Other projects focus on proteins interacting with or modifying RNA, such as yeast Arc1p, which forms multi-component complexes with met- and glu-tRNA and their cognate tRNA-synthetases. The X-ray structures of RNA complexes of Sm-related proteins from Archaea (*A. fulgidus* and *P. abyssi*) provided the first high-resolution picture of RNA-binding in an Sm core domain and revealed the close evolutionary relationship to eukaryotic snRNPs and bacterial Hfq proteins. Recently we solved the structure of the *E. coli* Hfq protein, also known as host factor I, which promotes interactions of small regulatory RNAs with their mRNA targets.

Another project in this category deals with yeast complexes involved in mRNA degradation. Pop2 – a homologue of the human Caf1 protein – is a component of the Ccr4-Not complex, which is responsible for mRNA deadenylation in yeast. We have solved the crystal structure of its non-canonical nuclease domain, representing the first structure of an RNase from the DEED superfamily and revealing its structural homology with the 3'-5' exonuclease domains of *E. coli* DNA polymerases I and III.

Future projects and goals

By analysing the 3D structures of protein-nucleic acid complexes, we want to reveal the structural basis of specificity and catalysis. Future projects include structure determinations of substrate complexes of Endo VII and other resolvases, Arc1p/tRNA/tRNA synthetase complexes, structure-function studies of Sm-related proteins and complexes involved in mRNA degradation.



Structure of the nuclease domain of the yeast Pop2 protein.

Selected references

Thore, S., Mauxion, F., Séraphin, B. & Suck, D. (2003). X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. *EMBO Rep.*, 4, 1150-1155

Ennifar, E., Meyer, J.E., Buchholz, F., Stewart, A.F. & Suck, D. (2003). Crystal structure of a wild-type Cre recombinase-loxP synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation. *Nucleic Acids Res.*, 31, 5449-5460

Sauter, C., Basquin, J. & Suck, D. (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Res.*, 31, 4091-4098

Thore, S., Mayer, C., Sauter, C., Weeks, S. & Suck, D. (2003). Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya. *J. Biol. Chem.*, 278, 1239-1247

Directors' Research

Directors' Research is unlike other EMBL Units in that it covers two independent research groups without an overall Coordinator. These groups are headed by the Director General of EMBL and the Executive Director of EMBO. As both appointments are made on the basis of a variety of skills (including research activities), it was thought rather important to define them as a separate Unit rather than placing them under the control of a different Unit Coordinator.

The Mattaj group studies diverse processes that are under the control of the Ran GTPase. During mitosis, Ran is needed for both mitotic spindle assembly and nuclear envelope (NE) formation. Remarkably, their studies suggest that Ran's mitotic functions occur by the same mechanism as nucleo-cytoplasmic transport, i.e. via regulation of interactions between transport receptors and factors involved in spindle or NE assembly. Using *in vitro* and *in vivo* methods, they are identifying the factors that mediate these processes and finding out how they function and how Ran controls their activity. NE assembly is a multi-stage process. Both the membranes that give rise to the NE and the proteins that form nuclear pore complexes (NPCs), through which transport across the NE occurs, associate with the chromatin surface early in NE assembly. Membrane fusion events and NPC assembly proceed in concert, and have to be regulated in an integrated way. The group has begun to understand how Ran controls NPC assembly, but has little information on how the NE membranes assemble, or how NPC insertion into the membranes takes place. In addition, although it is known that Ran regulates where NE assembly occurs in the cell, they do not know how the process is temporally regulated, i.e. why it occurs in telophase rather than at other times during mitosis. Understanding the spatial regulation of mitotic events by Ran and their temporal regulation during the cell cycle is an ambitious long-term goal.

The Gannon group focuses on the regulation of expression of genes regulated by the steroid hormone oestrogen. The physiological effects of oestrogen are transduced through specific nuclear proteins, the Oestrogen Receptors, which are oestrogen-dependent transcription factors. They have defined the promoter regions of ER- α from human, chicken and mouse, and described multiple start points of transcription. In addition, they have demonstrated that the 3'-untranslated region of the human oestrogen receptor has a role in de-stabilising the ER- α messenger RNA. More recently, they have comprehensively portrayed the sequence of events, instigated by Oestrogen Receptor, that achieve and then limit transcription of oestrogen responsive promoters. Ultimately, the group's aim is to understand how oestrogen regulates a number of diverse physiological processes, such as embryonic development, sex determination and reproduction. This information is essential to understanding the role of steroid hormone receptors in the onset and progression of a variety of pathological conditions such as cancer, osteoporosis and Alzheimer's disease.



Frank Gannon

PhD 1973, University of Leicester, UK. Postdoctoral research at the University of Wisconsin, Madison, USA and the University of Strasbourg, France.
Associate Professor and Director of the National Diagnostic Centre, University College Galway, Ireland.
Executive Director EMBO; Secretary General EMBC. At EMBL since 1994.

Selected references

Reid, G., Metivier, R., Lin, C.Y., Denger, S., Ibberson, D., Ivacevic, T., Brand, H., Benes, V., Liu, E.T. & Gannon, F. (2005).

Multiple mechanisms induce transcriptional silencing of a subset of genes, including oestrogen receptor- α , in response to deacetylase inhibition by valproic acid and trichostatin A. *Oncogene*, 24, 4894-4907

Metivier, R., Penot, G., Carmouche, R.P., Hubner, M.R., Reid, G., Denger, S., Manu, D., Brand, H., Kos, M., Benes, V. & Gannon, F. (2004). Transcriptional complexes engaged by apo-oestrogen receptor- α isoforms have divergent outcomes. *EMBO J.*, 23, 3653-3666

Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M. & Gannon, F. (2003). Oestrogen receptor- α directs ordered, cyclical and combinatorial recruitment of cofactors on a natural target promoter. *Cell*, 115, 751-763

Reid, G., Hubner, M.R., Metivier, R., Brand, H., Denger, S., Manu, D., Beaudouin, J., Ellenberg, J. & Gannon, F. (2003). Cyclic, proteasome-mediated turnover of unliganded and liganded ER- α on responsive promoters is an integral feature of oestrogen signalling. *Mol. Cell*, 11, 695-707

The expression and functional regulation of Oestrogen Receptor- α

Previous and current research

A major focus of my laboratory's work is the control of the expression of genes by the Oestrogen Receptor- α (ER- α). Although there have been very many studies on this topic, there is still an amazing amount of information that is uncovered every year. In the recent past we have turned our attentions specifically on the detailed description of the mechanism of action of the ER- α . The process that we have defined, in a very precise manner, is the cycling of the ER on the promoter of a target gene. The ER binds to the promoter in the presence or absence of oestradiol, and thereafter follows a sequence of recruitment of co-factors which ultimately results in the binding of polymerase II (in the presence of ligand) or in the non-productive cycle for the receptor when no ligand is present. Almost 50 factors related to transcription were monitored in this ChIP based experiment and the data showed not only that the receptor is cycling, but also that the process had many redundant aspects and that the steps that are normally viewed as being negative (e.g. degradation of the receptor by the proteasome) are in fact an inherent component of the successful action of the oestrogen receptor.

Following on the messages from these studies, we are now developing new tools to interfere with this process. The first of these was the use of histone deacetylase inhibitors, but others are following from chemogenomic approaches. We are also studying the process using proteomics and extending the studies to other target genes for the oestrogen receptor including those that are down regulated.

Future projects and goals

At all times our studies are linked to human disease and our current focus remains on human breast cancer and osteoporosis. The laboratory, therefore, links very fundamental studies with the potential for an impact on human medicine and in this way fits well into the molecular medicine aspirations of EMBL. In doing so, we continue to collaborate very actively with groups from Cell Biology and Biophysics, Gene Expression, Structural and Computational Biology Units and with many of the core facilities, particularly with the Genomics and Chemical Genomics facilities.

Iain Mattaj

PhD 1979, University of Leeds, UK. Postdoctoral research at the Friedrich Miescher Institute and the Biocenter, Basel, Switzerland. Group Leader since 1985. Gene Expression Unit Coordinator from 1990-1999. Scientific Director of EMBL from 1999-2005. Director-General of EMBL since 2005.



The RanGTPase as a spatial regulator

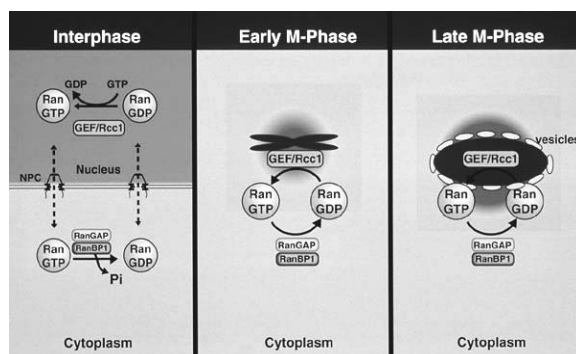
Previous and current research

The research in our group is centred on diverse processes that are under the control of the Ran GTPase. Ran requires regulators for its activity. These are a GEF (Guanine Nucleotide Exchange Factor) that loads Ran with GTP, and a GAP (GTPase Activating Protein) that is required for Ran to hydrolyse GTP to GDP. These two factors are asymmetrically distributed within cells, both in interphase and during mitosis (see figure). As a result, RanGTP is present in interphase at high concentration in the nucleus and at low concentration in the cytoplasm. In mitosis, RanGTP concentration is locally increased in the vicinity of chromatin because of the concentration of the GEF there. RanGTP interacts with the import and export receptors that mediate nucleo-cytoplasmic transport. In the former case, this interaction results in import cargo release. Thus, import receptors bind cargo in the cytoplasm (low Ran GTP) and release it in the nucleus (high RanGTP). Ran's effect on export receptors is the opposite, they only interact with their cargoes in the presence of RanGTP. These RanGTP-dependent binding events impart directionality to nuclear transport. During mitosis, Ran is needed for both mitotic spindle assembly and nuclear envelope (NE) formation (see figure). Remarkably, our studies suggest that Ran's mitotic functions occur by the same mechanism as nucleo-cytoplasmic transport, i.e. via regulation of interactions between transport receptors and factors involved in spindle or NE assembly.

Future projects and goals

In the case of spindle formation, we know that Ran regulates multiple aspects of spindle assembly; microtubule nucleation, microtubule stability, the production of anti-parallel microtubule arrays, the focusing of the spindle poles, etc. We are using in vitro and in vivo methods to identify the factors that mediate these processes and to find out how they function and how Ran controls their activity.

NE assembly is a multi-stage process. Both the membranes that give rise to the NE and the proteins that form nuclear pore complexes (NPCs), through which transport across the NE occurs, associate with the chromatin surface early in NE assembly. Membrane fusion events and NPC assembly proceed in concert, and have to be regulated in an integrated way. We have begun to understand how Ran controls NPC assembly but we have essentially no information on how the NE membranes assemble, or how NPC insertion into the membranes takes place. In addition, although we know that Ran regulates where NE assembly occurs in the cell, we do not know how the process is temporally regulated, i.e. why it occurs in telophase rather than in metaphase. Understanding the spatial regulation of mitotic events by Ran and their temporal regulation during the cell cycle is an ambitious long-term goal.



The figure depicts the local production of RanGTP, caused by the non-uniform distribution of the Ran GEF and Ran GAP in relation to the nuclear envelope or the chromosomes at various stages in the cell cycle.

Selected references

- Antonin, W., Franz, C., Haselmann, U., Antony, C. & Mattaj, I.W. (2005). The integral membrane nucleoporin pom121 functionally links nuclear pore complex assembly and nuclear envelope formation. *Mol. Cell*, 17, 83-92
- Mattaj, I.W. (2004). Sorting out the nuclear envelope from the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.*, 5, 65-69
- Walther, T.C., Askjaer, P., Gentzel, M., Habermann, A., Griffiths, G., Wilm, M., Mattaj, I.W. & Hetzer, M. (2003). RanGTP mediates nuclear pore complex assembly. *Nature*, 424, 689-694
- Walther, T.C., Alves, A., Pickersgill, H., Loiodice, I., Hetzer, M., Galy, V., Hulsman, B.B., Kocher, T., Wilm, M., Allen, T., Mattaj, I.W. & Doye, V. (2003). The conserved Nup107-160 complex is critical for nuclear pore complex assembly. *Cell*, 113, 195-206

Core Facilities

Biology is moving from an era in which research has focused on a small number of molecules or cellular events into the age of functional genomics – where a genome-wide approach is needed to study the fundamental processes of life. Researchers hope to identify all the molecules in a cell and explain their functions, striving towards a holistic view of how cells behave, how organisms develop and how diseases disrupt biological processes. This shift means that researchers need to access many types of efficient services, some of them operating at high throughput. During the past years, EMBL has established several core facilities to offer such services to its researchers, visitors and scientists from the laboratory's Member States. Each facility has tight links to industries which provide state-of-the-art equipment used to address the most interesting scientific questions. The results from these experiments feed directly back into improvements in instruments and methodologies, spawning the next generation of scientific questions and technology. Additionally, research at the EMBL core facilities goes hand in hand with the training of students, in-house staff and a constant stream of visitors, reaching a huge number of users and potential future customers. The facilities aim to cover the needs of scientists engaged in a wide spectrum of molecular biology research.

Claude Antony

PhD 1984, University Paris VI, France.
Postdoctoral research at EMBL.
Group Leader at CNRS.
Facility Head and Team Leader at EMBL since 2003.



Electron Microscopy Core Facility

The Electron Microscopy (EM) Core Facility gives EMBL scientists access to advanced electron microscopes, relevant sample preparation techniques and sophisticated instrumentation for their research needs. One main task of the EM Core Facility is to train new users to make best use of our advanced equipment and various techniques. We also work on the development of new approaches and methods in Electron Microscopy application to Cellular and Developmental Biology (e.g. tomography of intact cells).

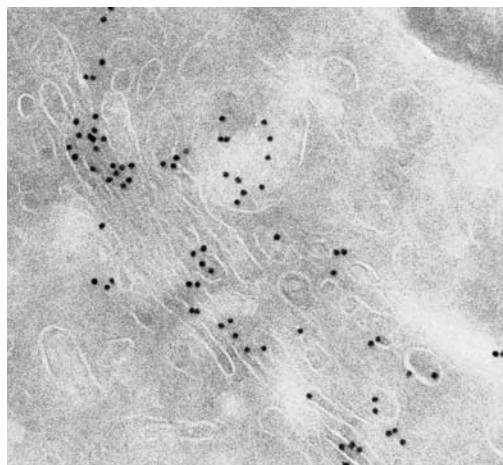
Services provided

- Providing up-to-date knowledge on EM methods for cell biology and immunocytochemistry, in particular the use of cryosectioning and cryofixation of various cell types or organisms.
- Maintaining the electron microscopy equipment and the laboratory for sample preparation, microtomy and various cryogenic methods.
- Supplying a range of reagents specific for the relevant EM methods and protocols.
- Assisting users in choosing the right methods and protocols for their research.
- Organising courses and lectures on EM methods in cell biology.

Technology partners

Access to the facility requires users to be trained in sample preparation techniques as well as in the use of microscopes.

- The FEI Company: supplier of advanced electron microscopes.
- Leica: supplier of the High Pressure Freezing Machine (EM PACT). A partnership has been established between Leica and the EM Facility to develop and promote this machine.



Detection of a tracer protein crossing the Golgi apparatus – colloidal gold labelled cryosection.

Selected references

- Müller-Reichert, T., Sassoon, I., O'Toole, E., Romao, M., Ashford, A.J., Hyman, A.A. & Antony, C. (2003). Analysis of the distribution of the kinetochore protein Ndc10p in *Saccharomyces cerevisiae* using 3-D modelling of mitotic spindles. *Chromosoma*, 111, 417-428
- Arrastua, L., San Sebastian, E., Quincoces, A.F., Antony, C. & Ugalde, U. (2003). In vitro fusion between *Saccharomyces cerevisiae* secretory vesicles and cytoplasmic-side-out plasma membrane vesicles. *Biochem. J.*, 370, 641-649
- Mallard, F., Tang, B.L., Galli, T., Tenza, D., Saint-Pol, A., Yue, X., Antony, C., Hong, W., Goud, B. & Johannes, L. (2002). Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform. *J. Cell Biol.*, 156, 653-664
- Popov, A.V., Pozniakovsky, A., Arnal, I., Antony, C., Ashford, A.J., Kinoshita, K., Tournibize, R., Hyman, A.A. & Karsenti, E. (2001). XMAP215 regulates microtubule dynamics through two distinct domains. *EMBO J.*, 20, 397-410
- Le Bot, N., Antony, C., White, J., Karsenti, E. & Vernos, I. (1998). Role of xklp3, a subunit of the *Xenopus* kinesin II heterotrimeric complex, in membrane transport between the endoplasmic reticulum and the Golgi apparatus. *J. Cell Biol.*, 143, 1559-1573



Joe Lewis

PhD 1991, Institute of Molecular Pathology, Vienna, Austria.

Postdoctoral research at EMBL.

Group Leader at the Wellcome Trust Centre for Cell Biology, Edinburgh, Scotland.

Group and Global HCV Project Leader at Anadys Pharmaceuticals, Heidelberg, Germany.

Facility Head at EMBL since 2004.

Chemical Biology Core Facility

Small molecules play essential roles in many areas of basic research and are often used to address important biological questions. The aim of our Chemical Biology Core Facility is to enable research groups to address biological questions by identifying and developing “biotool” compounds against novel targets. We can assist groups in the development of primary and secondary assays for screening against our in-house compound library and guide them through the process of developing tool compounds for their specific target. Chemical optimisation projects can be done in collaboration with our chemistry partners.

The facility is a collaboration between EMBL and the DKFZ (German Cancer Research Center, Heidelberg) to provide the infrastructure and expertise to open up small molecule development to research groups at both institutions.

Major projects and accomplishments

The facility was established at the beginning of 2004. Through a large team effort, it was fully functional in August 2004. We have a very strong pipeline of projects from EMBL and DKFZ, several of which have completed screening and are now in the lead optimisation phase.

Services provided

- Selection of appropriate assay technology platforms.
- Developing assays for medium-throughput screening.
- Small molecule screening library of greater than 50,000 compounds.
- Assisting in the design of secondary specificity assays.
- Compound characterisation.
- Managing compound acquisition through our chemistry partners.

Partners

- Technology partners: Perkin Elmer, IDBS.
- Chemistry partners: Tripos Inc. and Tripos Discovery Research Ltd.



Parallel pipetting of samples
in 384-well format.

Andrew Riddell

BSc, 1992, Paisley University, Scotland;
PgDip, 1993, Caledonian University, Glasgow, Scotland.
Work at the MRC LMB and CIMR and Hutchison/MRC, Cambridge, UK.
Facility Head at EMBL since 2003.



Flow Cytometry Core Facility

The Flow Cytometry Facility offers a range of flow cytometric techniques. The equipment adds flexibility in the preparation and execution of experiments, allowing different approaches to research problems. Our facility meets researchers' needs and enables better resolution in terms of analysis and product.

The goal of the facility is to proactively introduce flow cytometric methods into new research areas while supporting and extending current research.

Major projects and accomplishments

Our analysis of algae life cycle project relies on the intrinsic fluorophores in algae to identify life cycle stages. It requires photosaturation of the photosynthetic units in the algae. Single cell sorting for cloning requires a series of single cell sorts of a target population into 96-well plate in order to select a stable integrated gene into a cell line by Darwinian selection.

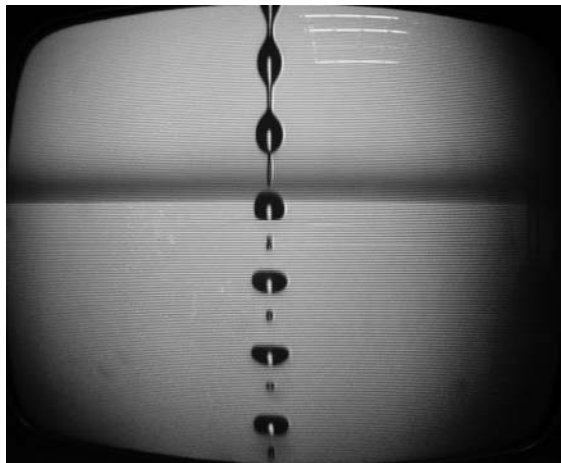
Our apoptosis project utilizes a novel FRET assay in order to identify apoptosis in a cell line. In our phosphorylation cascade project, we aim to have simultaneous detection of "ON/OFF" kinase states by flow cytometry.

Services provided

- Sorting heterogeneous single cell populations into homogeneous populations for experiments.
- Providing an analysis of single cell populations based on fluorescent probes and light intensities (including light scattering and polarisation).
- Providing expertise in flow cytometric techniques for use in experiments.
- Providing advice in the use of flow cytometry.
- Developing novel flow cytometric techniques for use in EMBL's scientific activities.

Technology partners

We work with equipment from DakoCytomation, Becton Dickinson, Union Biometrica and Miltenyi Biotec. We are open to testing new technological developments to better serve the needs of the scientific community.



A view of a point where a jet of fluid breaks into stable droplets. A strobe image of an actively perturbed fluid jet; 50,000 drops per second.



Vladimir Benes

PhD 1994, Czech Academy of Sciences,
Prague, Czech Republic.
Postdoctoral research at EMBL.
Facility Head since 2001.

Genomics Core Facility (GeneCore)

GeneCore is equipped with state-of-the-art hardware required for transcriptome studies and operated by highly qualified staff with all the necessary expertise. As training is an important part of GeneCore activities, its staff is involved in tutoring individual researchers as well as in organising practical courses on subjects pertinent to its orientation, like qPCR or gene expression data analysis. It consists of two teams, DNA sequencing and microarray.

The DNA sequencing team is able to determine primary nucleotide structure of all types of sequenceable templates like PCR fragments, plasmids, bacterial artificial chromosomes and related constructs. Due to the application of the robust Rolling Circle Amplification (RCA) protocol to generate sequencing template, highly-optimised sequencing protocol and optimally tuned capillary sequencer MegaBace by GE Healthcare (GE HC), we have achieved a 95% success rate in the single-pass sequencing, processing more than 20,000 reactions on samples from almost all the “wet lab” groups in Heidelberg and Grenoble, and some from Hamburg, from January to June 2005. In addition, another 10,000 or more sequencing reactions were prepared for various projects, for example full-length insert sequencing (from 3 kb to 48 kb) including primer walking and assembly of the target sequence, as well as sequencing of EST and SAGE (Serial Analysis of Gene Expression) libraries. We are also able to prepare the SAGE libraries for detailed characterisation of transcriptome. GeneCore is also providing an access to the ABI instruments required for sequence detection and quantification of gene expression by quantitative real-time RT-PCR (qPCR), including assay design and training for the first-time users. The instruments are used not only for corroboration of the microarray results but also to verify the ChIP (chromatin immunoprecipitation) experiments.

The DNA microarray team is covering two interrelated activities, microarrays and liquid handling robotics. The team provides capacity and competence to carry out complete microarray experiment whether the investigator wishes to use spotted custom-made microarrays or commercial system such as Affymetrix GeneChip or GE HC CodeLink bioarrays. Instrumentation available for microarray experiments include a high-precision spotter to prepare arrays, automated slide processor to hybridise them and a laser scanner for image acquisition. Evaluation of the results is an inseparable component of the whole process, and for that purpose we are equipped with several licenses of data mining software package GeneSpring (Agilent). We have generated and spotted PCR fragments for *Anopheles* (20,000 features), *Drosophila* (13,000 features) as well as for highly customised human and mouse “iron chip” (2,000 features) microarrays – in total over 2,500 slides. During the first half of 2005, the microarray facility hybridised over 300 samples onto GeneChips and CodeLink microarrays. The powerful suite of liquid-handling robots enables to set up a wide range of medium throughput applications like preparation and purification of PCR fragments, purification of plasmid DNA, replication of the clone sets, preparation of spotting plates, etc. It was also used for setting up a genome-scale RNAi screen at a nematode *C. elegans*.

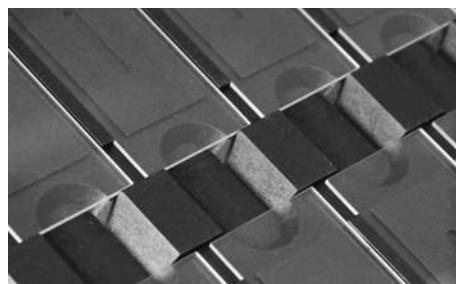
Selected references

Reid, G., Metivier, R., Lin, C.Y., Denger, S., Ibberson, D., Ivacevic, T., Brand, H., Benes, V., Liu, E.T. & Gannon, F. (2005).

Multiple mechanisms induce transcriptional silencing of a subset of genes, including oestrogen receptor- α , in response to deacetylase inhibition by valproic acid and trichostatin A. *Oncogene*, 24, 4894-4907

Bustin, S.A., Benes, V., Nolan, T. & Pfaffl, M.W. (2005). Quantitative real-time RT-PCR – a perspective. *J. Mol. Endocrinol.*, 34, 597-601

Religio, A., Ben-Dov, C., Baum, M., Ruggiu, M., Gemund, C., Benes, V., Darnell, R.B. & Valcarcel, J. (2005). Alternative splicing microarrays reveal functional expression of neuron-specific regulators in Hodgkin lymphoma cells. *J. Biol. Chem.*, 280, 4779-4784



A regular pattern is an important feature of any good microarray experiment.



A bundle of capillaries used to analyse most sequenced samples at GeneCore.

Alan Sawyer

BSc 1988, Paddington College, London, UK.
Work at the Dept. of Molecular Oncology,
Harvard Medical School, USA.
Facility Head at EMBL since 2001.



Monoclonal Antibody Core Facility

The Monoclonal Antibody Core Facility (MACF) was established to provide a service for EMBL researchers and EMBO Young Investigators, while continuously advancing the technology. The mission of our facility is to quickly produce high-affinity antibodies in a high-throughput manner while concentrating on quality of product and of service. The MACF manages over 200 projects per year with an annual theoretical capacity of up to 400 projects. Having developed novel high-throughput techniques and screening assays for the production of mouse-derived, high-affinity monoclonal antibodies, we have recently extended our services to external clients. Current commercial partners include large pharmaceutical companies as well as biotechnology companies.

Services provided

- Consulting on peptide and protein design.
- Producing monoclonal antibodies.
- Characterising antibodies by three separate assays.
- Isotyping of the produced antibodies.
- Scaling up antibody production (10-100 mg level).
- Advising on further characterisation of the produced antibodies.

Technology partners

The Monoclonal Antibody Core Facility works closely with Tecan Italia SA and Genomic Solutions, who provide advanced equipment and top-quality supplies.

The Core facility has an on-going collaboration with BioRad (USA and France) for the development of high-throughput production of complementary pairs of monoclonals for use in their BioPlex assays.

Abnova GmbH and Abnova Corporation (Taiwan) is a licensee of the Core Facility's technologies.



Facility Head Alan Sawyer and his team work on over 200 projects a year in their Monterotondo facility.

Selected references

- Carissimi, C., Baccon, J., Straccia, M., Chiarella, P., Maiolica, A., Sawyer, A., Rappsilber, J. & Pellizzoni, L. (2005). Unrip is a component of SMN complexes active in snRNP assembly. *FEBS Lett.*, 579, 2348-2354
- De Masi, F., Chiarella P., Wilhelm, H., Bullard, B., Ansorge, W. & Sawyer, A. (2005). High-throughput production of mouse monoclonal antibodies using antigen microarrays. *Proteomics*, in press
- Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R. & Georgopoulos, K. (1999). Ikaros DNA-binding proteins direct formation of chromatin remodelling complexes in lymphocytes. *Immunity*, 10, 345-355
- Andersen, S.S., Buendia, B., Dominguez, J.E., Sawyer, A. & Karsenti, E. (1994). Effect on microtubule dynamics of XMAP230, a microtubule-associated protein present in *Xenopus laevis* eggs and dividing cells [published erratum appears in *J. Cell Biol.* 1995, 128, following 988] *J. Cell Biol.*, 127, 1289-1299



Ario de Marco

PhD 1993, Biochemistry, University of Udine, Italy.
Postdoctoral research at Michigan Technical University, Houghton, USA,
University of Crete, Greece, IBMP, Strasbourg, France, Novartis, Basel, Switzerland and EMBL.
Facility Head at EMBL since 2000.

Protein Expression and Purification Core Facility

Our facility produces and purifies proteins from *E. coli*, insect cells and sera using a variety of chromatographic methods. Following each purification, we perform biophysical analyses to ensure the quality of the purified proteins in terms of correct folding and stability. Our group also tests new products, novel techniques, proposes advanced protocols for the purification of proteins and collaborates in the development of automated-compatible materials and time-saving solutions. Each year, the expression and purification steps of 100-200 proteins are optimised in our facility.

Major projects and accomplishments

- Successful development of the “thermovector”, a vector which allows the expressed proteins to be fused to a thermostable partner. This system allows for one-step purification – by simple heating – either manually or automatically.
- Significant increase of protein solubility by co-expressing molecular chaperones and adding chemical chaperones. The established protocols are now routinely used to optimise protein expression.
- Separation, characterisation and re-folding of recombinant protein aggregates.

Services provided

- Establishing and maintaining collections of expression vectors and bacterial strains.
- Developing new vectors and protocols.
- Providing bioanalytical characterisation of proteins.
- Maintaining a web-based flow chart of vectors and protocols.
- Giving scientific and technical assistance.
- Caring for equipment – protein production (2.5 and 42-litre fermentors, insect cell room), purification (FPLC) and analysis (ITC, AUC, DLS, CD, BIACORE).

Selected references

Schrodel, A. & de Marco, A. (2005). Characterisation of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem.*, 6, 10

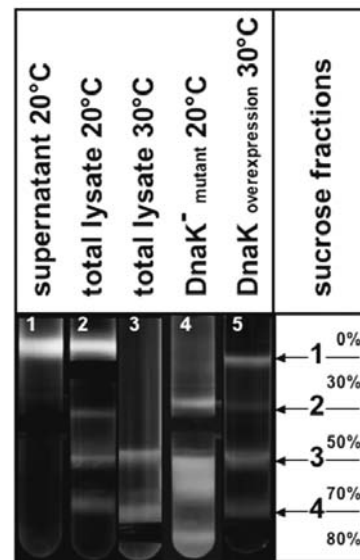
Stegemann, J., Ventzki, R., Schrodel, A. & de Marco, A. (2005). Comparative analysis of protein aggregates by blue native electrophoresis and subsequent sodium dodecyl sulfate-polyacrylamide gel electrophoresis in a three-dimensional geometry gel. *Proteomics*, 5, 2002-2009

De Marco, A. & De Marco, V. (2004). Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *J. Biotechnol.*, 109, 45-52

Technology partners

Our established collaborations with several scientific groups have always represented important opportunities for both partners. Our partners have their materials and ideas tested on a large scale while we gain access to materials and information. These collaborations have already produced patented items. The facility is also open to test new company products and instruments. In the future, we wish to recruit partners for the development of the recombinant antibody technology.

Separation of recombinant GFP-GST fractions with different degree of aggregation by a sucrose step gradient.



Thomas Franz

PhD 1992, Würzburg University, Germany.
Postdoctoral research at Chulabhorn Research Institute, Bangkok, Thailand and at the
Laboratory of Organic Chemistry, School of Agriculture, Nagoya University, Japan.
Facility Head at EMBL since 2001.



Proteomic Core Facility

To address the needs of the post genomic era, the Proteomic Core Facility was set up as a cooperation between EMBL and industry. We provide a complete functional proteomic line, including medium-throughput protein and peptide identification, for both in-house service and external visitors.

Major projects and accomplishments

- Completed differential proteomics of wild type and mutant *E. coli* (in cooperation with EMBL-Hamburg).
- Identified marker proteins for the rare cerebrovascular disorder Moya Moya, which affects the carotid circulation of the brain (in cooperation with the University Hospital Mannheim).
- Developed new MALDI target coating for direct on-target sample preparation.

Services Provided

- Preparative HPLC protein purification.
- IEF with different gradient ranges.
- 2D-gel electrophoresis including colloidal Coomassie, silver or fluorescence staining.
- Laser fluorescence imaging and quantification.
- High-resolution and high-sensitivity imaging densitometer for silver and Coomassie staining.
- PDQuest analysis and evaluation.
- Automatic gel spot excision (fluorescence, silver and Coomassie).
- Full automatic in gel digestion and MALDI target spotting.
- MALDI peptide mass fingerprinting and online database protein identification.
- MS and MS/MS nano electrospray of proteins and peptides.

Technology partners

- Waters/Micromass – sponsors mass spectrometers Q-ToF2, MALDI and MassPrep station.
- BIO-RAD – places advanced equipment at the facility's disposal, including the complete Proteome Works System, 2D-gel equipment, spot cutting robot, HPLC, Photometer, Fluorescence Imager and Densitometer.
- IBM – supplies the computing infrastructure for database searches and instrument operation.

Selected references

- Li, X., Wilm, M. & Franz, T. (2005). Silicone/graphite coating for on-target desalting and improved peptide mapping performance of matrix-assisted laser desorption/ionisation-mass spectrometry targets in proteomic experiments. *Proteomics*, 5, 1460-1471
- Schilling, O., Ruggeberg, S., Vogel, A., Rittner, N., Weichert, S., Schmidt, S., Doig, S., Franz, T., Benes, V., Andrews, S.C., Baum, M. & Meyer-Klaucke, W. (2004). Characterisation of an *Escherichia coli* *elaC* deletion mutant. *Biochem. Biophys. Res. Commun.*, 320, 1365-1373
- Sola, M., Bavro, V.N., Timmins, J., Franz, T., Ricard-Blum, S., Schoehn, G., Ruigrok, R.W., Paarmann, I., Saiyed, T., O'Sullivan, G.A., Schmitt, B., Betz, H. & Weissenhorn, W. (2004). Structural basis of dynamic glycine receptor clustering by gephyrin. *EMBO J.*, 23, 2510-2519
- Mavrikakis, M., Iseni, F., Mazza, C., Schoehn, G., Ebel, C., Gentzel, M., Franz, T. & Ruigrok, R.W. (2003). Isolation and characterisation of the rabies virus N degrees-P complex produced in insect cells. *Virology*, 305, 406-414



Rainer Pepperkok

PhD 1992, University Kaiserslautern, Germany.
Postdoctoral research at University of Geneva, Switzerland.
Lab Head at the Imperial Cancer Research Fund, London, UK.
At EMBL since 1998.

Advanced Light Microscopy Core Facility

The Advanced Light Microscopy Facility (ALMF) offers a collection of state-of-the-art light microscopy equipment and image processing tools. The ALMF was set up as a cooperation between EMBL and industry to improve communication between users and producers of high-end microscopy technology and to support in-house scientists and visitors in using light microscopy methods for their research. The facility also organises regular international courses to teach advanced light microscopy methods.

Major projects and accomplishments

- The ALMF presently manages 17 top-of-the-line microscope systems and four image analysis workstations from leading industrial companies.
- More than 30 visitors per year come to carry out their own experiments in the ALMF or to evaluate microscopy equipment.
- The ALMF was the seed for the European Light Microscopy Initiative (ELMI) that establishes links with light microscopy facilities throughout Europe.
- Usage of the facility has exceeded 24,000 hours per year.

Services provided

- Assisting all EMBL groups and visitors through the microscopic imaging process: project planning, sample preparation, staining, microscope selection and use, image processing, presentation, data transfer and storage.
- Developing accessory software and microscopy equipment, co-developments with industrial partners, pre-evaluation of commercial equipment.

Technology partners

The ALMF presently has collaborations with the following companies:

Applied Precision*
Bitplane*
Carl Zeiss*
Cell Biotrading*
Definiens
Eppendorf*
jpk Instruments*
Improvision
Leica Microsystems*
MetaSystems
Olympus Europe*
Olympus BioSystems*
Perkin Elmer*
Scientific Volume Imaging
T.I.L.L. Photonics
Visitron*

*The ALMF also serves as a reference site for these companies.

Selected references

Ben-Tekaya, H., Miura, K., Pepperkok, R. & Hauri, H.P. (2005). Live imaging of bidirectional traffic from the ERGIC. *J. Cell Sci.*, 118, 357-367

Brunet, S., Sardon, T., Zimmerman, T., Wittmann, T., Pepperkok, R., Karsenti, E. & Vernos, I. (2004). Characterisation of the TPX2 domains involved in microtubule nucleation and spindle assembly in *Xenopus* egg extracts. *Mol. Biol. Cell*, 15, 5318-5328

Vlachou, D., Zimmermann, T., Cantera, R., Janse, C.J., Waters, A.P. & Kafatos, F.C. (2004). Real-time, in vivo analysis of malaria ookinete locomotion and mosquito midgut invasion. *Cell Microbiol.*, 6, 671-685

Zimmermann, T., Rietdorf, J. & Pepperkok, R. (2003). Spectral imaging and its applications in live cell microscopy. *FEBS Lett.*, 546, 87-92

EMBL-Grenoble, France

The EMBL Outstation in Grenoble, France, is situated in one of Europe's most beautiful locations, the heart of the French Alps, with a view of snow-covered mountains and the ski slopes. The Outstation, a laboratory of about 90 people, shares a campus with the European Synchrotron Radiation Facility (ESRF), which produces some of the world's most intense X-ray beams, and the Institut Laue Langevin (ILL), which provides high-flux neutron beams. The Outstation collaborates very closely with these facilities in building and operating beam-lines for macromolecular crystallography, in developing the associated instrumentation and techniques and in providing biochemical laboratory facilities and expertise to help external visitors making measurements.

Within this exciting context, the Outstation has a very active in-house research programme in the structural biology of cellular processes, making use of a wide range of techniques including molecular biology, biochemistry, electron microscopy, light scattering, neutron scattering, X-ray crystallography and computing. The availability of such a range of techniques, combined with the neighbouring large-scale facilities, is vital to the success of ambitious projects in modern structural molecular biology.

A strong tradition in studying systems involving protein-nucleic acid complexes and viruses has contributed to making the Outstation a leader in international high-throughput structural genomics projects. The structural work on aminoacyl-tRNA synthetases is particularly well known. A number of synthetases were first cloned at EMBL-Grenoble and various different synthetase structures have been determined, including several in complex with cognate tRNA. Studies of protein-RNA interactions have been extended to the mammalian signal recognition particle and other proteins involved in translational regulation and RNA transport. The analysis of protein-DNA interactions and mechanisms of transcriptional regulation is another important topic here. Structural analysis of eukaryotic transcription factor DNA complexes like the first STAT/DNA complex is now moving towards the analysis of larger complexes involved in transcriptional regulation.

Another major focus is the study of RNA viruses, such as influenza, rabies and Ebola, with the aim of understanding how they replicate and assemble. In parallel, studies of the structure and function of proteins involved in viral and cellular membrane fusion is actively pursued (e.g. HIV gp41 and proteins involved in vesicle transport). Some of the projects at the Outstation depend on close interactions with colleagues at EMBL-Heidelberg and collaborations are underway on proteins involved in nucleocytoplasmic transport, translational regulation and RNA metabolism.

A new development at the Outstation is the introduction of automated, high-throughput methods to make structure determination more efficient. This is connected to the Outstation's involvement in the EU-funded SPINE project (Structural Proteomics in Europe) and to the establishment with the neighbouring ESRF, ILL and French national Institut de Biologie Structurale (IBS) of a Partnership for Structural Biology (PSB). A new team has been established at the Outstation who will implement robotic systems for crystallisation, protein expression and the development of selection methods for finding protein fragments or mutants with improved solubility properties.





Stephen Cusack

PhD 1976, Imperial College, London, UK.

Postdoctoral research at EMBL-Grenoble.

Staff Scientist at EMBL-Grenoble 1980-1989.

Group Leader and Head of Outstation since 1989. Joint appointment with the Gene Expression Unit.

Structural biology of RNA-binding and viral proteins

Previous and current research

We use X-ray crystallography as a central technique to study the structural biology of protein-RNA complexes involved in RNA maturation and translation. Additionally we work on the structure of adenovirus capsid proteins involved in host-cell entry.

In eukaryotic cells, nascent Pol II RNA transcripts (mRNA or snRNA) are rapidly given a m7Gppp cap at the 5' end. The nuclear cap-binding complex (CBC) binds to this tag and mediates interaction with nuclear RNA processing machineries for splicing, poly-adenylation and transport. We determined the structure of human CBC, a 90KDa heterodimeric protein and its complex with a cap analogue and are currently working on structures of several other proteins involved in cap-dependent processes.

We are interested in understanding the structure and function of the mammalian signal recognition particle (SRP), a ribonucleoprotein particle (300nt RNA and six proteins) essential for targeting of signal-peptide-containing proteins to the ER membrane. We have determined the crystal structure of the Alu-domain (heterodimer SRP14/9 complexed with 88nt RNA) which is responsible for ribosomal translation arrest during the targeting process.

Aminoacyl-tRNA synthetases play an essential role in protein synthesis by charging specifically their cognate tRNA(s) with the correct amino acid. We aim to obtain atomic resolution structural information to help us understand the catalytic mechanism of the enzymes and their substrate specificity for ATP, cognate amino acid and tRNA. Most recently we have solved the structures of two class I enzymes, tyrosyl-tRNA synthetase (and its tRNA complex) and leucyl-tRNA synthetase. The latter is particularly interesting as it contains a large editing domain able to hydrolyse mischarged amino acids.

We have been studying the structure of the adenovirus capsid proteins the fibre and penton base which carry primary and secondary receptor binding activity respectively to allow virus entry into cells. We have solved the structure of the receptor binding domain and part of the fibrous shaft of Ad2 fibre as well as the part of the adenovirus receptor (a human protein known as CAR) to which the fibre binds. We have used this structural information (in collaboration with a company) to engineer mutant adenoviruses that have altered tropism for use in gene therapy applications.

Future projects and goals

We are currently involved in a number of projects related to RNA metabolism, our goal being to obtain structures of the multi-protein, often transient, complexes involved. These include structural studies on capping and decapping enzymes, proteins involved in nonsense mediated decay (NMD) such as Upf1, Upf2 and Upf3, and proteins involved in snoRNP assembly. The work on the signal recognition particle is still focused on the structure of the Alu domain and how it interacts with the ribosome to arrest translation. Work is continuing on several aminoacyl-tRNA synthetase systems and their substrate complexes, notably the leucyl- and prolyl-systems, both of which have editing activities. We are also involved in drug design work on synthetases from pathogenic bacteria and tropical disease producing nematodes. The focus of the adenovirus work is now on the fibre proteins of serotypes which have alternative receptors to CAR, on exploiting the structure of the penton base that we have recently determined, and in combining our results with those of cryo-EM to obtain a quasi-atomic structure of the whole virus.

Selected references

Zubieta, C., Schoehn, G., Chroboczek, J. & Cusack, S. (2005). The structure of the human adenovirus 2 penton. *Mol. Cell*, 17, 121-135

Kadlec, J., Izaurralde, E. & Cusack, S. (2004). The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat. Struct. Mol. Biol.*, 11, 330-337

Lincecum, T.L., Tuko, M., Yaremchuk, A., Mursinna, R.S., Williams, A.M., Sproat, B.S., Van Den Eynde, W., Link, A., Van Calenbergh, S., Grotli, M., Martinis, S.A. & Cusack, S. (2003). Structural and mechanistic basis of pre- and posttransfer editing by leucyl-tRNA synthetase. *Mol. Cell*, 11, 951-963

Mazza, C., Segref, A., Mattaj, I.W. & Cusack, S. (2002). Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.*, 21, 5548-5557

Florent Cipriani

BSc 1974, Physics, University of Grenoble, France.
Senior engineer in nuclear and medical industries.
At EMBL-Grenoble since 1991.
Joint Team Leader since 2003.



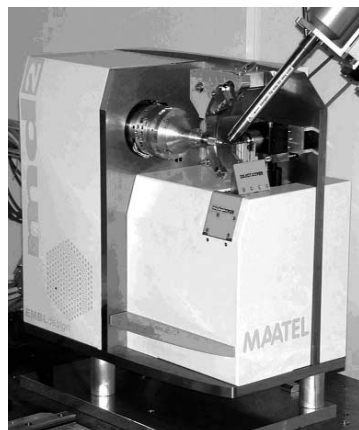
Diffraction Instrumentation Team

Previous and current research

Our activity is governed by the demands expressed by the EMBL crystallographers and by the need to maintain the EMBL/ESRF Joint Structural Biology Group (JSBG) MX beamlines at the best level. Our major themes are the improvement of data quality, the development of instruments for new data collection methods, and the development of technologies for beamline automation. We also develop Neutron Image Plate Diffractometers at our second neighbouring neutron large-scale facility, the Institute Laue-Langevin. The synergy between the Diffraction Instrumentation Team and the Synchrotron Crystallography Team gives the Instrumentation Team the capability of bringing to reality the most challenging scientific demands. We take a particular interest in making our technology available to the scientific community. Most of our developments are licensed to industry in association with EMBLEM. We are especially focused on the user-end of the MX beamlines and develop most of the devices used in experimental hutches: slit-boxes, fast low jitter piezoelectric shutters, air bearing goniometers, on-beam axis video microscopes and κ goniometers. We are strongly involved in the automation of the beamlines. Two projects are of special importance: the microdiffractometer or MD2 (figure 1) and the automatic sample changer (figure 2). A lighter version of the MD2 named minidiffractometer or MD2M has also been designed. Most of the JSBG MX beamlines will receive both the MD2/MD2M and the sample changer in the near future. Beamline automation requires automating all the crystal processing steps. Crystal alignment is one of the critical steps we are facing. A first implementation of our c3d alignment library for loop and crystal centring is currently being tested at the ESRF ID14-3 beamline on a MD2 microdiffractometer. We also have made a major contribution in the definition of the first European standard for storage of frozen crystals.

Future projects and goals

Our major goals are the improvement of data quality and automation of the MX beamlines. For example, the MiniKappa head should lead to new data collection strategies. Specific methods and software will be developed by the Synchrotron Crystallography Team to take full advantage of this new device. New crystal alignment methods like UV fluorescence and X-ray scanning will be explored to overcome the limits of the optical methods currently used. Two important issues for collecting good data are the optimisation of crystal quality at freezing time and the control of radiation damage. Future projects with the Synchrotron Crystallography Team are underway to develop new instruments for monitoring crystal humidity at freezing time and to track online the crystal radiation damage by spectrophotometry.



MD2 Microdiffractometer.



SC3 Sample changer.

Selected references

- Arzt, S., Beteva, A., Cipriani, F., Delageniere, S., Felisaz, F., Forstner, G., Gordon, E., Launer, L., Lavault, B. & Leonard, G. et al. (2005). Automation of macromolecular crystallography beamlines. *Prog. Biophys. Mol. Biol.*, 89, 124-152
- Andrey, P., Lavault, B., Cipriani, F., & Maurin Y. (2004). Using image analysis for automated crystal positioning in a synchrotron X-ray beam for high-throughput macromolecular crystallography. *J. Appl. Cryst.*, 37, 265-269
- Perrakis, A., Cipriani, F., Castagna, J.C., Claustre, L., Burghammer, M., Riek, C. & Cusack, S. (1999). Protein micro-crystals and the design of a microdiffractometer: current experience and plans at EMBL and ESRF/ID13. *Acta Crystallogr. D Biol. Crystallogr.*, 55, 1765-1770
- Cipriani, F., Castagna, J.C., Claustre, L., Wilkinson, C. & Lehmann, M.S. (1997). Large area neutron and X-ray image-plate detectors for macromolecular biology. *Nucl. Instr. Meth.*, A392, 471-474



Darren Hart

PhD 1996, Oxford University, UK.
 Postdoctoral research at Cambridge University, UK.
 Group Leader at Sense Proteomic Ltd., Cambridge, UK.
 Team Leader at EMBL-Grenoble since 2003.

High-throughput protein technologies

Previous and current research

Combinatorial methods in biology (e.g. directed evolution, phage display) are used to address problems that are too complex for rational approaches. First, large random libraries of variants are synthesised in which possible solutions exist at a very low frequency. Second, screening and selection processes are used to isolate the rare hits from the library. Previously, I have used these techniques to explore the function of transcription factors and in the development of functional protein microarrays. At EMBL, we have applied these methods to a common bottleneck of structural biology, that of soluble protein expression.

In our high-throughput process, ESPRIT (Expression of Soluble Proteins by Random Incremental Truncation), all truncations of a target protein, are synthesised as a random library. These are then screened using robots to print “protein arrays” and probed to identify soluble variants. In this way, high value proteins of biological and medicinal importance that have resisted structural analysis due to poor recombinant expression may yield soluble, well-expressed fragments for study by protein crystallography and NMR.

Future projects and goals

We are exploring “directed evolution” strategies involving gene fragmentation and point mutagenesis for solving the expression problems of previously intractable eukaryotic proteins. And as part of the European 3D Repertoire project, we are adapting the high-throughput robotic methods of structural genomics for the study of protein-protein interactions and applying them to a set of targets. The aim is to produce large amounts of high quality, recombinant complex and to proceed to structure determination.

Selected references

DeLumley, Hart, D.J., Cooper, M.A., Symeonides, S and Blackburn, J.M. (2004). A biophysical characterisation of factors controlling dimerisation and selectivity in the NJ-κB and NFAT families. *J. Mol. Biol.*, 339, 1059-1075.

Boutell, J.M., Hart, D.J., Godber, B.L., Kozlowski, R.Z. & Blackburn, J.M. (2004). Functional protein microarrays for parallel characterisation of p53 mutants. *Proteomics*, 4, 1950-1958

Speight, R.E., Hart, D.J., Sutherland, J.D. & Blackburn, J.M. (2001). A new plasmid display technology for the in vitro selection of functional phenotype-genotype linked proteins. *Chem. Biol.*, 8, 951-965

Hart, D.J., Speight, R.E., Sutherland, J.D. & Blackburn, J.M. (2001). Analysis of the NJ-κB p50 dimer interface by diversity screening. *J. Mol. Biol.*, 310, 563-575

A custom-built picking and gridding robot is used to format and analyse large libraries of clones with the aim of identifying members exhibiting improved recombinant protein expression.



José Márquez

PhD 1997, University of Valencia, Spain.
Postdoctoral research at EMBL.
Staff Scientist at EMBL-Grenoble since 2003.



The high-throughput crystallisation laboratory at EMBL-Grenoble

Previous and current research

Finding conditions in which biological macromolecules would form crystals is recognised as one of the major bottlenecks in structural biology. Once macromolecules are purified they need to be assayed for crystallisation with a collection of precipitants under different chemical environments. This leads to the need to perform hundreds of experiments, consuming large amounts of sample and taking time. At the EMBL-Grenoble Outstation we have established a high-throughput crystallisation platform with the aim to increase the success rate and speed up the process of crystal structure determination. In this platform the whole process of crystallisation screening is automated through the introduction of liquid handling, crystallisation and crystal imaging robots. The technology introduced allows us to perform experiments using extremely low volumes of sample, which makes it possible to perform extensive screening even when the amount of sample is limited. This platform, which started to operate in September 2003, has now more than a hundred registered users and more than half a million individual crystallisation experiments have already been performed. The high-throughput crystallisation laboratory is not only open to EMBL researchers but also to all the members of the Partnership for Structural Biology (PSB), which includes the ESRF, the ILL, the IBS and the IVMS, and represents one of its core technological platforms. In addition to offering automated crystallisation resources, the HTX lab is actively involved in the development of new methods and concepts in macromolecular crystallography and works in close co-ordination with the high-throughput protein expression and synchrotron instrumentation groups at the Outstation. We are also collaborating with the EBI and the Hamburg Outstation towards the development of a Laboratory Information Management System (LIMS) for macromolecular crystallography.

Future projects and goals

Structure and function of transcriptional regulators: in eukaryotes the RNA polymerase II (RNAPII) is responsible for transcription of most protein-coding genes. Though RNAPII is able by itself to recognise core promoter elements and confer low levels of transcription, regulated transcription requires the contribution of gene specific factors binding at promoter regions. These factors are commonly known as transcriptional activators or repressors. In 1990, Kornberg realised that transcriptional activators and RNAPII did not interact directly, but required the contribution of a third element that he called the mediator complex. This complex is composed of about 25 polypeptides and mediates interactions between activators/repressors and the core RNAPII. The mediator complex is conserved in all eukaryotic species including humans and yeast and is essential to support regulated transcription. Though different studies have revealed its composition and general organisation there is still very little structural information available at atomic resolution. We are currently applying high-throughput structural methods to understand the function of the mediator complex and other transcriptional regulatory proteins.

Selected references

- Márquez, J.A., Smith, C.I., Petoukhov, M.V., Lo Surdo, P., Mattsson, P.T., Knekt, M., Westlund, A., Scheffzek, K., Saraste, M. & Svergun, D.I. (2003). Conformation of full-length Bruton tyrosine kinase (Btk) from synchrotron X-ray solution scattering. *EMBO J.*, 22, 4616-4624
- Márquez, J.A., Hasenbein, S., Koch, B., Fieulaine, S., Nessler, S., Russell, R.B., Hengstenberg, W. & Scheffzek, K. (2002). Structure of the full-length HPr kinase/phosphatase from *Staphylococcus xylosus* at 1.95 Å resolution: Mimicking the product/substrate of the phospho transfer reactions. *Proc. Natl. Acad. Sci. USA*, 99, 3458-3463
- Márquez, J.A. & Saraste, M. (2001). Pleckstrin homology (PH) domains. In "Encyclopedia of Molecular Medicine," Creighton, T. (Ed.), John Wiley & Sons, New York
- Márquez, J.A., Pascual-Ahuir, A., Proft, M. & Serrano, R. (1998). The Ssn6-Tup1 repressor complex of *Saccharomyces cerevisiae* is involved in the osmotic induction of HOG-dependent and -independent genes. *EMBO J.*, 17, 2543-2553



Christoph W. Müller

PhD 1991, University of Freiburg, Germany.

Postdoctoral research at Harvard University, Cambridge, USA.

Group Leader at EMBL-Grenoble since 1995.

Senior Scientist since 2000. Joint appointment with the Gene Expression Unit.

Selected references

Petosa, C., Schoehn, G., Askjaer, P., Bauer, U., Moulin, M., Steuerwald, U., Soler-Lopez, M., Baudin, F., Mattaj, J.W. & Müller, C.W. (2004). Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex. *Mol. Cell*, 16, 761-775

Soler-Lopez, M., Petosa, C., Fukuzawa, M., Ravelli, R., Williams, J.G. & Müller, C.W. (2004). Structure of an activated Dictyostelium STAT in its DNA-unbound form. *Mol. Cell*, 13, 791-804

Grune, T., Brzeski, J., Eberharter, A., Clapier, C.R., Corona, D.F., Becker, P.B. & Müller, C.W. (2003). Crystal structure and functional analysis of a nucleosome recognition module of the remodelling factor ISWI. *Mol. Cell*, 12, 449-460

Cohen, S.X., Moulin, M., Hashemolhosseini, S., Kilian, K., Wegner, M. & Müller, C.W. (2003). Structure of the GCM domain-DNA complex: a DNA-binding domain with a novel fold and mode of target site recognition. *EMBO J.*, 22, 1835-1845

Müller, C.W. (2001). Transcription factors: global and detailed views. *Curr. Opin. Struct. Biol.*, 11, 26-32

Structural biology of transcriptional regulation

Previous and current research

Our group is interested in the mechanisms of transcriptional regulation and nuclear transport. We use structural information obtained by X-ray crystallography or electron microscopy combined with biochemical and other biophysical techniques to gain insight into these complex cellular processes.

Currently, we are pursuing three major research themes. We are studying structures of different eukaryotic transcription factor/DNA complexes involved in the regulation of fundamental physiological processes like immune response, growth control and cell differentiation. Our work has been focusing on Rel/NF κ B proteins, T-box transcription factors, STAT proteins and recently GCM transcription factors. Our research aims to obtain a detailed understanding of the global architecture of these transcription factors, their interaction with DNA and with other proteins.

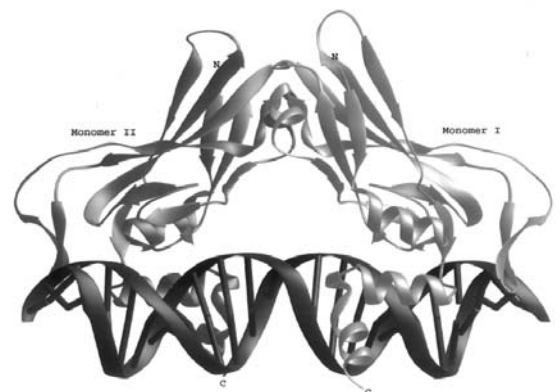
Many eukaryotic transcription factors are found in latent form in the cytosol and upon activation are imported into the nucleus, where they bind to their DNA target sequences. Regulated export from the nucleus is one important mechanism for their subsequent inactivation. Our research is concentrating on the structural and functional analysis of proteins of the import β superfamily and their interactions with different transport substrates and other effector molecules.

Our research is now moving towards the structural and functional analysis of multi-protein complexes involved in transcriptional regulation. Systems currently under investigation include yeast RNA polymerase III, the general transcription factor complex TFIIC from yeast, and nucleosome remodelling complexes.

Future projects and goals

Our work on transcription factor/DNA complexes, nuclear transport receptors and multi-protein complexes involved in transcriptional regulation will continue. We want to provide structural information about multi-protein and protein-DNA complexes in order to better understand their complicated spatial architecture but also the dynamics of these multi-component assemblies. The obtained structural information will help us to better understand their molecular function in the cellular context.

Work on RNA polymerase III transcription and the different protein complexes involved in this process will be further extended. We will continue the structural analysis of nucleosome remodelling factors but also start analysing other systems involved in the organisation and modification of chromatin structure. In the analysis of multi-protein complexes we will increasingly combine X-ray crystallography with electron microscopy. The detailed molecular information provided by our research will contribute to a deeper understanding of transcriptional regulation and nuclear transport.



Structure of the T domain-DNA complex of the Brachyury transcription factor at 2.5 Å resolution.

Winfried Weissenhorn

PhD 1991, University of Munich, Germany.
Postdoctoral research and Instructor at Harvard Medical School, Boston, USA.
Group Leader at EMBL-Grenoble since 1998.
Joint appointment with Cell Biology and Biophysics Unit.



Protein mediated membrane reorganisation

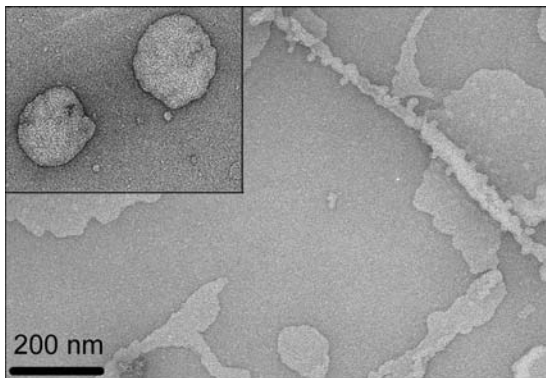
Previous and current research

Our research centres on the structural basis of protein-mediated changes in membrane structures with a specific focus on enveloped virus entry and egress. While the entry process involves only viral structural proteins, egress is a complex interplay between viral and cellular protein machines. We have previously determined the crystal structures of core domains from the HIV and Ebola virus membrane fusion proteins and we have recently shown that liposome bound trimeric gp41 fragments can inhibit membrane fusion, which implicates glycoproteins outside of the direct contact zone in the fusion process. We are currently working towards the understanding of intermediate conformations of HIV-1 gp41 postulated to drive the membrane fusion process as well as the final post-fusion conformation, including all membrane anchors.

Matrix proteins are the driving force for enveloped virus assembly by interacting with cellular membranes and recruiting cellular factors to the site of assembly and budding. We have previously determined crystal structures of the Ebola virus matrix protein VP40 in the monomeric state and in its octameric form in complex with single-stranded RNA. We have further shown that the monomeric conformation is metastable and can be converted into oligomeric forms, which are essential for the virus life cycle. Ebola VP40 like HIV-1 Gag also interacts with cellular factors such as Tsg101, which has been implicated in the down regulation of plasma membrane receptors and their sorting into multivesicular bodies (MVB), destined for lysosomal degradation. Enveloped viruses recruit all or part the MVB machinery such as ESCRT (Endosomal Sorting Complex Required for Transport) complexes, which is composed of at least 43 different protein-protein interactions to the site of budding. We have recently solved the structures of some components of the ESCRT machinery and continue our studies are to understand the role MVB proteins and complexes in membrane deformation, vesicle formation and membrane fission processes.

Future projects and goals

Structural and functional studies of cellular and viral proteins that mediate and control membrane interaction, membrane deformation, membrane fusion and membrane fission processes.



Alix/AIP1 is an essential component of multi vesicular body (MVB) formation and is recruited by viral late domains such as HIV-1 Gag to the site of budding. Negative staining electron microscopy shows initially round liposomes before (upper left panel) and after incubation with Alix/AIP1, which induces membrane deformation and tubulation.

Selected references

- Hoenen, T., Volchkov, V., Kolesnikova, L., Mittler, E., Timmins, J., Ottmann, M., Reynard, O., Becker, S. & Weissenhorn, W. (2005). VP40 octamers are essential for Ebola virus replication. *J. Virol.*, 79, 1898-1905
- Sola, M., Bavro, V.N., Timmins, J., Franz, T., Ricard-Blum, S., Schoehn, G., Ruigrok, R.W., Paarmann, I., Saiyed, T., O'Sullivan, G.A., Schmitt, B., Betz, H. & Weissenhorn, W. (2004). Structural basis of dynamic glycine receptor clustering by gephyrin. *EMBO J.*, 23, 2510-2519
- Schibli, D.J., & Weissenhorn, W. (2004). Class I and class II viral fusion protein structures reveal similar principles in membrane fusion. *Mol. Membr. Biol.*, 6, 361-371
- Gomis-Ruth, F.X., Dessen, A., Timmins, J., Bracher, A., Kolesnikowa, L., Becker, S., Klenk, H.D. & Weissenhorn, W. (2003). The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. *Structure (Camb)*, 11, 423-433

EMBL-Hamburg, Germany

The Hamburg Outstation is situated in one of the most beautiful areas of this Hanseatic city in the northern part of Germany. Hamburg is centred around its harbour, which has a long tradition of overseas trade, and offers a wide variety of cultural activities. In the close vicinity of EMBL, there are diverse neighbourhoods ranging from residential, wealthy suburbs like Blankenese, to areas with a lot of students, entertainment and a multicultural ambience like St. Pauli, Altona and Ottensen.

EMBL-Hamburg is located on the site of DESY (German Synchrotron Research Centre) that provides synchrotron radiation (SR) through its DORIS positron storage ring. This radiation is used to study the structure and function of proteins using state-of-the-art equipment and methods. The Outstation operates seven SR beamlines, five of which are dedicated to biocrystallography, one to small angle X-ray scattering of biological samples and one to X-ray absorption spectroscopy (EXAFS). EMBL-Hamburg has a well-established record for the development of novel, innovative technologies in biological applications of SR. These started with the first diffraction and scattering experiments in molecular structural biology in Hamburg more than thirty years ago. Later, the first imaging plate scanner, now standard equipment in any biocrystallographic laboratory, was developed by the EMBL instrumentation group, and was subsequently commercialised by the spin-off company MarResearch. At present, world-leading software packages for the automation of data interpretation have been developed and are used internationally in a large number of projects. One is the ARP/wARP package that, in its current version, allows automatic X-ray structure determination with X-ray data higher than 2.5 Å resolution within a few hours. The other one is called ATSAS and allows the automatic interpretation of small angle X-ray scattering data for structural shape determination.

These research developments are paralleled by an integrated approach to carry out scientifically demanding projects in structural biology. For this, EMBL-Hamburg offers facilities in molecular biology, heterologous expression in prokaryotic and eukaryotic hosts, protein purification, biophysical characterisation and crystallisation, complementing our X-ray data acquisition and processing infrastructure.

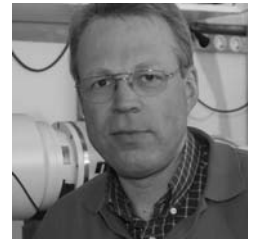
Our biological interests include, inter alia, regulation of transcription and translation, viral replication, protein-ligand interactions in signalling proteins and protein kinases, giant muscle proteins and protein translocation into peroxisomes. Our activities in small angle X-ray scattering and EXAFS provide unique opportunities to provide insight, for instance, into large protein complexes and protein with metal centres.

EMBL-Hamburg is one of the nodes of the European Structural Genomics Project, SPINE. EMBL is also coordinating a Structural Genomics consortium targeting 150 genes of *Mycobacterium tuberculosis*, by using novel technologies in comparative proteomics. As part of the latter project a high-throughput crystallisation facility was installed during 2003. DESY is planning to dedicate the PETRA ring for future SR applications at the beginning of 2007 and our aim is to continue to provide world-leading synchrotron radiation beamline facilities for structural biology on the PETRA ring during August 2007.



Matthias Wilmanns

PhD 1990, University of Basel, Switzerland.
Postdoctoral research at University of California, Los Angeles, USA.
Group Leader at EMBL-Heidelberg 1993-1997.
Head of EMBL-Hamburg since 1997.



Protein X-ray crystallography

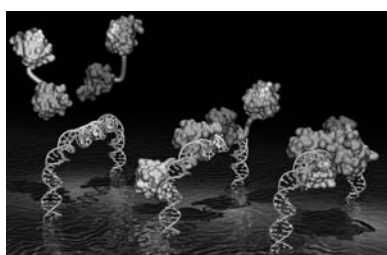
Previous and current research

Our research aims are driven by the availability of synchrotron radiation at EMBL-Hamburg and DESY, and by the emerging opportunities to use these facilities for automated and high-throughput approaches. We are involved in two structural proteomics projects, the first on the depository of the 3D proteome of SH3 domains from yeast. This proteome comprises 30 genes, and our activities are on the structures of the apo SH3 domains, SH3 domain-ligand peptide complexes and SH3 domain-protein complexes. The second project is the determination of structures from the proteome of *Mycobacterium tuberculosis*. Our focus is entirely on targets that have been selected by functional genomics approaches monitoring the gene/protein expression patterns of Mtb targets under different growth conditions, including infected human lung material. Highlights of other research activities include:

The giant muscle protein titin: titin is the largest gene product of the human genome, and it comprises up to 38,000 residues in its largest isoform. It is known as the third filament of the muscle sarcomere and is involved in multiple functions, such as acting as a “molecular ruler” keeping major components of the sarcomere in place, muscle development, passive elasticity of the muscle sarcomere and muscle signalling. Titin comprises 300 predicted protein domains, and we consider it a structural proteomics type project. At present, we have determined X-ray structures from the Z-disk segment of titin (the N-terminus in complex with a ligand called telethonin), from the I-band (the IG structures I1 and I27) and from the A-band (titin kinase, a double IG domain construct 167-A168). Several other projects are in the pipeline, and we are moving more and more in the direction of solving structures of titin domain-protein ligand complexes.

Calcium/calmodulin regulated kinases: the human “kinome” consists of about 2,000 protein kinases and a comparable number of phosphatases. Protein phosphorylation/dephosphorylation is a simple key regulation mechanism of protein activities found in almost any biological process. Our interest is in calcium/calmodulin regulated protein kinases, and our goal is to determine protein kinase-regulator complexes. Recently we determined a novel structure of the Death Related Protein Kinase1, which is involved in a number of tumour progression and metastasis events if malfunctioning.

Protein translocation into peroxisomes: there is only limited knowledge about the molecular events associated with protein translocation into different cell organelles. We are interested in the import of proteins into peroxisomes, which contain a C-terminal PTS1 signal recognition motif. These proteins are liganded by the peroxisome receptor Pexp5, which in turn docks on a multiple component complex at the peroxisomal membrane, leading to translocation of the cargo protein by a basically unknown process. Our approach is to determine the PTS1 docking complex in several iterations. First, we started solving the structure of the SH3 domain of one of its components, Pex13p, in complex with peptides from two other components, Pex5p and Pex14p. More recently, we determined the first structure of a peroxisome receptor (Pex5p)-cargo protein (SCP2) complex, demonstrating that the cargo remains folded and functionally intact. Further binary, ternary etc. complexes are in the pipeline.



Schematic cartoon of the assembly process of Oct POU factors onto specific DNA response motifs. Each of the two POU domain molecules consists of two sub-domains, a POU-homeo domain and a POU-specific domain. The DNA-binding surface patches are indicated in different grey shades. This figure was made with the help of Ansgar Philippsen, University of Basel, Switzerland.

Selected references

- Kuper, J., Doenges, C. & Wilmanns, M. (2005). Two-fold repeated ($\beta\alpha$)(4) half-barrels may provide a molecular tool for dual substrate specificity. *EMBO Rep.*, 6, 134-139
- Vega, M.C., Zou, P., Fernandez, F.J., Murphy, G.E., Sterner, R., Popov, A. & Wilmanns, M. (2005). Regulation of the heterooctameric ATP phosphoribosyl transferase complex from *Thermotoga maritima* by a tRNA synthetase-like subunit. *Mol. Microbiol.*, 55, 675-686
- Barabas, O., Pongracz, V., Kovari, J., Wilmanns, M. & Vertessy, B.G. (2004). Structural insights into the catalytic mechanism of phosphate ester hydrolysis by dUTPase. *J. Biol. Chem.*, 279, 42907-42915
- Remenyi, A., Schöler, H.R. & Wilmanns, M. (2004). Combinatorial control of gene expression. *Nat. Struct. Mol. Biol.*, 11, 812-815



Christoph Hermes

PhD 1981, Technical University,
Munich, Germany.
Group Leader at EMBL-Hamburg since 1982.

Instrumentation for structural research at EMBL-Hamburg

Previous and current research

Research carried out at the Hamburg Outstation since 1974 covers a broad spectrum of X-ray methods used for structural investigations in biology. The early diffraction studies of insect flight muscle has developed into a more general approach for the observation of time-dependent phenomena by applying small angle scattering techniques in combination with fast detectors and sophisticated data acquisition methods. X-ray absorption spectroscopy (EXAFS) permits high-precision investigation of the environment around specific metal atoms in proteins, while the complete three-dimensional picture can be obtained by protein X-ray crystallography, a method which has become the dominant structural research tool in molecular biology not only at synchrotron radiation sources.

Each of the above-mentioned methods has specific instrumentation needs and our group designs, constructs and builds the appropriate equipment. Our activities include mechanical engineering, vacuum technology, X-ray optics, and data acquisition and control electronics. Recently we rebuilt three beamlines on bending magnet fan K of the DORIS storage ring. Two of the stations for data collection at fixed wavelength from protein crystals have been operating successfully for a couple of years, whereas the third one, which will be mainly used for multi-wavelength anomalous diffraction (MAD) data collection, was opened for external user operation in 2005. Its key component is a fixed exit focusing double monochromator system, which allows a quick change of the X-ray energy over a wide range.

As a pilot project we have installed a sample mounting robot on one of our beamlines following the general trend to increase the level of automation of experiments leading to high-throughput facilities.

Selected references

Pohl, E., Ristau, U.,
Gehrmann, T., Jahn, D.,
Robrahn, B., Malthan, D.,
Dobler, H. & Hermes, C.
(2004). Automation of the
EMBL-Hamburg protein
crystallography beamline
BW7B. *J. Synchrotron
Radiat.*, 11, 372-377

Hermes, C. Gehrmann, T.,
Jahn, D., Ristau U.,
Robrahn, B. & Siambanis,
T. (2004). New beamlines
for protein crystallography
at the EMBL Outstation
Hamburg. *AIP Conference
Proceedings*, 705,
384-388.

Pohl, E., Gonzalez, A.,
Hermes, C. & van Silfhout,
R.G. (2001). Overview of
the tunable beamlines for
protein crystallography at
the EMBL-Hamburg
Outstation; an analysis of
current and future usage
and developments. *J
Synchrotron Radiat.*, 8,
1113-1120

Future projects and goals

We will continue to improve the spectral quality of our existing beamlines, while at the same time we plan to make our lines more “user-friendly”. This also means that we will continue to increase the level of automation of our experimental stations, which is a necessary condition to perform high-throughput data collection.

The transformation of one octant of the PETRA accelerator into a dedicated Synchrotron Radiation source by 2007 has been funded and EMBL has decided to build and operate beamlines on this unique radiation source. In this context major challenges and opportunities in the field of beamline instrumentation are coming up.

Our goal is to create optimal conditions for state-of-the-art experiments in structural biology at a modern synchrotron radiation source.

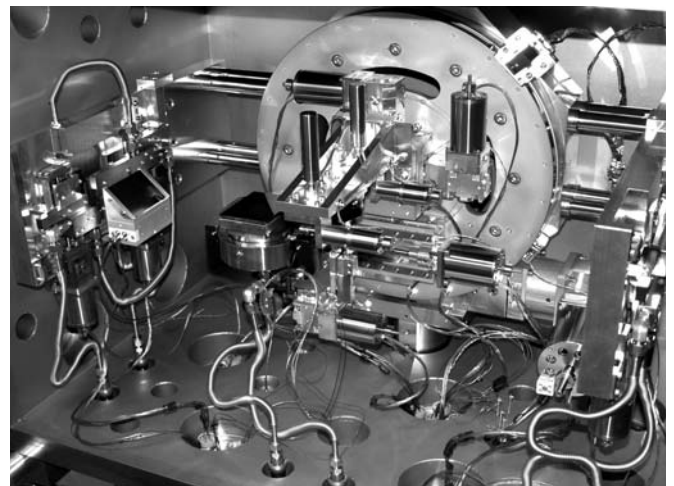


Image of the double
monochromator set-up of
the MAD beamline X12.

Victor S. Lamzin

PhD 1987, Moscow State University, Russia.
 Scientist, Inst. Biochemistry, Russian Academy of Sciences, Moscow, until 1991.
 Postdoctoral research at EMBL-Hamburg until 1995; Staff Scientist until 1997.
 Group Leader and Deputy Head of EMBL-Hamburg since 1997.



Development of methodology for macromolecular X-ray crystallography

Previous and current research

All macromolecular recognition and enzymatic processes are based on chemical principles involving valence electrons of interacting molecules, functional groups or atoms. Knowledge of their detailed electronic structure is essential for a deeper understanding of the chemical reactions underlying biological processes. High-resolution protein crystallography is currently the only experimental technique that allows the detailed atomic and electronic structure of biological substances to be studied. More than 600 atomic resolution structures at 1.2 Å or higher resolution are currently available from the PDB and the last decade has seen spectacular methodological developments in this field. The group's highlights include the experimental observation of bonding electrons in macromolecules, the pH titration of the enzyme ribonuclease A in the crystal state, and determination of protonation state of active site residues in a series of enzymes.

An ongoing development of methods in crystallography has been the software suite ARP/wARP for an automated modelling and refinement of biological macromolecules. ARP/wARP uses density-driven procedures complemented by stereochemical recognition algorithms for the model-building steps involving pattern recognition techniques and artificial intelligence methods. The software is routinely used to automatically build protein models in experimental maps, where good quality data are available at sufficient resolution. The time needed for building and refinement of a protein structure can be shortened from several days or even months to a few CPU hours on inexpensive workstations. ARP/wARP has been a key part of many structure solution pipelines in both academic and industrial laboratories. Judging from the PDB statistics, ARP/wARP is now applicable to about 75% of the structures.

A further research direction is the systematic enhancement of the diffraction quality of protein crystals by selective mutagenesis, as well as the protection of the crystals from radiation damage in the X-ray beam in order to ensure their survival through data collections that need long exposure, such as Multiple Anomalous Dispersion and ultra-high resolution data collection. For this the availability of good quality protein crystals is required. Therefore another research project is focusing on the prediction of protein-crystal-contacts to ease the crystal growing process.

Future projects and goals

We will further concentrate on structural studies of enzymatic catalysis. This will involve the development of relevant methodology in order to meet the demands set by ultra-high resolution data collection, structure refinement and interpretation, as well as the combination of X-ray crystallography with complementary quantum chemical methods.

Development of automated objective procedures for structure solution and refinement (ARP/wARP) will remain the focus of crystallographic research. Macromolecular structures include not only proteins, but also DNA, RNA or a variety of cofactors. Therefore, we will specifically address automatic modelling of such compounds and implement this within the software suite. Crystallographic studies on proteins carrying surface residue mutations and development of an approach for prediction crystal contact formation of proteins will also be pursued.

Selected references

- Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. & Tucker, P.A. (2005). Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr. D Biol. Crystallogr.*, 61, 449-457
- Morris, R.J., Perrakis, A. & Lamzin, V.S. (2003). ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.*, 374, 229-244
- Schmidt, A., Jelsch, C., Ostergaard, P., Rypniewski, W. & Lamzin, V.S. (2003). Trypsin revisited: crystallography at (sub)atomic resolution and quantum chemistry revealing details of catalysis. *J. Biol. Chem.*, 278, 43357-44336
- Meijers, R., Morris, R.J., Adolph, H.W., Merli, A., Lamzin, V.S. & Cedergren-Zeppeauer, E.S. (2001). On the enzymatic activation of NADH. *J. Biol. Chem.*, 276, 9316-9321



Wolfram Meyer-Klaucke

PhD 1996, Hamburg University, Germany.
Postdoctoral research at the Medical
University of Lübeck, Germany.
Team Leader at EMBL-Hamburg since 1998.

The role of metals in biology

Previous and current research

Metal ions play a key role in the structure and function of about 30% of all proteins. Many biocatalytic processes depend on the presence of metal ions. Research of the Meyer-Klaucke team deals with metal functionality, metal binding and metal selectivity in biological systems with a strong focus on the continuous development of X-ray absorption spectroscopy, our key technique. X-ray absorption spectroscopy allows us to individually target a specific metal and elucidate its local environment and electronic structure. X-ray absorption spectroscopy can use a multitude of sample states, ranging from whole organisms to protein solutions. Our group combines X-ray absorption spectroscopy with molecular biology, biochemistry and further methods aiming at a complete understanding of metal-related biological processes. Current research includes:

Proteins of the metallo- β -lactamase superfamily. This superfamily with an active site capable of binding up to two metal ions is utilised by nature in a variety of enzymatic processes. Beside the global metal binding motif the overall fold of α -sheets and β -helices is conserved within the superfamily. The enzymes can be grouped as follows:

- zinc dependent enzymes: class B lactamases, zinc phosphodiesterase (ZiPD);
- enzymes flexible in metal used: Glyoxalase II;
- iron dependent redox enzymes: rubredoxin oxygen oxidoreductase (ROO);
- yet uncharacterised members.

Their physiological importance varies, from putative association with cancer and antibiotic resistance to different roles in cellular detoxification. Our current aims include a deeper understanding of metal selectivity.

New metal binding motifs. The characterisation of new metal binding motifs is a challenging task in structural biology. Our work performed on CO-dehydrogenase (CODH) in parallel to crystallographic studies by the group of Prof. R. Huber resulted in the determination of the electronic and geometric structure of the new CuSMo metal binding motif. Using XAS, the previously assumed presence of Se at the active site of CODH was excluded. Instead, an Cu atom bridged by an S atom at a distance of 2.29 Å to the Mo was identified and the presence of a SCuS residue proximal to the Mo was shown. XAS analysis of CODH preparations with different activities revealed that optimal active CODH contains a binuclear (CuISMoVI(=O)₂) cluster, whereas preparations with lower activity contain in addition Cu-deficient CODH with a (HSMo(=O)OH(2)) centre.

The positions of the absorption edges determine the metal oxidation states in air-oxidised CODH (MoVI and CuI) and reduced CODH (MoIV and CuI). The Mo ion represents the redox active centre and is the likely coordination site for the oxygen atom transferred during the reaction. It is known that CuI is able to form copper carbonyls and that CO can reversibly bind to CuI solutions. Thus, we could postulate that Cu is involved in substrate binding.

Future projects and goals

The characterisation of isolated proteins has dominated research for the last decades. Prompted by the enormous technical improvements over recent years we are now learning more and more about the complex interplay of proteins. This moves topics like post-translational modifications of metal centres and regulations of metal concentrations in cells to the focus of our attention.

Selected references

Svetlitchnyi, V., Dobbek, H., Meyer-Klaucke, W., Meins, T., Thiele, B., Romer, P., Huber, R. & Meyer, O. (2004). A functional Ni-Ni-(4Fe-4S) cluster in the monomeric acetyl-CoA synthase from Carboxydotherrmus hydrogenoformans. *Proc. Natl. Acad. Sci. USA*, 101, 446-451

Gnida, M., Ferner, R., Gremer, L., Meyer, O. & Meyer-Klaucke, W. (2003). A novel binuclear (CuSMo) cluster at the active site of carbon monoxide dehydrogenase: characterisation by X-ray absorption spectroscopy. *Biochemistry*, 42, 222-230

Vogel, A., Schilling, O., Niecke, M., Bettmer, J. & Meyer-Klaucke, W. (2002). ElaC encodes a novel binuclear zinc phosphodiesterase. *J. Biol. Chem.*, 277, 29078-29085

Jochen Müller-Dieckmann

PhD 1994, Albert-Ludwigs-University, Freiburg, Germany.
Postdoctoral research at University of California, Berkeley, USA.
Associate Director, SGX, San Diego, USA until 2004.
Team Leader at EMBL-Hamburg since 2004.



High-throughput crystallisation of biological macromolecules

Previous and current research

Development of high-throughput technologies in crystallisation. X-ray crystallography is the most powerful and most widely applied technique in structural biology. Today, there are literally thousands of protein sequences from genome sequencing efforts, with little or no functional information. Many of these have important biological or medical implications. All steps within the gene-to-structure pipeline have benefited from the introduction and development of parallelised and highly-automated methods. One of the most challenging aspects of crystallography remains the growth of single and well-ordered crystals. Our group will implement and operate a high-throughput crystallisation platform accessible both to EMBL and, eventually, to the greater life sciences community. In addition to managing and running this platform, we will develop new crystallisation methods in order to further improve this process.

Proteins from Mycobacterium tuberculosis. Tuberculosis is the most prevalent infectious disease, affecting about one third of the world's population. Every year, more than eight million new cases occur and two million people die from the disease. Despite the increasing danger of tuberculosis due to increasing global mobility, only one new drug has been developed since 1972. At the same time, the number of strains with multiple resistance is on the rise. EMBL-Hamburg participates in the international Mycobacterium tuberculosis (MTb) structural genomics consortium as the coordinator of XMTB (German MTb structural genomics consortium). The goal of the XMTB is to determine the structure of 25 relevant targets, and the co-crystal structures of two small molecule compounds with each of the 25 structures.

Future projects and goals

Automation of the crystallisation process has created new bottlenecks. In order to take full advantage of the increase in throughput, a number of problems have to be addressed in the future. The three most important issues are:

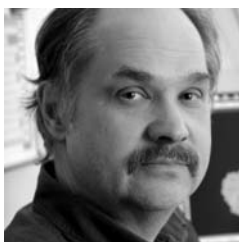
- Development and implementation of an automated crystal recognition and scoring algorithm (with V. Lamzin)
- Closing of software and hardware gap to enable the continuous optimisation of crystallisation experiments
- Development of methods to determine the quality of initial crystallisation hits in situ, without the need of prior mounting.

Additionally, we will develop a web-based application for remote viewing and scoring of the crystallisation database to make the facility accessible to the general user community by 2006.

We are currently working on 10 proteins of *M. tuberculosis* that have been validated as potential drug targets by the MPI for infection biology in Berlin or that constitute anti-targets. Most if not all of these will be screened against the small molecule library of the FMP Berlin to identify lead compounds. We plan to subsequently co-crystallise the identified compounds with their corresponding targets.

Selected references

- Müller-Dieckmann, J., Weiss, M.S. & Wilmanns, M. (2004). Protein crystallisation for X-ray. In "Encyclopedic reference of genomics and proteomics in molecular medicine" Ruckpaul, K. & Ganten, D. (Eds.). Springer Verlag Heidelberg, in press
- Müller-Dieckmann, H.J. & Kim, S.H. (2002). Structure function relationships: Chemotaxis and ethylene receptors. histidine kinase in signal transduction Elsevier Science USA. Inouye, M. (Ed.), 123-141
- Müller-Dieckmann, H.J., Grantz, A.A. & Kim, S.H. (1999). The structure of the signal receiver domain of the Arabidopsis thaliana ethylene receptor ETR1. Structure Fold Des., 7, 1547-1556
- Noland, B.W., Newman, J.M., Hendle, J., Badger, J., Christopher, J.A., Tresser, J., Buchanan, M.D., Wright, T.A., Rutter, M.E., Sanderson, W.E., Müller-Dieckmann, H.J., Gajiwala, K.S. & Buchanan, S.G. (2002). Structural studies of Salmonella typhimurium ArnB (PmrH) amino-transferase: A 4-amino-4-deoxy-L-arabinose lipopolysaccharide-modifying enzyme. Structure, 10, 1569-1580



Dmitri Svergun

PhD 1982, Institute of Crystallography, Moscow, Russia.
Dr. of Science 1997, Institute of Crystallography, Moscow, Russia.
At EMBL since 1991.

Selected references

Rosano, C., Zuccotti, S., Cobucci-Ponzano, B., Mazzone, M., Rossi, M., Moracci, M., Petoukhov, M.V., Svergun, D.I. & Bolognesi, M. (2004). Structural characterisation of the nonameric assembly of an Archaeal α -L-fucosidase by synchrotron small angle X-ray scattering. *Biochem. Biophys. Res. Commun.*, 320, 176-182

Fukuhara, N., Fernandez, E., Ebert, J., Conti, E. & Svergun, D. (2004). Conformational variability of nucleocytoplasmic transport factors. *J. Biol. Chem.*, 279, 2176-2181

Marquez, J.A., Smith, C.I., Petoukhov, M.V., Lo Surdo, P., Mattsson, P.T., Knekt, M., Westlund, A., Scheffzek, K., Saraste, M. & Svergun, D.I. (2003). Conformation of full-length Bruton tyrosine kinase (Btk) from synchrotron X-ray solution scattering. *EMBO J.*, 22, 4616-4624

Svergun, D.I. & Koch, M.H.J. (2003). Small angle scattering studies of biological macromolecules in solution. *Rep. Progr. Phys.*, 66, 1735-1782

Development and application of new methods for X-ray and neutron solution scattering data analysis

Previous and current research

Fundamental biological processes, such as cell-cycle control, signalling, DNA duplication, gene expression and regulation and some metabolic pathways, depend on supra-molecular assemblies and their changes over time. There are objective difficulties in studying such complex systems, especially their dynamic changes, with high resolution structural techniques like X-ray crystallography or NMR.

Small-angle X-ray scattering (SAXS) allows us to study native biological macromolecules, from individual proteins to large complexes, in solution under nearly physiological conditions. SAXS not only provides low resolution three-dimensional models of particle shapes but yields answers to important functional questions. Thus, kinetic SAXS experiments allow us to analyse structural changes in response to variations in external conditions, protein-protein and protein-ligand interactions, and to study kinetics of assembly/dissociation or folding/unfolding.

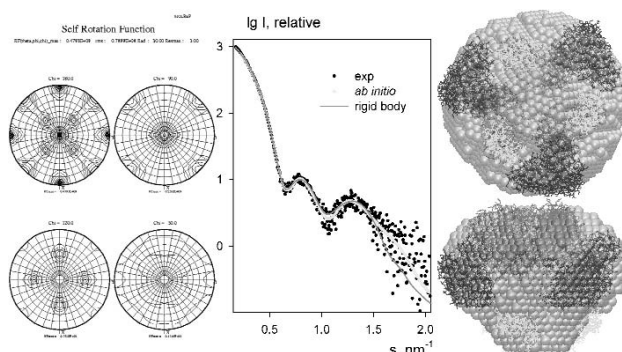
Our group runs the EMBL SAXS beamline X33 at synchrotron DESY and develops advanced methods for the analysis of X-ray and neutron scattering. Advanced mathematical methods (regularisation, non-linear optimisation, heuristic algorithms, neural networks, etc.) are employed for extracting structural information from the scattering data. Most of the ongoing projects of the external user groups at the SAXS beamline are collaborative projects, whereby the members of the group are applying the new methods to solve biological problems using SAXS. Special emphasis is put to the joint use of the results of X-ray crystallography, NMR and EM with SAXS data.

Future projects and goals

The present and future work of the group includes:

- development of algorithms for ab initio analysis of the tertiary and quaternary structure of proteins from X-ray and neutron scattering data;
- methods for rigid-body modelling of macromolecular complexes using high-resolution structures of individual domains from crystallography or NMR;
- maintenance and upgrade of the existing X33 beamline and collaborative user projects;
- design of a new high-brilliance biological SAXS beamline at the planned third-generation PETRA storage ring at DESY, Hamburg.

An example of structural analysis using SAXS. A nonameric assembly of L-fucosidase constructed by ab initio shape determination (semi-transparent beads) and rigid body modelling (Ca-traces) accounting for the non-crystallographic 3-fold symmetry observed in the crystal.



Paul Tucker

PhD 1972, University of Essex, Colchester, UK.
 Postdoctoral research at the University of Groningen, the Netherlands, University of Leicester, UK,
 Australian National University, Canberra and University of Exeter, UK.
 Lecturer in Inorganic Chemistry, University of the West Indies.
 Group Leader at EMBL-Heidelberg 1985-1998; at EMBL-Hamburg since 1998.



Structural studies of proteins from pathogens

Previous and current research

A major focus has been trying to understand the mechanism by which classical two-component systems regulate gene expression by responding to the environment. We are concentrating our efforts on systems found in *Mycobacterium tuberculosis* and have determined the structures of several response regulators. We have also identified a new type of two-component system in MtB that acts at the level of transcriptional anti-termination rather than transcriptional activation/repression.

A second area of study has been on proteins that are involved in replication of the viral genome. We have been working on the structural and functional aspects of single-stranded DNA binding proteins found in the dsDNA viruses, and are interested in how these proteins interact with other components of the replication machinery. More recently we have shifted our attention to proteins that are important for replication of the RNA viruses.

A third focus is on improving the structure determination process by, for example, developing improved phasing methods and automated structure determination.

Future projects and goals

We are making a major push on viral replication proteins, which involves participation in a major EU-funded Integrated Project (VIZIER), and we are also expanding our interest in MtB proteins involved in other signalling systems.



A cartoon of the molecular structure of the HSV1 ssDNA binding protein ICP8 with, on the right, the model of how multiple copies of the protein are expected to bind to ssDNA.

Selected references

- Sallai, L. & Tucker, P.A. (2005). Crystal structure of the central and C-terminal domain of the sigma(54)-activator ZraR. *J. Struct. Biol.*, in press
- Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. & Tucker, P.A. (2005). Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr. D Biol. Crystallogr.*, 61, 449-457
- Mapelli, M., Panjikar, S. & Tucker, P.A. (2005). The crystal structure of the herpes simplex virus 1 ssDNA-binding protein suggests the structural basis for flexible, cooperative single-stranded DNA binding. *J. Biol. Chem.*, 280, 2990-2997
- Butan, C., Van Der Zandt, H. & Tucker, P.A. (2004). Structure and assembly of the RNA binding domain of bluetongue virus non-structural protein 2. *J. Biol. Chem.*, 279, 37613-37621
- Morth, J.P., Feng, V., Perry, L.J., Svergun, D.I. & Tucker, P.A. (2004). The crystal and solution structure of a putative transcriptional antiterminator from *Mycobacterium tuberculosis*. *Structure (Camb)*, 12, 1595-1605



Manfred Weiss

PhD 1992, Albert-Ludwigs-University Freiburg, Germany.
Postdoctoral research at the University of California, Los Angeles, USA.
Senior Research Assistant at the IMB Jena.
Team Leader at EMBL-Hamburg since 2001.

Selected references

Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. & Tucker, P.A. (2005). Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr. D Biol. Crystallogr.*, 61, 449-457

Singh, R.K., Kefala, G., Janowski, R., Mueller-Dieckmann, C., von Kries, J.P. & Weiss, M.S. (2005). The high-resolution structure of LeuB (Rv2995c) from *Mycobacterium tuberculosis*. *J. Mol. Biol.*, 346, 1-11

Mueller-Dieckmann, C., Panjikar, S., Tucker, P.A. & Weiss, M.S. (2005). On the routine use of soft X-rays in macromolecular crystallography. Part III. The optimal data-collection wavelength. *Acta Crystallogr. D Biol. Crystallogr.*, in press

Weiss, M.S., Mander, G., Hedderich, R., Diederichs, K., Ermiler, U. & Warkentin, E. (2004). Determination of a novel structure by a combination of long-wavelength sulfur phasing and radiation-damage-induced phasing. *Acta Crystallogr. D Biol. Crystallogr.*, 60, 686-695

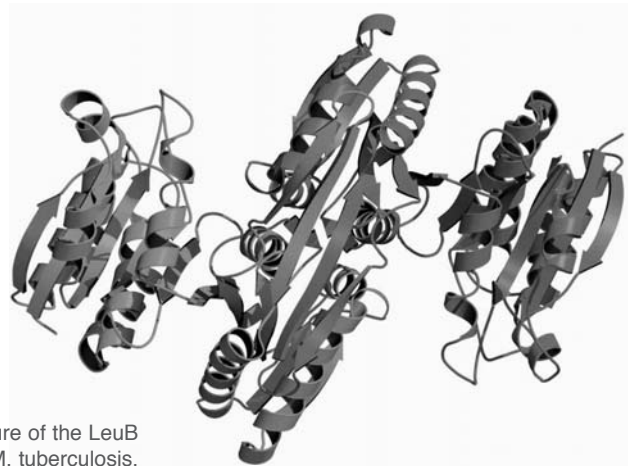
X-ray crystallography of biological macromolecules

Previous and current research

Structural analysis of proteins from *Mycobacterium tuberculosis*. According to a recent WHO report, one third of the world's population is infected with the pathogenic bacterium *Mycobacterium tuberculosis*, the causative agent of tuberculosis. The international Mtb Structural Genomics consortium, a worldwide Structural Genomics project, has been founded to determine the structures of as many *M. tuberculosis* proteins as possible. Within the Hamburg group of the consortium (M. Wilmanns, V. Lamzin, P. Tucker, W. Meyer-Klaucke and myself), proteins from the leucine and lysine biosynthesis pathways of this organism are under investigation. These proteins are potential drug targets for new anti-mycobacterial drugs.

Recently, we have been able to determine the structures of four enzymes from *M. tuberculosis*: one from the leucine biosynthesis and three from the lysine biosynthesis pathway. We are currently in the process of co-crystallising these enzymes with substrates, cofactors, inhibitors, etc. in order to better understand the metabolic pathways in *M. tuberculosis*. These structures will also provide the basis for the design of new drugs.

Development of new methods for macromolecular structure determination. Since the beginning of the 90s, when soft X-rays were first applied to address specific problems in biological crystallography, the interest in using them has been slowly but steadily growing. A few years ago, we started to collect highly redundant data sets at wavelengths between 1.5 Å and 3.0 Å on a number of model proteins, e.g. thermolysin, elastase, lysozyme, thaumatin, DNA and others. Properly scaled, the very small anomalous differences provided by the intrinsically present sulfur or phosphorus atoms turned out to be sufficiently accurate for structure determination. More recently, we have further refined the approach, and were able to determine the structure of the γ -subunit of Dsvc, which has resisted structure determination by conventional methods for a long time. Especially in the light of the large numbers of structure determination projects, which will be provided by the various Structural Genomics initiatives, this seemingly simple approach has the potential to become an important tool in the still small arsenal of methods for automated structure determination, and will be developed further.



Dimeric structure of the LeuB protein from *M. tuberculosis*.

EBI-Hinxton, UK

The European Bioinformatics Institute (EBI) lies in the fifty-five acres of landscaped parkland in rural Cambridgeshire that make up the Wellcome Trust Genome Campus, which also houses the Wellcome Trust Sanger Institute. Together, these institutes provide one of the world's largest concentrations of expertise in genomics and bioinformatics.

The mission of the EBI is to ensure that the growing body of information from molecular biology and genome research is placed in the public domain and is accessible freely to all facets of the scientific community in ways that promote scientific progress.

The EBI serves researchers in molecular biology, genetics, medicine and agriculture from academia, and the agricultural, biotechnology, chemical and pharmaceutical industries. The EBI does this by building, maintaining and making available databases and information services relevant to molecular biology, as well as carrying out research in bioinformatics and computational molecular biology.

A number of active research groups are working on problems such as classifying and understanding proteins and their interactions, analysing genomic information to discover networks and other functional entities within cells, mathematical analyses of evolutionary models, systems neurobiology and the development of novel computational annotation methods to attach knowledge about biological functions to molecules described in the databases.

Some of the key resources developed, curated and provided by the EBI include EMBL Bank (genes), UniProt (proteins), Ensembl (eukaryotic genomes), the Macromolecular Structures Database, and ArrayExpress (microarray expression data). Over the past few years, the EBI has become a world leader in integrating these types of data and has become a key partner in European initiatives to collect and understand all types of biological information.





Janet Thornton

PhD 1973, King's College and NIMR, London, UK. Postdoctoral research at Oxford University, NIMR and Birkbeck College, London. Lecturer, Birkbeck College, London, 1983. Professor, University College London, 1990. Professor & Director of Centre for Biomolecular Sciences, Birkbeck College and University College London, 1990-2001. Group Leader and Director of EMBL-EBI since 2001.

Proteomics: structure, function and evolution

Previous and current research

We use computational approaches to understand and discover more about fundamental principles governing biological processes, including the evolution of proteins and pathways, the structural basis of how proteins function and interact with other molecules, and the molecular processes involved in aging. We analyse genomes, proteomes (especially three-dimensional structures) and, most recently, transcriptomes and metabolomes, bringing together data from many sources to address specific questions. We have three major projects:

- The analysis and prediction of protein function from structural data, especially relating to structural genomics projects. We are developing new algorithms to predict protein function from structure.
- Understanding the structure, function and mechanisms of enzymes and metabolites. By gathering data on many enzymes, we aim to understand more about how enzymes work and evolve to perform new reactions. Towards this goal, we have analysed the *Escherichia coli* metabolome and are relating it to the proteome. Detailed analysis of enzyme active sites and catalytic residues should lead to improved annotation of protein sequences.
- The molecular basis of aging and disease. In collaboration with a consortium of experimental groups, we are analysing expression data from flies, worms and mice to understand the molecular processes involved in aging.

The group includes biologists, chemists, physicists, mathematicians and computer scientists, who work together to solve specific problems.

Future projects and goals

Our goal is to elucidate fundamental principles by analysing many proteins, their sequences, structures and expression profiles. The key questions we seek to address are: can we predict function from structure? Can we understand more about catalytic mechanisms and elucidate the evolution of proteins, pathways and organisms? Can we understand the molecular processes involved in aging?

Selected references

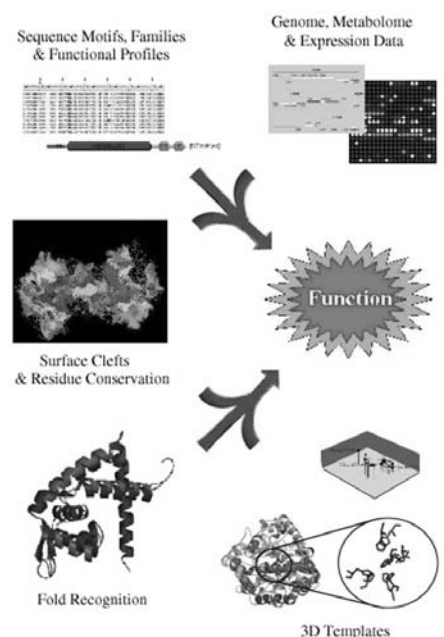
Macchiarulo, A., Nobeli, I. & Thornton, J.M. (2004). Ligand selectivity and competition between enzymes in silico. *Nat. Biotechnol.*, 22, 1039-1045

Ranea, J.A., Buchan, D.W., Thornton, J.M. & Orengo, C.A. (2004). Evolution of protein superfamilies and bacterial genome size. *J. Mol. Biol.*, 336, 871-887

Jones, S. & Thornton, J.M. (2004). Searching for functional sites in protein structures. *Curr. Opin. Chem. Biol.*, 8, 3-7

Porter, C.T., Bartlett, G.J. & Thornton, J.M. (2004). The Catalytic Site Atlas: a resource of catalytic sites and residues identified in enzymes using structural data. *Nucleic Acids Res.*, 32, D129-D133

Shanahan, H.P., Garcia, M.A., Jones, S. & Thornton, J.M. (2004). Identifying DNA-binding proteins using structural motifs and the electrostatic potential. *Nucleic Acids Res.*, 32, 4732-4741



Methods used in the Thornton group to predict function sequence, structure, metabolomic and transcriptomic data.

Rolf Apweiler

PhD 1994, University of
Heidelberg, Germany.
At EMBL since 1987.



Sequence annotation and mining of high-throughput genomics and proteomics datasets

Previous and current research

The activities of the Sequence Database group focus on the production of protein sequence, protein family and nucleotide sequence databases at the EBI. The main research activities are on automatic annotation of proteins, genes and genomes, as well as the mining of high-throughput genomics and proteomics datasets.

With the rapid growth of sequence databases, there is an increasing need for reliable functional characterisation and annotation of newly predicted proteins. One approach is automatic large-scale functional characterisation and annotation, which is generated with limited human interaction.

We use InterPro to recognise domains and to classify all protein sequences in the UniProt Knowledgebase (UniProtKB) into families and superfamilies. For automatic annotation, various systems of standardised transfer of annotation from well-characterised proteins stored in the Swiss-Prot section of the UniProtKB (UniProtKB/Swiss-Prot) to non-annotated entries stored in the TrEMBL section of UniProtKB (UniProtKB/TrEMBL) have been implemented. One system, RuleBase, uses a semi-automatic approach, while Spearmint is completely automated and is based on decision trees. Both systems use UniProtKB/Swiss-Prot as the source to generate the annotation rules, which are stored and managed in RuleBase or Spearmint. InterPro is then used to assign UniProtKB entries into groups. The annotation shared by the functionally characterised UniProtKB/Swiss-Prot proteins of a group is then extracted and is assigned to the non-annotated UniProtKB/TrEMBL entries of this group. These systems have been used to improve the annotation in 32% (RuleBase) and 55% (Spearmint) of UniProtKB/TrEMBL entries.

However, a part of the automatically added data will be erroneous, as are parts of the information coming from other sources. Therefore we introduced a post-processing system called Xanthippe that is based on a simple exclusion mechanism and a decision-tree approach using the C4.5 data-mining algorithm. Xanthippe detects and flags a large proportion of annotation errors and considerably increases the reliability of both automatically generated data and pre-existing annotation inherited from the underlying nucleotide sequence source data.

The highly reliable automatic annotation has already been incorporated into the UniProtKB/TrEMBL flat files, while additional automatic annotation is available from the extended UniProtKB view at www.ebi.uniprot.org/.

Future projects and goals

In addition to major improvements of our current systems described above, we will add mining of high-throughput genomics and proteomics datasets to our automatic annotation toolset. Despite the abundance of data from large-scale experimentation on a genome-wide level, such as expression profiling, protein-protein interaction screens or protein localisation, the systematic and integrated use of this type of information for high-throughput annotation of proteins remains largely unexplored. We therefore intend to build on ongoing research activities at the EBI to develop and assess new protocols to integrate and analyse functional genomics datasets for the purpose of high-throughput annotation of uncharacterised proteins. This will include the analysis of different data types regarding their suitability for the approach, development of data structures that allow the efficient integration and mining of data of different types and quality as well as benchmarking of the obtained results and the application of the new methodologies to annotation of UniProtKB/TrEMBL records.

Selected references

- Orchard, S., Hermjakob, H. & Apweiler, R. (2005). Annotating the human proteome. *Mol. Cell Proteomics*, 4, 435-440
- Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bradley, P., Bork, P., Bucher, P., Cerutti, L., et al. (2005). InterPro, progress and status in 2005. *Nucleic Acids Res.*, 33, D201-D205
- Wieser, D., Kretschmann, E. & Apweiler, R. (2004). Filtering erroneous protein annotation. *Bioinformatics*, 20, Suppl. 1, I342-I347
- Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N. & Yeh, L.S. (2004). UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.*, 32, D115-D119
- Kretschmann, E., Fleischmann, W. & Apweiler, R. (2001). Automatic rule generation for protein annotation with the C4.5 data mining algorithm applied on SWISS-PROT. *Bioinformatics*, 17, 920-926



Ewan Birney

PhD 2000, Sanger Institute,
Hinxton, Cambridge, UK.
Team Leader at EMBL-EBI since 2000.

Selected references

Slater, G.S. & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6, 31

Hubbard, T., Andrews, D., Caccamo, M., Cameron, G., Chen, Y., Clamp, M., Clarke, L., Coates, G., Cox, T., Cunningham, F., et al. (2005). Ensembl 2005. *Nucleic Acids Res.*, 33, D447-D453

Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G.R., Wu, G.R., Matthews, L., Lewis, S., Birney, E. & Stein, L. (2005). Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.*, 33, D428-D432

Hillier, L.W., Miller, W., Birney, E., Warren, W., Hardison, R.C., Ponting, C.P., Bork, P., Burt, D.W., Groenen, M.A., Delany, M.E., et al. (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, 432, 695-716

Algorithms and data mining techniques for genomic data

Previous and current research

Our service group runs both the Ensembl system for metazoan genome annotation and Reactome, a human pathway database.

Ensembl – a joint project between the EBI and the Wellcome Trust Sanger Institute – provides a framework for working with the genomes of higher animals, focusing on vertebrates. It presents, via an interactive website (www.ensembl.org), the human genome together with other genomes that are important for addressing questions in medical research and molecular biology. As of May 2005, there are 16 genomes available in Ensembl; these include the human, several other mammals, the chicken, two species of fish, several insect species and a nematode.

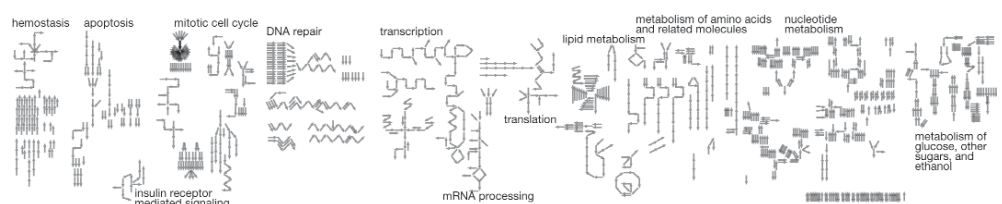
Reactome (www.reactome.org), a curated database of biological processes in humans, is produced in a partnership between the EBI and Cold Spring Harbor Laboratory (CSHL). It is a dual-purpose project that can be used by general biologists as an online textbook of biology, or by bioinformaticians to make discoveries about biological pathways. The basic information in Reactome is provided by bench biologists who are experts on a particular pathway. The information is then managed by groups of curators at CSHL and EBI, peer-reviewed by other researchers and published on the web. Its coverage ranges from the basic processes of metabolism to complex regulatory pathways such as hormonal signalling. While Reactome is targeted at human pathways, it also includes many individual biochemical reactions from non-human systems such as rat, mouse, pufferfish and zebrafish. This makes the database relevant to the many researchers who work on model organisms. All the information in Reactome is backed up by its provenance: either a literature citation or an electronic inference based on sequence similarity. Reactome has grown steadily since its launch in 2004. New pathways include blood clotting, programmed cell death and the start of the insulin receptor pathway.

Research in the Birney group is directed in these broad areas, with a focus on developing novel algorithms and data-mining techniques to discover new biological information, such as finding new cis-regulatory elements or discovering how a pathway has evolved.

Future projects and goals

Future projects include developing new methods in a variety of areas, including:

- sequence manipulation methods in graph structures to represent assembly and efficient search strategies;
- integration of expression and protein-protein interaction data into Reactome;
- examination of the evolution of pathways in metazoans;
- development of novel hidden Markov models for gene prediction.



Reactome's "starry sky" user interface.

Alvis Brazma

PhD 1987, Computer Science, Moscow State University, 1987.
Postdoctoral research at New Mexico
State University, Las Cruces, USA.
At EMBL-EBI since 1997.



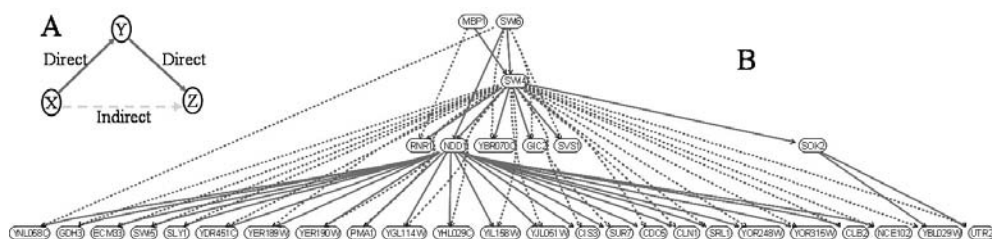
Microarray data informatics, analysis and gene networks

Previous and current research

Our group was among the first to use microarray data to study transcription regulation mechanisms on a genomic scale. In 1999 we realised the importance of standards in microarray data reporting and started working on establishing ArrayExpress – a public repository for microarray data. Now the ArrayExpress repository holds data from over 20,000 microarrays and has become a major bioinformatics resource. More recently, we have extended our work to other domains of high-throughput data, such as proteomics and metabolomics. ArrayExpress development is the major focus of many experienced software engineers and data curators in the group, while our PhD students and postdocs focus mostly on analysing and exploiting these data for systems biology, as well as on developing new algorithms. We have introduced a new approach to modelling gene networks based on four layers of increasing complexity: the network parts list, architecture, logics and dynamics. We are applying new algorithms to extending the parts lists to new regulatory elements, and network architecture for discovering new gene functions. For instance, we have shown that high-throughput data enable us to distinguish between direct and indirect connections in the networks (see figure).

Future projects and goals

We plan to continue developing ArrayExpress, populating it with high-quality data that describe basic biological processes, disease and toxicity, and improving the ease with which users can submit, access and mine the data. We will continue applying cutting-edge software engineering methodologies to achieve these goals. New data-analysis algorithms will be developed and implemented in Expression Profiler, our online data-analysis tool. Integration and meta-analysis of different datasets, with the goal of understanding and modelling basic biological processes, will continue to be the major focus of our research. We have recently started several major collaborative projects, such as integrating transcriptomics, proteomics and metabolomics data to help us understand the molecular mechanisms of type 2 diabetes. A major collaborative project, which includes the Wellcome Trust Sanger Institute, is the study and modelling of cell cycle. We will be using microarray data to reconstruct combinatorial relationships in transcription regulation. A different approach will be dynamic modelling using the recently developed finite state linear model. As part of our involvement in the EMBL Centre for Computational Biology we have started building and analysing a four-dimensional gene expression atlas to integrate gene expression data from different technologies, organisms, tissue types and developmental stages.



A shows how two direct physical interactions from genes X to Y and Y to Z induce an indirect one from X to Z: if gene Z is disrupted, not only gene Y is affected, but also gene X. **B** represents part of a direct and indirect interaction network in yeast obtained from genome-wide transcription factor binding, location and gene-disruption microarray data.

Selected references

- Rustici, G., Mata, J., Kivinen, K., Lio, P., Penkett, C.J., Burns, G., Hayles, J., Brazma, A., Nurse, P. & Bahler, J. (2004). Periodic gene expression program of the fission yeast cell cycle. *Nat. Genet.*, 36, 809-817
- Schlitt, T., Palin, K., Rung, J., Dietmann, S., Lappe, M., Ukkonen, E. & Brazma, A. (2003). From gene networks to gene function. *Genome Res.*, 13, 2568-2576
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., et al. (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.*, 29, 365-371
- Brazma, A., Jonassen, I., Vilo, J. & Ukkonen, E. (1998). Predicting gene regulatory elements in silico on a genomic scale. *Genome Res.*, 8, 1202-1215



Kim Henrick

PhD 1974, University of Western Australia, Perth, Australia.
 Postdoctoral research at the Polytechnic of North London, Imperial College of Science and Technology, CCLRC Daresbury Laboratory, Cambridge Centre for Protein Engineering and National Institute for Medical Research, London.
 Team Leader at EMBL-EBI since 1996.

Macromolecular Structure Database

Previous and current research

Three-dimensional structures give us mechanistic insight into how macromolecules work, and help to explain how their functions are disrupted by mutation or interaction with small molecules. As the results from the structural genomics projects begin to accumulate, the need for better ways of analysing three-dimensional structures, and for integrating structural information with other data types, becomes ever greater. The focus of the Macromolecular Structure Database group (www.ebi.ac.uk/msd) is to accept depositions to the Protein Data Bank, produce an integrated 3D protein relational database and to provide a service to the scientific community.

We have developed an integrated, scalable informatics infrastructure based upon very well understood data warehousing technologies. Data-mining and visualisation tools are being developed alongside the MSD data warehouse to identify non-obvious trends and relationships hidden in the PDB. The MSD, together with the Research Collaboratory for Structural Bioinformatics (RCSB) and the Protein Data Bank Japan (PDBj), are partners in the Worldwide Protein Data Bank (wwPDB, www.wwpdb.org), which oversees the maintenance of the PDB data as a single, publicly accessible data archive. The MSD also runs the Electron Microscopy Database (EMDB) and the web-based deposition system EMDep, for macromolecular volumes determined by electron microscopy.

The group is involved in several database integration projects including Structure Integration with Function, Taxonomy and Sequences (SIFTS), aimed at fully integrating and cross-referencing between MSD, UniProt, Interpro, Gene Ontology (GO) and Enzyme Commission (EC) data. The eFamily project (www.efamily.org.uk) aims to integrate Pfam and Interpro sequence family data with the structure-oriented databases of SCOP and CATH. Other projects include specific NMR-related areas for a toolkit, reference data and extensions to the MSD to hold NMR experimental data, and the RECOORD (www.ebi.ac.uk/msd-srv/docs/NMR/recoord/main.html) database containing recalculated structures for 500+ protein entries from the Protein Data Bank (PDB). We are also working on new methods to visualise and analyse biological sequences, alignments and structures in collaboration with the University of Dundee. Other joint projects are concerned with data management in the high-throughput process of taking a protein target sequence through to a protein sample. We are partners in the UK-funded EHTPX (www.ebi.ac.uk/msd-srv/docs/ehtpx/lims) and PIMS (www.pims-lims.org) projects and the EU-funded FP6 projects BIOXHIT (icarus.embl-hamburg.de:8080/bioxhit/index1.html) and SPINE (www.spineurope.org).

The group has produced a range of search interfaces to the MSD database including general and advanced search interfaces, residue, active-site, and target search interfaces and a secondary-structure-matching service.

Future projects and goals

The MSD group will continue to enhance its services to meet the demands of the structural biology community. Services currently in development include search interfaces for small 3D motifs in proteins, common local residue interactions and protein-protein interfaces, as well as a second-generation structure/sequence viewer. Other research areas include extensions to the MSD database and web services for analysis of protein cavities and interfaces; water clusters in protein structures and ligand interactions; structure-sequence relationships; chemoinformatics analysis of the PDB; integrating information on alternate splice forms and protein interactions with structures from the PDB; and new visualisation techniques.

Selected references

Pajon, A., Ionides, J., Diprose, J., Fillon, J., Fogh, R., Ashton, A.W., Berman, H., Boucher, W., Cygler, M., Deleury, E., et al. (2005). Design of a data model for developing laboratory information management and analysis systems for protein production. *Proteins*, 58, 278-284

Golovin, A., Dimitropoulos, D., Oldfield, T., Rachedi, A. & Henrick, K. (2005). MSDsite: a database search and retrieval system for the analysis and viewing of bound ligands and active sites. *Proteins*, 58, 190-199

Velankar, S., McNeil, P., Mittard-Runte, V., Suarez, A., Barrell, D., Apweiler, R. & Henrick, K. (2005). E-MSD: an integrated data resource for bio-informatics. *Nucleic Acids Res.*, 33, D262-D265

Krissinel, E. & Henrick, K. (2004). Common subgraph isomorphism detection by backtracking search. *Software – Practics and Experience*, 34, 591-607

Peter Rice

BSc 1976, University of Liverpool, UK.
Previously at EMBL-Heidelberg, Sanger
Institute and LION Bioscience.
Group Leader at EMBL-EBI since 2003.



e-Science, Grid and integrated bioinformatics development

Previous and current research

The group's focus is on the integration of bioinformatics tools and data resources. We have the remit to investigate and advise on the e-Science and Grid technology requirements of the EMBL-EBI, through application development, training exercises and participation in international projects and standards development. Our group is responsible for the EMBOSS open source sequence analysis package, the Taverna bioinformatics workflow system (originally developed as part of the myGrid UK eScience project) and for various projects (including EMBRACE and ComparaGrid) that integrate access to bioinformatics tools and data content.

Future projects and goals

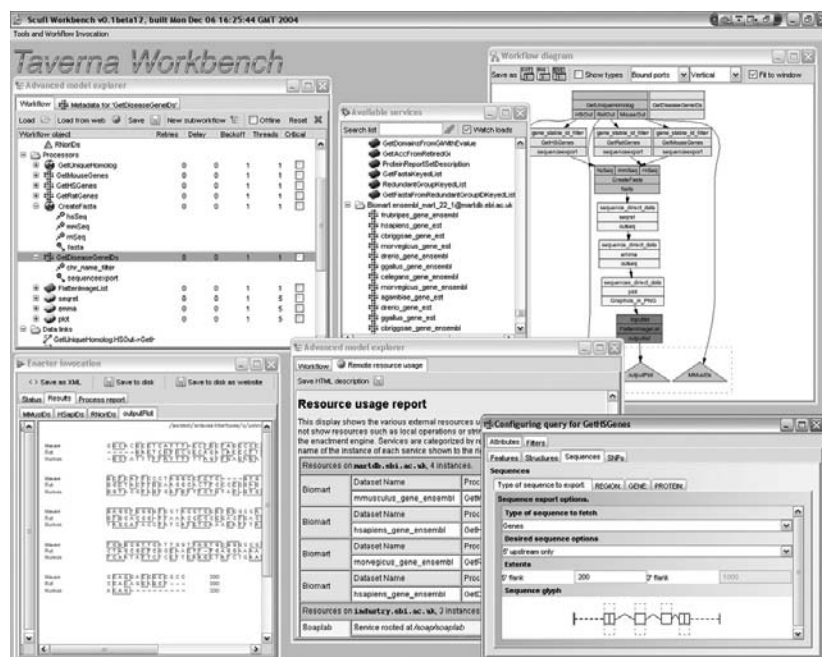
The EMBRACE project is a five-year EC-funded Network of Excellence in which the partners will provide a common interface to a broad range of bioinformatics tools and data content. This will build on our EMBOSS, SoapLab and Taverna projects and lead to interoperable data and tools with Grid-enabled access in collaboration with other European Grid projects.

Selected references

Oinn, T., Addis, M., Ferris, J., Marvin, D., Senger, M., Greenwood, M., Carver, T., Glover, K., Pocock, M.R., Wipat, A. & Li, P. (2004). Taverna: a tool for the composition and enactment of bioinformatics workflows. *Bioinformatics*, 20, 3045-3054

Senger, M., Rice, R., & Oinn, T. (2003). Soaplab – a unified Sesame door to analysis tools. *Proceedings UK OST e-Science 2nd All Hands Meeting 2003, Nottingham, UK 2-4 Sept, 2003.*

Rice, P., Longden, I. & Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.*, 16, 276-277



A Taverna session showing a workflow that integrates data from BioMart with SoapLab's EMBOSS services.



Nick Goldman

PhD 1992, Cambridge University, UK.

Postdoctoral research at National Institute for Medical Research, London, and Cambridge University, UK. Wellcome Trust Senior Fellow since 1995.

Group Leader at EMBL-EBI since 2002.

Evolutionary analysis of sequence data

Previous and current research

The group's main interests are in developing better methods for analysing DNA and amino acid sequences to study evolution, and in using information from evolutionary studies to improve other data analyses in bioinformatics. Our work includes the theoretical basis of phylogenetic analyses of sequence data, aiming to understand and improve methods, and developing statistical methods that test the accuracy of our mathematical models. We also concentrate on devising better mathematical models to describe sequence evolution. Typically this involves incorporating greater mathematical complexity in the hope of adding biological reality. By seeing which biological features add most to the accuracy of the models, we hope to find out which proposed evolutionary forces are responsible for the patterns of sequence variation within and among species.

Increased understanding of the processes of molecular evolution has come from models that acknowledge the constraints placed on evolution by protein secondary structure, models that relax the restriction that DNA substitutions occur only as point mutations (i.e. affecting single sites) and models that can represent the forces of natural selection acting on nucleic acid sequences' protein products. In the last example, it is now possible to use evolutionary inferences to detect specific residues in proteins that seem to be evolving under strong positive selection. These mathematical models have been devised, coded and applied in our group and in collaboration with scientists worldwide.

The methods we develop are almost invariably based on probabilistic modelling. Analytical (closed form) solutions of the modelling equations are rarely available, and much of our work is based on computationally intensive iterative algorithms, with statistical analyses performed via repeated data simulation and analysis.

Our group also maintains and develops the Pandit database of DNA and protein sequence alignments and phylogenies (www.ebi.ac.uk/goldman-srv/pandit), a resource for both evolutionary modellers and data-miners.

Future projects and goals

The increasing availability of genomic data is having an effect on all of our work. These data create new possibilities for extending our understanding of molecular evolution. Our data analysis work is moving from the small-scale evolutionary analyses of the past towards large-scale studies involving multiple gene sequences from many organisms. Work on statistical methods and mathematical modelling of evolution can now be validated using vastly more data than was the case previously. Projects are currently underway to scan large sequence databases for previously undetected functional elements and for examples of positive selection, to analyse the evolution of multiple mammalian genomes, and implement complex models of DNA and protein sequence evolution. We are also developing multiple sequence alignment methods that use both phylogenetic information and our existing understanding of the evolutionary processes that operate across large genomic regions.

Selected references

Massingham, T. & Goldman, N. (2005). Detecting amino acid sites under positive selection and purifying selection. *Genetics*, 169, 1753-1762

Kosiol, C. & Goldman, N. (2005). Different versions of the Dayhoff rate matrix. *Mol. Biol. Evol.*, 22, 193-199

Wong, W.S., Yang, Z., Goldman, N. & Nielsen, R. (2004). Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics*, 168, 1041-1051

Whelan, S. & Goldman, N. (2004). Estimating the frequency of events that cause multiple-nucleotide changes. *Genetics*, 167, 2027-2043

Henning Hermjakob

Dipl. Inf. (MSc) 1996, Bioinformatics, University of Bielefeld, Germany.
Research assistant in the Transfac Database team at the National Research Centre
for Biotechnology (GBF), Braunschweig, Germany.
At EMBL-EBI since 1997.



Proteomics services – standards, databases and tools for proteomics

Previous and current research

Publicly available proteomics data are currently highly fragmented across public databases, authors' websites, PDF tables in the supplementary material of publications and the text of publications. This is due partly to the highly dynamic and diverse development of experimental technologies, and partly to the lack of standards and repositories for proteomics.

The Proteomics Services Team provides databases and tools for the deposition, distribution and analysis of proteomics and proteomics-related data. We contribute to the development of community standards for proteomics data in the context of the HUPO Proteomics Standards Initiative (PSI), and develop reference implementations for these standards.

The HUPO PSI Molecular Interaction standard (PSI MI) has been jointly developed by major interaction-data providers from both the academic and commercial sector, among them BIND, Cellzome, DIP, Hybrigenics, IntAct, MINT and MIPS, and is now widely accepted as the standard format for exchange of protein-interaction data. The IntAct (www.ebi.ac.uk/intact) project provides an open source, production-quality, fully portable molecular interaction database and toolset implementing the PSI MI standard. Beyond IntAct, we are now collaborating with other major providers to establish a regular exchange of molecular interaction data between large public data providers.

The PSI mass spectrometry work group (PSI MS) has developed the mzData format, a vendor-independent representation of mass spectra, providing a unified format for data archiving, exchange and search engine input. It has been jointly developed by academic users, commercial users and instrument vendors. The PRIDE (www.ebi.ac.uk/pride) database, developed in collaboration with the University of Gent, provides a PSI MS-compatible database for protein identifications.

In addition to the MI and MS work groups, HUPO PSI has started to develop a standard for the representation of a full proteomics experiment, the Global Proteomics Standards (GPS). Based on the PEDRo schema, this work group develops the "Minimum Information About a Proteomics Experiment (MIAPE)" document analogue to the MIAME requirements for a microarray experiment, and an object model (PSI-OM) and XML format (PSI-ML) to fully represent a proteomics experiment.

In the context of the EU-funded BioSapiens (www.biosapiens.info) and Transfop projects, the Proteomics Services Team contributes to the development of tools for proteomics data integration and distribution based on the lightweight DAS (Distributed Annotation System) protocol, in particular the DAS UniProt reference server (www.ebi.ac.uk/uniprot-das) and Dasty client (www.ebi.ac.uk/dasty).

In the framework of the IntEnz integrated Enzyme database (www.ebi.ac.uk/IntEnz), the ChEBI small molecules database (www.ebi.ac.uk/chebi), and the IUPHAR receptor database (www.iuphar-db.org), the Proteomics Services Team contributes to the standardisation of proteomics-related nomenclature and data representation.

Selected references

- Jones, P., Vinod, N., Down, T., Hackmann, A., Kahari, A., Kretschmann, E., Quinn, A., Wieser, D., Hermjakob, H. & Apweiler, R. (2005). Dasty and UniProt DAS: a perfect pair for protein feature visualisation. *Bioinformatics*, 21, 3198-3199
- Martens, L., et al., (2005). PRIDE: The PRoteomics IDentifications database. *Proteomics*, in press
- Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Ceol, A., Moore, S., Orchard, S., Sarkans, U., von Mering, C., et al., (2004). The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. *Nat. Biotechnol.*, 22, 177-183
- Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., Vingron, M., Roehert, B., Roepstorff, P., Valencia, A., Margalit, H., Armstrong, J., Bairoch, A., Cesareni, G., Sherman, D. & Apweiler, R. (2004). IntAct: an open source molecular interaction database. *Nucleic Acids Res.*, 32, D452-D455



Wolfgang Huber

PhD 1998, University of Freiburg, Germany.

Postdoctoral Research at IBM Almaden Research Center, San Jose, California and German Cancer Research Center (DKFZ), Heidelberg, Germany.

Group Leader at EMBL-EBI since 2004. Joint appointment with Gene Expression Unit.

Computational statistics in functional genomics

Previous and current research

The group's interest is in developing advanced mathematical and statistical methods for the understanding of functional genomics data. This includes:

- analysis methods for emerging microarray technologies;
- statistical methods for the integration of functional genomic data sets;
- analysis of high content phenotype data from systematic functional screens and of genetic interaction networks;
- statistics of graph-like data and graphical models.

In addition, we regard the publication of high-quality scientific software as complementary to the publication of methodical insight and a core contributor to the Bioconductor project (www.bioconductor.org).

Future projects and goals

Biology and its applications to human health will continue to be driven by advances in experimental technologies. Of particular interest to us are array-based techniques (DNA, mRNA, protein), genome-scale functional cellular assays, and high-content phenotyping. To make these fruitful for systematic models of biological processes, we aim to stay at the forefront of method development in experimental design, data analysis, statistical software and mathematical modelling. An emphasis is on project-oriented collaborations with experimenters.

Selected references

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A. & Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, 21, 3439-3440

Gentleman, R., Carey, V., Huber, W., Irizarry, R. & Dudoit, S. (2005). *Bioinformatics and computational biology solutions using R and Bioconductor*. Springer Verlag

Arlt, D., Huber, W., et al. (2005). Functional profiling: from microarrays via cell-based assays to novel tumor relevant modulators of the cell cycle. *Cancer Res.*, in press

Huber, W., et al. (2003). Parameter estimation for the calibration and variance stabilization of microarray data. *Statistical Applications in Genetics and Molecular Biology*, 2, 3

Nicolas Le Novère

PhD 1998, Pasteur Institute, Paris, France.
Postdoctoral research at Cambridge University, UK
and CNRS, Paris, France.
Group Leader at EMBL-EBI since 2003.



Applying computational systems biology to neuronal signalling

Previous and current research

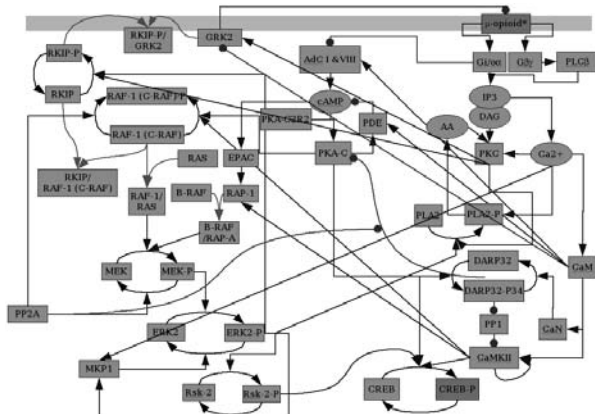
The group's interests revolve around signal transduction in neurons, ranging from the molecular structure of membrane proteins involved in neurotransmission to modelling signalling pathways. A strong focus is the molecular and cellular basis of neuroadaptation in neurons of the basal ganglia. We also provide services that facilitate our research, including database production and software development.

The dynamic behaviour of the post-synaptic machinery probably strongly influences signal transduction. By building detailed and realistic models of the post-synaptic machinery, we try to decipher how neurotransmitter-receptor movement and clustering in the excitatory synapse influence synaptic signalling. Downstream from the transduction machinery, we build quantitative models of the integration of signalling pathways known to mediate the effects of neurotransmitters, neuromodulators and drugs of abuse. This modelling requires various types of software, including design environments and simulators. We are involved in the development of the Systems Biology Markup Language (SBML) and its software support. Moving from the form to the content, we are also developing standards for model curation and annotation, and controlled vocabularies to increase the information content of models. In parallel, we are developing the BioModels Database, a data resource that allows biologists to store, search and retrieve published mathematical models of biological interest.

A significant proportion of the proteins involved in neuronal signalling are embedded in biological membranes. Their three-dimensional structures are seldom known and this precludes a clear understanding of their function. Our group maintains the computing infrastructure of the consortium E-MeP, a research platform that focuses on solving high-resolution structures of membrane proteins and membrane protein complexes. Most of the project's 300 eukaryotic targets are involved in neurotransmission, many of them implicated in devastating pathologies.

Future projects and goals

In forthcoming years, the activity of the group will unfold in two directions. Our work on modelling neuronal signalling will expand to tackle problems such as the role of scaffolding proteins or the synchronisation of calcium waves and phosphorylation gradients. Taking advantage of the growth of BioModels, we will also carry out research on model composition, with the aim of improving component identification and reaction matching to build large-scale models of cellular compartments such as dendritic spines.



Dynamical model of how the post-synaptic density is organised. The AMPA glutamate receptor, stargazin, PSD95, cytoskeletal proteins and adaptor proteins are represented. The formula describes the probability function used to control the linear movement of receptors in the membrane.

Selected references

- Le Novère, N. & Donizelli, M. (2004). The Molecular Pages of the mesotelencephalic dopamine consortium (DopaNet). *BMC Bioinformatics*, 5, 174
- Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19, 524-531
- Le Novère, N., Grutter, T. & Changeux, J.P. (2002). Models of the extracellular domain of the nicotinic receptors and of agonist- and Ca^{2+} -binding sites. *Proc. Natl. Acad. Sci. USA*, 99, 3210-3215
- Le Novère, N. & Shimizu, T.S. (2001). STOCHSIM: modelling of stochastic biomolecular processes. *Bioinformatics*, 17, 575-576



Nicholas Luscombe

PhD 2000, University College London, UK.

Postdoctoral research at the Department of Molecular Biophysics & Biochemistry, Yale University, USA.
Group Leader at EMBL-EBI since 2005.

Joint appointment with Gene Expression Unit.

Genome-scale analysis of regulatory systems

Previous and current research

An organism's viability depends on the right cells making the right proteins at the right times. Much of this regulation is achieved at the level of transcription, and a breakdown in this system results in numerous diseases. Thus it is important to characterise regulatory systems – and problems arising from associated defects – at a molecular level. We apply computational techniques to analyse transcriptional regulatory systems on a genomic scale.

The continued flood of biological data means that many interesting questions require the application of computational methods to answer them. The strength of bioinformatics is its ability to uncover general principles providing global descriptions of entire systems. Armed with these biological data we are now in a great position to do this.

Much of our work has focused on the regulatory system in yeast. By integrating diverse information – from its genome sequence to the results of functional genomics experiments – we can study the regulatory system at a whole-organism level. The ultimate goal is to analyse the human genome, but yeast's core similarities with humans and relative simplicity make it an ideal starting point. Nevertheless, the yeast system involves thousands of regulatory interactions, and we need to use a network perspective to examine it (figure 1).

Some of our projects have included:

- experiments to establish the regulatory binding sites of transcription factors in yeast;
- studying the structure and robustness of the regulatory network to identify specific points of vulnerability that have potential implications for diseases;
- integrating data to visualise how the system controls gene expression;
- analysing how the regulatory network is used dynamically in response to diverse biological demands (figures 1 and 2).

Future projects and goals

We will develop new techniques to advance our understanding of regulatory systems, and expand our approaches towards more complex organisms. Finally, we will interact closely with research groups performing functional genomics experiments.

Selected references

Luscombe, N.M., Babu, M.M., Yu, H., Snyder, M., Teichmann, S.A. & Gerstein, M. (2004). Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, 431, 308-312

Babu, M.M., Luscombe, N.M., Aravind, L., Gerstein, M. & Teichmann, S.A. (2004). Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.*, 14, 283-291

Yu, H., Luscombe, N.M., Qian, J. & Gerstein, M. (2003). Genomic analysis of gene expression relationships in transcriptional regulatory networks. *Trends Genet.*, 19, 422-427

Horak, C.E., Luscombe, N.M., Qian, J., Bertone, P., Piccirillo, S., Gerstein, M. & Snyder, M. (2002). Complex transcriptional regulatory circuitry at the G1/S transition in *Saccharomyces cerevisiae*. *Genes Dev.*, 16, 3017-3033

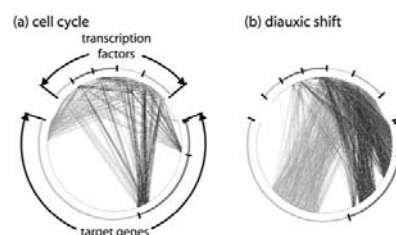


Figure 1. A network representation displays the dynamic usage of the yeast regulatory system during (a) the cell cycle and (b) diauxic shift. Distinct regions of the network are clearly used, and this is accompanied by fundamental changes in network structure.

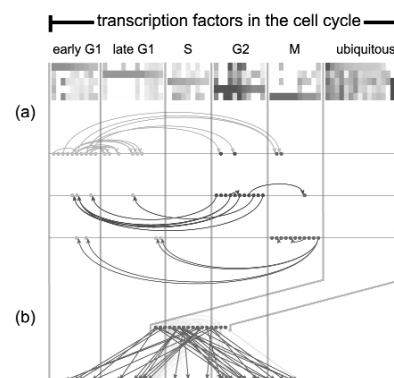


Figure 2. There is substantial inter-regulation between the transcription factors that control the cell cycle. (a) In serial inter-regulation, factors active in one phase of the cycle control the expression of those that are active in subsequent ones. (b) In parallel inter-regulation, ubiquitously active factors control the phase-specific ones.



Dietrich Rebholz-Schuhmann

Master in Medicine 1988, and PhD 1989, immunology, University of Düsseldorf, Germany.
Master in Computer Science 1993, Passau, Germany.
Postdoctoral research at gsf, Munich, and LION bioscience AG, Heidelberg, Germany.
Group Leader at EMBL-EBI since 2003.

Extracting facts from the scientific literature in molecular biology (text mining)

Previous and current research

Mining the biomedical literature, together with open access publishing, will lead to a faster turnover of scientific hypotheses. Shortly after a publication becomes available, automatic information extraction techniques seek the document and its facts, extract the data and put it into electronic databases for ready access. These data then inspire new scientific work. To realise this vision, we have to solve a number of problems. First we have to identify a large number of biomedical terms, many of which have context-dependent meanings. These terms must be disambiguated using contextual information. Special nomenclatures for the description of mutations, chemical compounds and their experimental results must be recognised. Finally, the semantic relationships between identified concepts must be extracted.

The Rebholz group uses text-mining methods to assess information provided by the scientific literature in the context of the EBI's databases and to make new information-extraction solutions available to the public and to curation teams. We apply pattern matching (finite state automata), shallow parsing and support vector machines (SVMs) for disambiguation of terms. One new solution developed by the group is the automatic generation of a dictionary of abbreviations from Medline: each abbreviation has a list of expanded forms grouped together according to their linguistic similarity to provide all the meanings of an abbreviation. SVMs automatically identify from the context of an abbreviation its correct meaning with >98% precision.

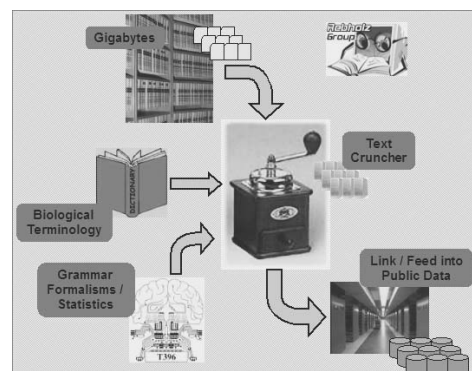
We work together with curation teams to mutual benefit, but every new solution has the potential to provide a service to the public. Whatizit is a modular text-mining tool that analyses submitted text and identifies biomedical facts. Apart from extracting a large number of biological terms, it extracts protein-protein interactions and other types of information. The concepts that it returns are linked to relevant database entries.

EbiMed is a new type of information extraction tool. It retrieves several thousand abstracts from the EBI's in-house Medline installation, analyses them and then presents a summary of biomedical associations. The Medline query defines the domain and scope of the analysed documents, whereas the analysis is predetermined and tuned to molecular biology.

Future projects and goals

We need to develop more solutions to extract facts correctly, including selected relationships between concepts. Such components can be based on text-mining techniques such as co-occurrence, full parsing, hidden Markov models or support vector machines. Topics of particular interest include: mapping terms from text to ontological resources; identifying related facts; disambiguating conflicting findings in natural-language text; approaches for the automatic identification of patterns; and definition of estimates for the quality of the extracted data.

The core components of Whatizit. Biological terminology are gathered from electronic databases; grammar formalisms are kept in a modular library, statistical approaches are used to classify and disambiguate findings, and all components are part of an XML-based pipeline that is distributed on a cluster of Linux computers. Input can be any type of Unicode, PDF or HTML text and the output is either accessible through a web interface or can be integrated into special biological databases after a consistency check.



Selected references

Rebholz-Schuhmann, D., Kirsch, H. & Couto, F. (2005). Facts from text – is text mining ready to deliver? PLoS Biol., 3, e65

Kirsch, H. & Rebholz-Schuhmann, D. (2005). Distributed modules for text annotation and IE applied to the biomedical domain. J. Biomed. Informat., in press

Rebholz-Schuhmann, D., Marcel, S., Albert, S., Tolle, R., Casari, G. & Kirsch, H. (2004). Automatic extraction of mutations from Medline and cross-validation with OMIM. Nucleic Acids Res., 32, 135-142

Albert, S., Gaudan, S., Knigge, H., Raetsch, A., Delgado, A., Huhse, B., Kirsch, H., Albers, M., Rebholz-Schuhmann, D. & Koegl, M. (2003). Computer-assisted generation of a protein-interaction database for nuclear receptors. Mol. Cell. Endocrinol., 17, 1555-1567

EMBL-Monterotondo, Italy

The EMBL Monterotondo campus is situated in a green park on the banks of the Tiber river, 20 km north of central Rome. The Eternal City is an easy train ride away, and the nearby Lazio countryside features medieval hill towns, spectacular mountain terrain and lakes for hiking, biking, skiing and swimming. EMBL shares the Monterotondo campus with Italian national research groups (IBC-CNR) and the headquarters of the European Mouse Mutant Archive (EMMA).

The EMBL Mouse Biology Unit is centred in Monterotondo but includes groups at other EMBL sites. The Unit was established to capture new opportunities in mouse functional genomics, and to exploit applications of mouse genetic manipulation to biomedical problems. Alliances with other European academic research and clinical centres have established EMBL-Monterotondo as a hub for the international mouse research community. Participation of the Outstation in several EU-wide mouse research and information initiatives serves to link genetics/genomics, phenotyping, physiology and translational features.

Recent expansion of the Outstation has retained the original emphasis on developmental mechanisms and extended capabilities to cover adult mammalian physiology and disease. Naturally converging interests between the groups include new collaborative explorations in developmental neurobiology, genetic and pharmacologic manipulations of adult behaviour, inflammation and regenerative processes, stem cell biology and biomedical applications.

A new, state-of-the-art animal facility provides a full range of mouse transgenic and gene knockout production, rederivation and cryopreservation services, and a fully-equipped behaviour phenotyping suite. The continued refinement of gene regulation in vivo is focused on generating more accurate models of human pathologies and multigenic disorders through the development of conditional and inducible mouse mutations. Other centralised core facilities provide a wide array of technologies to the scientific community. Interaction with clinical groups through the newly formed EMBL Centre for Disease Mechanisms provides exciting opportunities for application of basic research to advanced disease diagnosis and treatment.

Development of an EMBL Monterotondo course curriculum, focusing on mouse genetic manipulation in collaboration with local faculty at CNR, EMMA and The Jackson Laboratory, reflects EMBL's longstanding tradition of transmitting new information and expertise. A dynamic seminar series and a visiting researcher programme, together with active collaborations with research groups throughout the world, integrates the science at EMBL-Monterotondo with the international mouse biological community.





Nadia Rosenthal

PhD 1981, Harvard Medical School, USA.
 Postdoctoral research at the National Cancer Institute, USA.
 Assistant Professor, Boston University Medical Center, USA.
 Associate Professor, Massachusetts General Hospital, USA.
 Group Leader and Unit Coordinator at EMBL-Monterotondo since 2001.

Muscle development and disease

Previous and current research

Current advances in mouse genetics provide exciting new opportunities to study transcriptional regulatory pathways in the vertebrate embryo. By analysing the cis- and trans-regulatory components of skeletal muscle fibre-restricted gene expression using transgenic and knockout mouse models, we uncovered an epigenetic mechanism for perpetuating positional cues in developing skeletal muscles, and an essential function for E protein family of transcription factors in fibre type specification. We have recently extended our interest in muscle development to explore the anteroposterior specification of cell fate in the embryonic heart. Retinoic acid (RA) plays an important role in heart anteroposterior patterning and growth, and we have begun to characterise the transcriptional regulation of chamber-specific gene expression in cardiac development. A second phase of RA activity in the developing heart is initiated in the proepicardial organ and persists in migratory epicardial cells that completely envelop and invade the heart to form the coronary vasculature, suggesting a role for RA in expansion and morphogenesis of underlying myocardial tissues.

We are also interested in the molecular pathology of adult muscle. Selective muscle fibre loss in aging and diseased skeletal muscle can be blocked by transgenic or viral delivery of insulin-like growth factors, which promote recruitment of stem cells to sites of injury. A novel calcineurin-mediated signalling pathway responsible for the hypertrophic action of IGF-1 on skeletal muscle led to an unexpected role for GATA transcription factors in skeletal muscle hypertrophic gene expression. We extended these studies to cardiac muscle hypertrophy and the role of specific calcineurin isoforms in human heart failure. Mechanisms of IGF-1 action in both skeletal and cardiac muscle suggest novel therapeutic approaches to the prevention of muscle atrophy and degeneration, through enhancement of stem cell-mediated regeneration.

Future projects and goals

Using gene mutational strategies in the mouse we will continue to investigate:

- the integration of anteroposterior decisions in the developing heart with mechanisms that confer positional information in the embryo;
- the role of RA in cardiac muscle growth and coronary vessel formation;
- the signal transduction pathways responsible for regulating anabolic and hypertrophic gene expression programs in skeletal and cardiac muscle tissues;
- the origin and function of stem cells in muscle regeneration.

Our goals are to understand the molecular genetics of pattern formation in embryonic development, to explore the nature of regeneration in the adult vertebrate body and to develop clinically relevant strategies for therapeutic intervention in neuromuscular and cardiac aging and disease.

Selected references

Barton, E.R., Morris, L., Musaro, A., Rosenthal, N. & Sweeney, H.L. (2002). Muscle-specific expression of insulin-like growth factor I counters muscle decline in mdx mice. *J. Cell Biol.*, 157, 137-148

Paul, A.C. & Rosenthal, N. (2002). Different modes of hypertrophy in skeletal muscle fibres. *J. Cell Biol.*, 156, 751-760

Grounds, M.D., White, J.D., Rosenthal, N. & Bogoyevitch, M.A. (2002). The role of stem cells in skeletal and cardiac muscle repair. *J. Histochem. Cytochem.*, 50, 589-610

Xavier-Neto, J., Rosenthal, N., Silva, F.A., Matos, T.G., Hochgreb, T. & Linhares, V.L. (2001). Retinoid signalling and cardiac anteroposterior segmentation. *Genesis*, 31, 97-104

Musaro, A., McCullagh, K., Paul, A., Houghton, L., Dobrowolny, G., Molinaro, M., Barton, E.R., Sweeney, H.L. & Rosenthal, N. (2001). Localised Igf-1 transgene expression sustains hypertrophy and regeneration in senescent skeletal muscle. *Nat. Genet.*, 27, 195-200

Cornelius Gross

PhD 1995, Yale University, USA.
Postdoctoral research at Columbia University, USA.
Group Leader at EMBL-Monterotondo since 2003.



Regulation of anxiety behaviour

Previous and current research

Anxiety disorders are debilitating mental illnesses characterised by excessive or inappropriate fear. Epidemiological studies of monozygotic and dizygotic twins suggest that both genetic and environmental factors contribute to the prevalence of these disorders. However, we know very little about the neural circuits that underlie susceptibility to anxiety. We do not know which brain structures are involved nor what sort of structural or molecular changes might cause a predisposition to anxiety. In order to better understand the neural substrates of anxiety, we are undertaking studies to define the molecular and anatomical defects that lead to altered anxiety behaviour in the mouse.

Recently, we have used a conditional and tissue-specific genetic strategy to identify brain circuits that mediate the anxiety modulating effects of serotonin. We found that the forebrain contains a critical anxiety circuit that is modulated by serotonin and that functions during development to establish normal anxiety behaviour in the adult animal. This result argues that critical plastic changes occurring in the forebrain during the first few weeks of life are able to determine life-long emotional behaviour. Importantly, environmental influences such as changes in maternal care can also play a critical role during this developmental period in shaping anxiety behaviour.

Future projects and goals

The major focus of future research will be aimed at understanding how genetic and environmental influences act during development to modulate anxiety behaviour in the mouse.

This research will involve a number of approaches, including:

- use of genetic, pharmacological and anatomical tools to determine the critical brain structures and time period as well as identify relevant morphological and molecular changes via which serotonin modulates anxiety behaviour;
- examination of the interaction between genetic and environmental influences on anxiety during the postnatal period;
- identification of novel genetic factors that can impact anxiety behaviour during development;
- creation of mouse models of specific human genetic variations that have been associated with anxiety by genetic linkage studies.

Together these approaches are aimed at discovering specific long-term plastic mechanisms that underlie susceptibility to anxiety. A better understanding of the molecular signals that trigger these plastic changes will allow us to form specific hypotheses about how human anxiety is determined and may lead to improved diagnostic and therapeutic tools in the clinic.

Selected references

- Klemenhagen, K.C., Gordon, J.A., David, D.J., Hen, R. & Gross, C. (2005). Increased fear response to contextual cues in mice lacking the 5-HT1A receptor. *Neuropsychopharmacology*, in press
- Gross, C. & Hen, R. (2004). The developmental origins of anxiety. *Nat. Rev. Neurosci.*, 5, 545-552
- Santarelli, L., Saxe, M., Gross, C., Surget, A., Battaglia, F., Dulawa, S., Weisstaub, N., Lee, J., Duman, R., Arancio, O., Belzung, C. & Hen, R. (2003). Requirement of hippocampal neurogenesis for the behavioural effects of antidepressants. *Science*, 301, 805-809
- Gross, C., Zhuang, X., Stark, K., Ramboz, S., Oosting, R., Kirby, L., Santarelli, L., Beck, S. & Hen, R. (2002). Serotonin1A receptor acts during development to establish normal anxiety-like behaviour in the adult. *Nature*, 416, 396-400



Liliana Minichiello

PhD 1993, University of Naples, Italy, and the National Cancer Institute, NIH, Bethesda, USA.

Postdoctoral research at EMBL-Heidelberg.

Team Leader at EMBL-Monterotondo since 2000. Joint appointment with the Developmental Biology Unit.

Signalling mechanisms and gene regulation in the nervous system

Previous and current research

Nerve cells are the elementary signalling units of the nervous system. The structural and functional integrity of the nervous system is controlled by various signals that are in part transduced by transmembrane receptor tyrosine kinases. Extensive studies have established that neurotrophin receptor tyrosine kinases (called Trk) specifically and potently regulate diverse neuronal responses during development. Work with mice carrying deletions in one or two different Trk receptors have revealed their fundamental role in promoting survival and stabilizing the phenotype of neurons of both the peripheral (PNS) and the central nervous systems (CNS).

The main research focus of my laboratory is studying mechanisms of action of Trk receptors in the mouse nervous system. Our primary aim is to correlate the biological functions of neurotrophins with signal transduction events. For example, to investigate the mechanism of TrkB action in hippocampal LTP (Long Term Potentiation) function, we have chosen an approach that combines sophisticated genetic tools allowing to interfere with single phosphorylation sites on a large receptor protein (by generating mice with a targeted mutation in the PLC γ docking site of TrkB or in the Shc site of TrkB receptor) with biochemical and electrophysiological approaches. This provided us with a relatively clean way to tease out specific roles of the various signalling players involved in TrkB-dependent synaptic plasticity (see figure). To further dissect the mechanisms by which TrkB regulates the diverse neuronal responses we are looking now at downstream effectors induced by TrkB activation using the microarray approach and generating mouse genetic models of particular identified genes.

Future projects and goals

For further comprehensive understanding of how signalling molecules interlink with each other in the formation of a transduosome downstream TrkB receptor we are employing the co-precipitation/mass spectrometry approach to identify signalling complexes (in particular we are using the TAP tag strategy, in which a tandem affinity purification tag will be inserted into the mouse gene of interest by homologous recombination in ES cells. This will allow protein complexes to be directly purified from mouse tissues and subjected to mass spectrometric analysis for the identification of in vivo associated proteins.

Current areas of interest include also the generation of mouse models for neurodegenerative diseases like Alzheimer's and Huntington's by using conditional mutagenesis. Our goal is to define the molecular mechanisms by which neurotrophin receptor tyrosine kinases perform their various functions in the developing and adult vertebrate nervous system.

Selected references

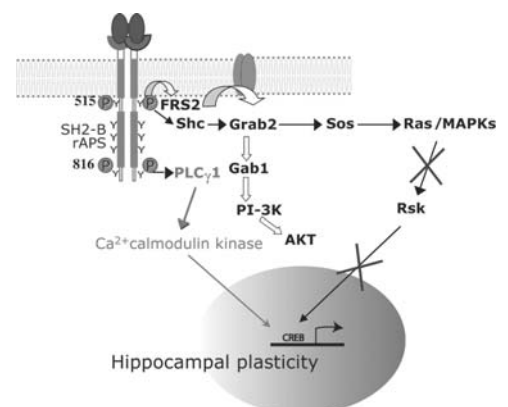
Medina, D.L., Sciarretta, C., Calella, A.M., Von Bohlen Und Halbach, O., Unsicker, K. & Minichiello, L. (2004). TrkB regulates neocortex formation through the Shc/PLC γ -mediated control of neuronal migration. *EMBO J.*, 23, 3803-3814

Rivera, C., Voipio, J., Thomas-Crusells, J., Li, H., Emri, Z., Sipila, S., Payne, J.A., Minichiello, L., Saarma, M. & Kaila, K. (2004). Mechanism of activity-dependent downregulation of the neuron-specific K-Cl cotransporter KCC2. *J. Neurosci.*, 24, 4683-4691

Minichiello, L., Calella, A.M., Medina, D.L., Bonhoeffer, T., Klein, R. & Korte, M. (2002). Mechanism of TrkB-mediated hippocampal long-term potentiation. *Neuron*, 36, 121-137

Postigo, A., Calella, A.M., Fritsch, B., Knipper, M., Katz, D., Eilers, A., Schimmang, T., Lewin, G.R., Klein, R. & Minichiello, L. (2002). Distinct requirements for TrkB and TrkC signalling in target innervation by sensory neurons. *Genes Dev.*, 16, 633-645

The PLC γ site, and subsequent phosphorylation of calcium calmodulin kinase/s and CREB is responsible for the hippocampal synaptic plasticity downstream of TrkB, whereas the Ras/MAPK pathway activated by TrkB is not critical for hippocampal LTP.



Claus Nerlov

PhD 1995, the University of Copenhagen, Denmark.
 Postdoctoral research at EMBL-Heidelberg.
 Head of Laboratory of Gene Therapy Research, Copenhagen University Hospital, Denmark.
 Group Leader at EMBL since 2001.



Transcriptional regulation in cellular differentiation and disease

Previous and current research

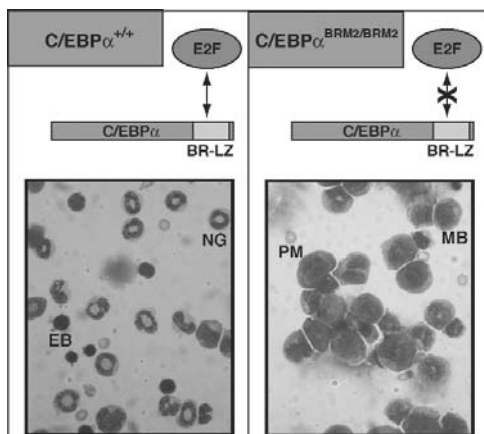
The goal of our laboratory is to understand how physiological processes, such as cell proliferation, differentiation, metabolism and cancer, are regulated by transcription factors. By using a combination of conditional knockout and knockin strategies in the mouse, as well as in vivo proteomics, we are analysing the role of C/EBP transcription factors and interacting regulatory proteins in differentiation and function of several cell types and tissues, including of adipocytes, hematopoietic cells, skin and liver.

We have previously shown that the one of these transcription factors, C/EBP α , co-ordinates cell differentiation with cell growth arrest by interfering with the E2F complex of cell cycle regulators. This regulatory interaction is essential for adipogenesis and granulopoiesis in vivo. We now find that mice in which this regulatory interaction is missing develop granulocyte-type leukaemia (see figure), consistent with the presence in acute myeloid leukaemia patients of acquired C/EBP α mutations that impair E2F repression.

Several aspects of liver function, including nitrogen, glucose and lipid metabolism are affected by loss of C/EBP α . We are eliminating specific phospho-acceptor sites in the C/EBP α transactivation domain, whose phosphorylation is regulated by the nutritional state, and analysing the effect on expression of genes controlling glucose metabolism and lipid biosynthesis using DNA microarray technology on mouse tissues. Similar approaches are being used to study regulation of C/EBP α function by signalling-regulated phosphorylation. Finally, by simultaneously mutating C/EBP α and C/EBP β we have found that these proteins cooperate in controlling differentiation of a number of cell types not affected by loss of a single factor, including skin and neuronal cell types.

Future projects and goals

The future projects of the laboratory include simultaneous conditional knockouts of C/EBP proteins in tissues co-expressing multiple C/EBPs. This will be used to study redundant C/EBP functions in skin, liver and hematopoietic cells as well as in neuronal differentiation and cognitive processes (learning, fear). To understand how C/EBP proteins carry out distinct functions in different tissues we are introducing affinity purification tags into the endogenous C/EBP loci as well as those of regulatory transcription factors (PU.1, GATA factors, Friend of GATA) with which they interact in order to delineate their cell type specific proteomes by tandem affinity purification and mass spectrometry.



Bone marrow transformation in mice lacking E2F repression by C/EBP α . In the left panel is a bone marrow cytopsin from a wildtype mouse, showing the presence of differentiated neutrophil granulocytes (NG) and erythroid progenitors (EB: erythroblasts). In the left panel is shown a transformed bone marrow from a mouse homozygous for the BRM2 mutation, which disrupts the interaction between the C/EBP α basic region and the E2F complex. In this bone marrow mainly undifferentiated myeloblasts (MB) and promyelocytes (PM) are observed, indicative of a leukaemic phenotype.

Selected references

- Porse, B.T., Bryder, D., Theilgaard-Monch, K., Hasemann, M.S., Anderson, K., Damgaard, I., Jacobsen, S.E. & Nerlov, C. (2005). Loss of C/EBP α cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage. *J. Exp. Med.*, 202, 85-96
- Nerlov, C. (2004). C/EBP α mutations in acute myeloid leukaemias. *Nat. Rev. Cancer*, 4, 394-400
- Pedersen, T.A., Kowenz-Leutz, E., Leutz, A. & Nerlov, C. (2001). Cooperation between C/EBP α TBP/TFIIB and SWI/SNF recruiting domains is required for adipocyte differentiation. *Genes Dev.*, 15, 3208-3216
- Porse, B.T., Pedersen, T.A., Xu, X., Lindberg, B., Wewer, U.M., Friis-Hansen, L. & Nerlov, C. (2001). E2F repression by C/EBP α is required for adipogenesis and granulopoiesis in vivo. *Cell*, 107, 247-258



Manolis Pasparakis

PhD 1997, the Hellenic Pasteur Institute (awarded by the University of Athens), Athens, Greece.
Postdoctoral research at the Institute of Genetics, University of Cologne, Germany.
Group Leader at EMBL-Monterotondo since 2002.

The NF- κ B signalling pathway in development and inflammatory and degenerative diseases

Previous and current research

The NF- κ B transcription factor family controls the expression of many genes that are important for cell survival and proliferation and for the regulation of immune, inflammatory and stress responses. NF- κ B factors are kept inactive by association with inhibitory proteins belonging to the I κ B family. NF- κ B activating stimuli induce the phosphorylation, polyubiquitination and proteasome degradation of I κ Bs allowing NF- κ B to accumulate in the nucleus and activate target genes. The inducible phosphorylation of I κ B is mediated by the I κ B kinase (IKK) complex, composed of two catalytic subunits termed IKK1 (or IKK α) and IKK2 (or IKK β) and a regulatory protein named NEMO or IKK γ .

NF- κ B has been implicated in the pathogenesis of inflammatory and degenerative diseases. Our aim is to understand the function of IKK-induced NF- κ B activation *in vivo*, focusing particularly on the role of the NF- κ B signalling pathway in disease pathogenesis. To address these questions we are using recombinase-assisted conditional gene targeting to generate mice with tissue-specific mutations in different subunits of the IKK complex.

We have previously shown that targeted disruption of the X-linked NEMO gene in mice causes male embryonic lethality and inflammatory skin lesions in heterozygous females, resembling the human disease incontinentia pigmenti (IP) that is caused by mutations in the human NEMO gene. Using conditional gene targeting we also showed that deletion of IKK2 specifically in epidermal keratinocytes results in TNF-mediated inflammatory skin disease. These experiments revealed that IKK-mediated NF- κ B activity in the epidermis is required for the maintenance of immune homeostasis in the skin.

By combining cell-specific modulation of IKK activity with disease modelling in the mouse, current experiments in our laboratory focus on addressing the role of NF- κ B in inflammatory and degenerative diseases such as Multiple Sclerosis, Alzheimer's disease, type I diabetes and atherosclerosis.

Future projects and goals

Our goal is to understand the mechanisms of NF- κ B function in development and in disease pathogenesis.

Selected references

Schmidt-Suppran, M., Courtois, G., Tian, J., Coyle, A.J., Israel, A., Rajewsky, K. & Pasparakis, M. (2003). Mature T cells depend on signalling through the IKK complex. *Immunity*, 19, 377-389

Pasparakis, M., Schmidt-Suppran, M. & Rajewsky, K. (2002). I κ B kinase signalling is essential for maintenance of mature B cells. *J. Exp. Med.*, 196, 743-752

Pasparakis, M., Courtois, G., Hafner, M., Schmidt-Suppran, M., Nenci, A., Toksoy, A., Krampert, M., Goebeler, M., Gillitzer, R., Israel, A., Krieg, T., Rajewsky, K. & Haase, I. (2002). TNF-mediated inflammatory skin disease in mice with epidermis-specific deletion of IKK2. *Nature*, 417, 861-866

Schmidt-Suppran, M., Bloch, W., Courtois, G., Addicks, K., Israel, A., Rajewsky, K. & Pasparakis, M. (2000). NEMO/IKK γ -deficient mice model incontinentia pigmenti. *Mol. Cell*, 5, 981-992

Walter Witke

PhD 1991, University of Munich, Germany.
Postdoctoral research at the Brigham and Women's
Hospital, Harvard Medical School, USA.
Group Leader at EMBL since 1996.



The actin cytoskeleton and cell migration in mouse development and physiology

Previous and current research

The actin cytoskeleton and its associated proteins represent the organisers and motors for the dynamic shape changes observed in cell crawling, polarisation and cytokinesis. The importance of controlling cell shape is best illustrated by the dramatic cytoskeletal changes observed upon malignant cell growth as well as programmed cell death. Actin binding proteins comprise a large family of proteins known to regulate the properties and dynamics of the actin cytoskeleton. The biochemical mechanisms defined by in vitro experiments can be manifold including crosslinking, severing, capping or nucleation of filaments. To what extent these events modulate the actin cytoskeleton and its functions in motility, endocytosis and cytokinesis is a central question in our group. We are trying to learn how cells employ the actin cytoskeleton in order to perform these specific tasks.

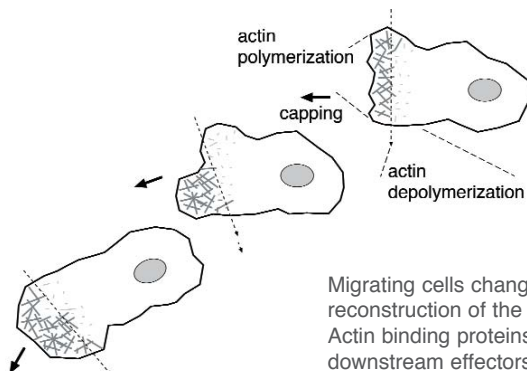
To address these questions we follow an interdisciplinary approach using genetics, cell biology and biochemistry. Mutant mice with specific defects for certain actin binding proteins (Gelsolin, CapG, profilin I) have been generated and analysed, helping us to better understand the role of actin cytoskeleton dynamics in mammalian physiology as well as cellular functions. We apply video microscopy to study cytoskeletal dynamics, chemotaxis and cell motility of primary cells isolated from mutant mice. In vitro assays are used to identify and characterise components of the signalling pathways to the actin cytoskeleton.

Our current and future research will focus on the mechanisms through which cells induce local actin polymerisation and filament breakdown is regulated. Recent work in our laboratory suggests that membrane flow (endocytosis, synaptic recycling, etc.) and actin polymerisation are tightly linked by common regulatory mechanisms, most likely through clusters of specific membrane lipids and membrane associated proteins.

Future projects and goals

Future research in our laboratory will focus on:

- the use of mouse genetics, cell biology and biochemistry to understand the role of profilin I and profilin II in lymphatic organs and in brain;
- regulated expression of mutated profilins using the "cre-lox" technique and generation of conditional mouse mutants;
- biochemical characterisation of profilin ligands and analysis of their cellular functions;
- the study of the regulated breakdown of actin filaments through members of the ADF/cofilin family – the generation of conditional mouse mutants for this set of essential proteins.



Selected references

- Gurniak, C.B., Perlas, E. & Witke, W. (2005). The actin depolymerizing factor n-cofilin is essential for neural tube morphogenesis and neural crest cell migration. *Dev. Biol.*, 278, 231-241
- Witke, W. (2004). The role of profilin complexes in cell motility and other cellular processes. *Trends Cell. Biol.*, 14, 461-469
- Witke, W., Sutherland, J.D., Sharpe, A., Arai, M. & Kwiatkowski, D.J. (2001). Profilin I is essential for cell survival and cell division in early mouse development. *Proc. Natl. Acad. Sci. USA*, 98, 3832-3836
- Witke, W., Podtelejnikov, A.V., Di Nardo, A., Sutherland, J.D., Gurniak, C.B., Dotti, C. & Mann, M. (1998). In mouse brain profilin I and profilin II associate with regulators of the endocytic pathway and actin assembly. *EMBO J.*, 17, 967-976

Index

Akhtar, Asifa	31	Kulozik, Andreas E.	39
Antony, Claude	9, 57	Ladurner, Andreas	34
Apweiler, Rolf	83	Lamzin, Victor S.	75
Arendt, Detlev	22	Le Novère, Nicolas	91
Bastiaens, Philippe	10	Lewis, Joe	58
Benes, Vladimir	60	Luscombe, Nicholas	92
Birney, Ewan	84	Márquez, José	69
Bork, Peer	42	Mattaj, Iain	55
Böttcher, Bettina	43	Meyer-Klaucke, Wolfram	76
Boulin, Christian	11	Minichiello, Liliana	98
Brazma, Alvis	85	Molecular Medicine Partnership Unit	39
Brunner, Damian	12	Mulder, Nicola	93
Cell Biology and Biophysics Unit	7	Müller, Christoph W.	70
Cipriani, Florent	67	Müller, Jürg	35
Cohen, Stephen	21	Müller-Dieckmann, Jochen	77
Conti, Elena	44	Nédélec, François	16
Core Facilities	56	Nerlov, Claus	99
Cusack, Stephen	66	Neumann, Carl	25
de Marco, Ario	62	Pasparakis, Manolis	100
Developmental Biology Unit	20	Pepperkok, Rainer	17, 64
Directors' Research	53	Rebholz-Schuhmann, Dietrich	94
EBI-Hinxton, UK	81	Rice, Peter	87
Ellenberg, Jan	32	Riddell, Andrew	59
EMBL-Grenoble, France	65	Rørth, Pernille	26
EMBL-Hamburg, Germany	72	Rosenthal, Nadia	96
EMBL-Heidelberg, Germany	6	Russell, Rob	48
EMBL-Monterotondo, Italy	95	Sattler, Michael	49
Ephrussi, Anne	23	Sawyer, Alan	61
Frangakis, Achilleas S.	45	Schneider, Reinhard	51
Franz, Thomas	63	Schultz, Carsten	36
Furlong, Eileen	24	Serrano, Luis	41
Gannon, Frank	54	Scheffzek, Klaus	50
Gene Expression Unit	29	Stelzer, Ernst H.K.	18
Gibson, Toby	46	Steinmetz, Lars M.	37
Gilmour, Darren	13	Structural and Computational Biology Unit	40
Goldman, Nick	88	Suck, Dietrich	52
Griffiths, Gareth	14	Surrey, Thomas	19
Gross, Cornelius	97	Svergun, Dmitri	78
Hart, Darren	68	Thornton, Janet	82
Henrick, Kim	86	Treier, Mathias	27
Hentze, Matthias W.	33, 39	Tucker, Paul	79
Hermes, Christoph	74	Weiss, Manfred	80
Hermjakob, Henning	89	Weissenhorn, Winfried	71
Hoenger, Andreas	47	Wilm, Matthias	38
Huber, Wolfgang	90	Wilmanns, Matthias	73
Izaurralde, Elisa	30	Witke, Walter	101
Karsenti, Eric	8	Wittbrodt, Jochen	28
Knop, Michael	15		

Notes

Notes