# EMBL Research at a Glance 2006-2007

EMBL

**EMBL Heidelberg**
Meyerhofstraße 1
69117 Heidelberg
Germany
Tel. +49 (0)6221 387 0, Fax +49 (0)6221 387 8306
www.embl.org
info@embl.org

**EMBL Grenoble**
6, rue Jules Horowitz, BP181
38042 Grenoble, Cedex 9
France
Tel. +33 (0)4 76 20 72 69, Fax +33 (0)4 76 20 71 99

**EMBL Hamburg**
c/o DESY
Notkestraße 85
22603 Hamburg
Germany
Tel. +49 (0)40 89 90 20, Fax +49 (0)40 89 90 21 04

**EMBL-EBI (European Bioinformatics Institute)**
Wellcome Trust Genome Campus, Hinxton
Cambridge CB10 1SD
United Kingdom
Tel. +44 (0)1223 494444, Fax +44 (0)1223 494468

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

**EMBL's Member States:**

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

**Texts:**
EMBL Group and Team Leaders and Heads of Units

**Layout and editing:**
Vienna Leigh, EMBL Office of Information and Public Affairs

**Cover Design:**
Vienna Leigh

**Exposure and printing:**
ColorDruck Kurt Weber GmbH, Leimen

European Molecular Biology Laboratory

Heidelberg • Grenoble • Hamburg • Hinxton • Monterotondo

# Contents

# 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, with Outstations in Grenoble, Hamburg, Hinxton and Monterotondo) has a deep and lasting impact on the wider community. The vast majority of our researchers, who work at EMBL for a fixed length of time, 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. Numerous courses and conferences, an energetic technology transfer activity 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 organise cells 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 allows 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 as well as organisational flexibility. This mix 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 all 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. 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, they combine to stage 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. Interestingly, living cells occupy the precise midpoint between the molecular and macroscopic scales. Thus, in order to understand how organisms are built and how they function, we need to understand 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 cytoskeleton and the secretory pathway. These compartments are permanently renewed by mechanisms that are still poorly understood.

Research in the Cell Biology and Biophysics Unit focuses on the mechanisms and principles that underlie the organisation of these different compartments and the distribution of specific molecules to each cellular sub-system. Cell biologists and physicists at EMBL are therefore trying to define the role of targeting events, as well as that of more complex self-organisation processes in organising cellular space. As a cell enters mitosis, all the microtubules suddenly depolymerise to reassemble into the mitotic spindle. At the same time, the nuclear envelope disappears, the Golgi complex fragments and membrane traffic ceases. 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. A reorganisation of cellular architecture takes place, guided by rules that begin to be unravelled. The elucidation of such rules and 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. 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. Novel developments in microscopy and computer simulations are a particular strength of the unit.


## 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.


# Self-organisation principles in cell morphogenesis

## Previous and current research

We have been working on the mechanism of mitotic spindle assembly in *Xenopus* egg extracts and cells using live cell imaging, FRET technology and computer simulations, in collaboration with physicists and colleagues in the Cell Biology and Biophysics Unit. We found that in large cells and frog eggs, spindle assembly involves two self-organisation principles: 1) the establishment of a gradient of regulators that affect microtubule nucleation and dynamics in the cytoplasm surrounding the chromosomes, and 2) the collective behaviour of dynamic microtubules and motors that results in the assembly of a steady state bipolar structure.

These observations reveal the existence of general principles in the formation of dynamic large cellular structures like the mitotic spindle. We can define such principles as: 1) the stochastic random motion of molecules interacting specifically with each other in the cytoplasm; 2) reaction-diffusion processes that define spatial distributions of regulatory molecules; and 3) collective behaviour of cytoskeletal systems that self-organise in large structures. Principle 2) may coordinate 1) and 3) to generate a precise dissipative structure that maintains its shape over time.

Initial experiments carried out on muscle cell morphogenesis suggest that similar principles are at work for the determination of cell shapes during cell differentiation.

Another important question concerns the nature of interphase and mitotic cytoplasmic "states". Frog eggs are interesting because following fertilisation they undergo a series of twelve divisions during which the cell cycle alternates only between two states: S-phase, during which a nucleus forms and DNA is replicated, and mitosis, during which a spindle forms and chromosomes are segregated. The switch between the two states involves only the continuous synthesis of cyclin coupled to its periodic degradation. When cyclin accumulates to a threshold, it triggers the activation of cdc2 kinase that sends the egg cytoplasm in the mitotic state. The problem is to define what we mean by "cytoplasmic state". We need to take one morphological effect of the interphase and mitotic states and check what is different in the cytoplasm and how such a difference can be brought about by cdc2 kinase. We have chosen to look at microtubule dynamic states. In interphase, microtubules are stable and long; in mitosis, they are short and dynamic. To characterise what could determine this difference, we looked at the interaction state of several molecules involved in the regulation of microtubule dynamics and correlated the state of this network with the effect of removing each component one-by-one on microtubule dynamics in interphase and mitosis. We found that two microtubule stabilisers need to interact in the mitotic situation to permit microtubule growth, whereas in interphase only one stabiliser is enough to generate very long and stable microtubules. On the other hand, a strong microtubule destabiliser is active in mitosis and almost entirely inactive in interphase, because its activity is damped through an interaction with several of the stabilisers. Therefore, the cdc2 kinase changes the state of interaction between a series of regulators when cells enter mitosis and the interphase and mitotic "states" can therefore be defined by the state of this network as far as microtubule dynamics are concerned.

## Future projects and goals

Concerning the mechanism of spindle assembly, there are issues that still need to be addressed generally concerning the principles underlying spindle morphogenesis during evolution. We will use computer simulations in collaboration with the Nédélec lab to establish the parameter space within which various spindles may form. We hope, in the near future, to have a full predictive description of the conditions that would support spindle assembly in vertebrate cells. Similar approaches will be used to examine the conditions of spindle assembly in various species.

Concerning the establishment of interphase and mitotic microtubule dynamic states, we are in the process of identifying the phosphorylation sites involved in the change in the pattern of MAP interactions between interphase and mitosis. We therefore hope to characterise not only the structure of the MAP interaction network but also the structure of the regulatory system, including the nature of the kinase and phosphatase networks involved.

## Selected references


Caudron, M., Bunt, G., Bastiaens, P. & Karsenti, E. (2005). Spatial coordination of spindle assembly by chromosome-mediated signaling gradients. *Science* 309, 1373-1376

Karsenti, E. & Vernos, I. (2001). The mitotic spindle: a self-made machine. *Science* 294, 543-547

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

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

## 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 microtubule arrays in fission yeast

## Previous and current research

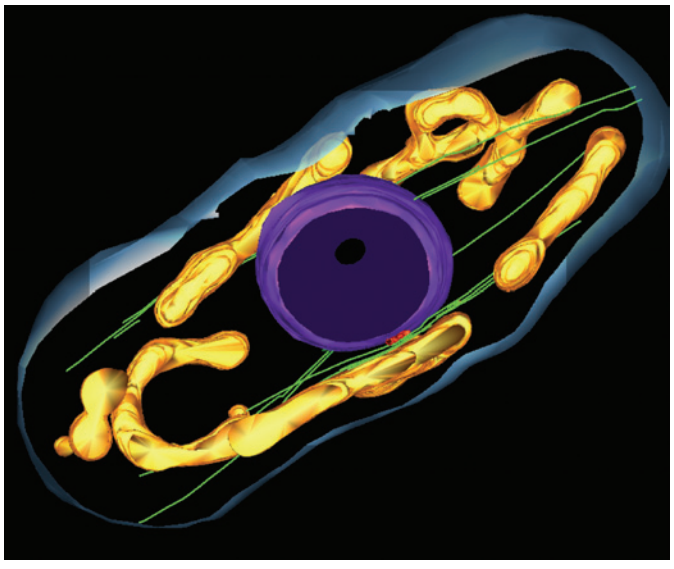
Yeast cells are ideally suited to electron microscopy (EM) of microtubule arrays because the cells are small and are a relatively simple system to study compared to higher eukaryotic cells. In fission yeast we are investigating the structure of the interphase microtubule (MT) bundles and their interaction with membrane organelles using an electron tomography approach on cryofixed cells. This way, detailed morphological features like MT ends or MT bridging factors can be solved (see figure). We have also developed a particular approach to explore and reconstruct systematically large cell volumes that enable to track microtubules their way through the cell. The combination of serial thick sections and systematic montaging during acquisition of each section was the method adopted to collect information from very large and even full cell volumes in the case of yeast cells. Precious structural information can therefore be derived from WT and mutant cells analysis both qualitatively and quantitatively.

## Future projects and goals

After solving the features of microtubule bundles organisation in interphase WT cells in fission yeast, we want to describe various mutant phenotypes, in particular in mutants affected in proteins that control MT dynamics like Mal3 (EB1) and Tip1 (clip-170) MAPs (in collaboration with the Brunner laboratory).

We also intend to study Ase1 deletion strain (which misses this essential MT bundling factor). We will reconstruct metaphase spindles at various stages during their assembly by ET using dual-axis tilt series. We will then focus our analysis of WT cells on the overlap zone of metaphase spindles in order to identify the bonds established between the MTs that emanate from opposite poles and overlapping in the central region of the spindles. In the deletion strain, 3D models of whole metaphase spindles should show the precise alterations induced by the loss of the stabilising factors.

Finally we plan to study MT organisation at the interphase to metaphase transition in the Alp14 ts mutant (Alp14_TOG/XMAP215 family, which is a MT dynamics regulator acting both in interphase and mitosis). We plan to study this mutant at early stages of G2 to M transition, focusing in particular on nuclear MTs initial occurence (in collaboration with the Brunner group).



*3D model of a Wildtype MT cytoskeleton (green) in fission yeast. The volume presented here corresponds to eight thick sections joined together (2um). Such a model is suitable to study the interaction of microtubule bundles with the mitochondriome (gold) and with the nucleus (purple).*

## 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

Carazo-Salas, R.E., Antony, C. & Nurse, P. (2005). The kinesin Klp2 mediates polarization of interphase microtubules in fission yeast. *Science* 309, 297-300

Franz, C., Askjaer, P., Antonin, W., Iglesias, C.L., Haselmann, U., Schelder, M., de Marco, A., Wilm, M., Antony, C. & Mattaj, I.W. (2005). Nup155 regulates nuclear envelope and nuclear pore complex formation in nematodes and vertebrates. *EMBO J.* 24, 3519-3531

Tanaka, K., Mukae, N., Dewar, H., van Breugel, M., James, E.K., Prescott, A.R., Antony, C. & Tanaka, T.U. (2005). Molecular mechanisms of kinetochore capture by spindle microtubules. *Nature* 434, 987-994


## 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 needs to polarise and to 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 correct spatial organisation of the microtubules, which involves the coordinated activity of a number of conserved microtubule-associated factors and varies tremendously between different cell types. Little is known about how this variability is achieved.

We are investigating the molecular mechanisms controlling spatial microtubule organisation using two model organisms – the unicellular fission yeast *Schizosaccharomyces pombe* and the fruit fly *Drosophila melanogaster*. Fission yeast cells are cylindrical. Their microtubules serve to deposit marker proteins that control the exact, antipodal positioning of the two growth sites. In each cell up to 10 anti-parallel microtubules are arranged into 3-6 bundles. The plus ends grow towards the cell poles where they switch to shrinkage, an event termed catastrophe. We have shown that microtubules correctly find and identify cell ends with the help of conserved proteins localising to the growing microtubule plus ends (+TIPs). Tip1p, a homolog of human Clip170, acts by preventing premature microtubule catastrophe if inappropriate regions of the cell cortex are encountered. Tip1p is transported to the MT plus ends by the kinesin motor protein Tea2p. Another +TIP is Mal3p, a homolog of human EB-1. Mal3p promotes microtubule growth throughout the cell. Its removal from microtubule tips is a pre-requisite for catastrophe to occur and is influenced by Tip1p.

We are focusing our fission yeast work on two topics.

1) We investigate why microtubules behave differently when they encounter the cell end cortex as opposed to the remaining cell cortex and how Tip1p contributes. Therefore, we have screened for Tip1p interacting proteins, which are currently being analysed. We have also analysed Pmo25p, a homologue of human MO25. Pmo25p is essential for polarised, cylindrical growth. Intriguingly, it localises to the newly forming cell ends during cell division, making it a candidate cell end identity factor. However, Pmo25p eventually showed two independent functions during cytokinesis and polarised growth.

2) Tip1p and Mal3p are phospho-proteins. We are trying to identify phosphorylation sites and the relevant kinases; we hope this will lead us to the mechanisms that control protein activity locally.

Using *Drosophila* we have started exploring how microtubules are organised in a multicellular organism. Because the +TIPs are highly conserved they are likely to have comparable functions in flies and yeasts. However, in multicellular organisms, the cells in addition need to integrate environmental cues, cell-cell signals and cell adhesion activities, to coordinate microtubule organisation with specific temporal and spatial requirements.

So far we have shown that microtubules are organised in a novel way in epithelial cells during dorsal closure (DC), a developmental process where an epithelial gap is closed. Surprisingly, microtubules are essential only for a particular step of the process, the final sealing.

## Future projects and goals

A central future question we are addressing is how the activity of +TIPs is spatially regulated. We also are interested in the question of whether and how +TIPs interact with the different microtubule targets and microtubule associated organelles (e.g. cell membrane, nucleus, vesicles, mitochondria).

DC-specific microtubule organisation in *Drosophila* is similar to that in fission yeast. 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 differs between cell types. We will also investigate how this machinery is influenced by the signalling and cell adhesion events that coordinate cellular behaviour in multicellular organisms.

## Selected references


Brunner, D. & Nurse, P. (2000). CLIP170-like tip1p spatially organizes microtubular dynamics in fission yeast. *Cell* 102, 695-704

Busch, K.E., Hayles, J., Nurse, P. & Brunner, D. (2004). Tea2p kinesin is involved in spatial microtubule organization 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

## 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.

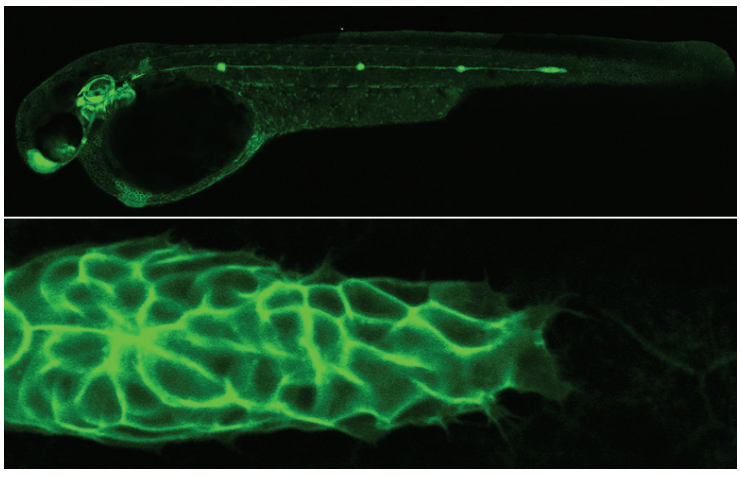

# The role of collective cell migration during organ morphogenesis

## 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 lead to the isolation of a number of molecules required for this process, most notably CXCR4, a chemokine-receptor required for the guidance and shaping 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., Knaut, H., Maischein, H.M. & Nüsslein-Volhard, C. (2004). Towing of sensory axons by their migrating target cells *in vivo*. *Nat Neurosci.* 7, 491-2

Gilmour, D.T., Maischein, H.M. & Nüsslein-Volhard, C. (2002). Migration and function of a glial subtype in the vertebrate peripheral nervous system. *Neuron* 34, 577-88

Haas, P. & Gilmour, D. (2006). Chemokine-signaling mediates self-organizing tissue migration in the zebrafish lateral line. *Developmental Cell* 10, 673-680

Lecaudey, V. & Gilmour, D. (2006). Organizing moving groups during morphogenesis. *Curr. Opin. Cell Biol.* 18, 102-7


## 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

**Phagosomes and mycobacteria:** In our first sub-group we use latex bead phagosomes (LBP) as a simple model system to analyse complex phagosome functions *in vitro* and in macrophages. We recently demonstrated that the assembly of actin by phagosomes facilitates phagosome fusion with late endocytic organelles. Phagosome actin assembly was shown to be dependent on ezrin/moesin and on the phosphoinositides $P1_4P$ and $P1_{415}P_2$. Dozens of lipids and proteins were subsequently identified which, when added to the LBP, stimulated or inhibited the actin assembly process. Our bioinformatic colleagues used pathway analysis algorithms to predict biochemical pathways operating in the phagosomal membrane. This modelling led to predictions, that we experimentally verified, that some lipid kinases can run backwards to a small extent to synthesise ATP. Our data argue that this ATP may be transported into the lumen of the phagosome where it may bind to membrane receptors that signal back to the (cytoplasmic) membrane signalling machinery for actin assembly. A detailed analysis of the role of ezrin in this process is ongoing.

A related project focuses on mycobacterial phagosomes. Non-pathogenic mycobacteria are killed within fully matured phagosomes, while those enclosing the pathogens, such as *M. tuberculosis*, block phagosome fusion with lysosomes. This inhibition could be related to the live bacteria's ability to block phagosome actin assembly machinery. Lipids and other factors that stimulate phagosome actin assembly reduce killing of pathogens in macrophages whereas inhibition of actin leads to more growth. We are now addressing the question, how do macrophages kill mycobacteria? Our data argue that a minor role is played by iNOS-derived nitric oxide (and derivatives) and a more important role is due to the fusion of phagosomes with at least four different late endocytic compartments. The killing factors remain to be identified, however.

**Vaccinia virus project:** This sub-group, under the supervision of Jacomine Krijnse-Locker, addresses two questions. First, how is the complex cytoplasmic life cycle of VV organised within the infected cells and what is the role of cellular membranes and the cytoskeleton? Second, what is the structure and assembly of the intact virion? In the past year two techniques have considerably contributed to these main questions; live cell imaging and electron tomography. We could show that VV-induced cellular motility contributes significantly to the cytoplasmic organisation of distinct stages in the VV life cycle. Electron tomography (ET) has enabled us to make considerable progress in understanding the biogenesis of the precursor membranes of the intact virion. The development of cryo-ET of whole intact cells now enables us to study the structural intermediates of entry and disassembly of VV at the plasma membrane in detail.

## Future projects and goals

**Phagosome project:** A detailed global understanding of the signalling networks regulating phagosome actin assembly is one of our long-term goals. The more we understand about these membrane signalling networks the more likely we are to be in a position of comprehending how pathogens such as *M. tuberculosis* actively modulate these networks to its advantage. The second long-term goal is to identify macrophage phago-lysosome killing factors. We hope to provide a rational basis for boosting the innate capacity of these cells in humans to kill pathogens.

**Poxvirus project:** Our goal is to continue our two main projects. In the immediate future we would like to understand how VV-induced cell motility is regulated. Cell motility is regulated in time, as the motile events take place early in infection, whereas late in infection motility ceases. The induction of motility is regulated by viral early proteins, whereas its cessation may be regulated by late viral proteins. Since MVA does not induce cell motility and since this virus misses about 60 genes compared to VV, it provides us with an additional genetic system to search for the molecules involved. Efforts in tomography, especially cryo ET, will be continued in order to unravel the complex structure and assembly of VV.

## Selected references


Anes, E., Kuhnel, 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

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

Kjeken, R., Egeberg, M., Habermann, A., Kuehnel, M., Peyron, P., Floetenmeyer, M., Walther, P., Jahraus, A., Defacque, H., Kuznetsov, S.A. & Griffiths, G. (2004). Fusion between phagosomes, early and late endosomes: A role for actin in fusion between late, but not early endocytic organelles. *Mol. Biol. Cell* 15, 345-358

Meiser, A., Sancho, C. & Krijnse-Locker, J. (2003). Plasma membrane budding as an alternative release mechanism of the extracellular enveloped form of vaccinia virus from HeLa cells. *J. Virol.* 77, 9931-9942

Palacios, S., Perez, L.H., Welsch, S., Schleich, S., Chmielarska, K., Melchior, F. & Krijnse-Locker, J. (2005). Quantitative SUMO-1 modification of a vaccinia virus protein is required for its specific localization and prevents its self-association. *Mol. Biol. Cell.* 16, 2822-35

Schramm, B. & Krijnse Locker, J. (2005). The cytoplasmic organization of Poxvirus DNA-replication. *Traffic* 6, 839-846

## Michael Knop

PhD 1995, University of Stuttgart, Germany.
Postdoctoral research at the Max Planck Institute 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.


# Meiosis and sex in yeast

## Previous and current research

Our laboratory is interested in the sexual cell division (meiosis) and reproduction with a focus on the connection of the evolutionary/population genetic implications with the underlying molecular mechanisms. Primarily we use the bakers yeast *Saccharomyces cerevisiae*, but also other yeast species, as model systems.

The sexual cell division evolved as a mechanism that allows organisms to better face the challenges of changing environments. Consequentially, the sexual cycle is adapted to fulfill the particular adaptive needs of a species with respect to the characteristics of the habitat it is occupying.

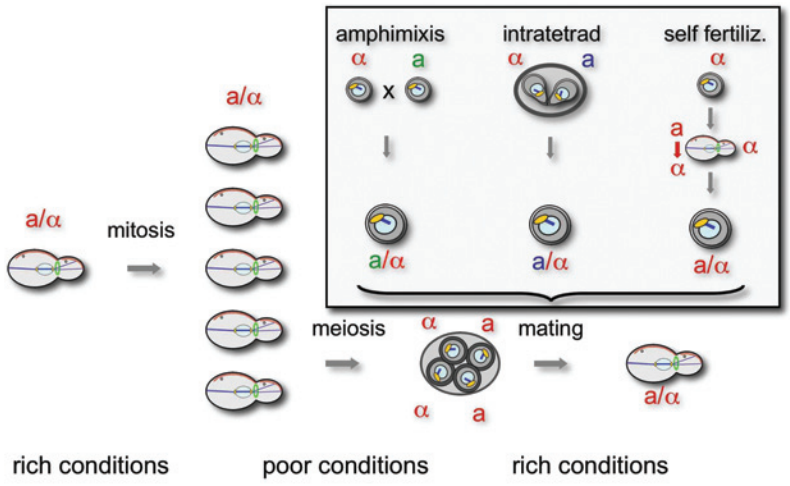
For budding yeast, we demonstrated that meiosis is governed by a regulatory mechanism that generates a particular outcome (spores) in such a way, that their number is adjusted to available resources (nutrients) and that simultaneously they match a particular and advantageous mating behaviour. We used cell biological and genetic methods in combination with computer simulation to unravel the function of a self-organising molecular mechanism that is able to generate this behaviour.

We are also interested in other molecular processes that govern the meiotic type of cell division in budding yeast. In particular, we are studying the modes in which basic molecular machinery, such as that involved in vesicle fusion or organisation of the cytoskeleton, becomes regulated during meiosis, so that it generates the different output seen for the meiotic type of cell division as compared to mitosis.

## Future projects and goals

Among other projects, we aim to understand the following processes:

- regulation of spindle pole body function during meiosis;
- cell shape generation during meiosis;
- impact of different breeding strategies, in particular inbreeding, on the genome organisation.




*Regulation of the yeast life cycle. The life cycle of budding yeast is adopted to fluctuating environments in which poor and rich phases follow each other. The spores constitute the life cycle stage adapted to adverse conditions. Upon return to a rich environment, spores germinate and seek mating partners in order to return to a diploid life cycle stage and to prepare for a subsequent period of poor nutrient supply. Upon encountering it, they induce meiosis and form spores.*

## Selected references


Knop, M. (2006). Evolution of the hemiascomycete yeasts: on life styles and the importance of inbreeding. *BioEssays*, in press

Knop, M., Miller, K.J., Mazza, M., Feng, D., Weber, M., Keranen, S. & Jantti, J. (2005). Molecular Interactions position Mso1p, a novel PTB domain homologue, in the interface of the exocyst complex and the exocytic SNARE machinery in yeast. *Mol. Biol. Cell* 16, 4543-4556

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

Riedel, C.G., Mazza, M., Maier, P., Korner, R. & Knop, M. (2005). Differential requirement for phospholipase D/Spo14 and its novel interactor Sma1 for regulation of exocytotic vesicle fusion in yeast meiosis. *J. Biol. Chem.* 280, 37846-37852

Taxis, C., Keller, P., Kavagiou, Z., Jensen, L.J., Colombelli, J., Bork, P., Stelzer, E.H.K. & Knop, M. (2005). Spore number control and breeding in *Saccharomyces cerevisiae*: a key role for a self-organizing system. *J. Cell Biol.* 171, 627-640


## 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


Goshima, G., Nédélec, F. & Vale, R.D. (2005). Mechanisms for focusing mitotic spindle poles by minus end-directed motor proteins. *J. Cell Biol.* 171, 229-240

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. *Journal of Cell Biology* 6, 1005-1015

Surrey, T., Nédélec, F., Leibler, S. & Karsenti, E. (2001). Physical properties determining self-organization 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 detail. 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). Gamma-BAR, a novel AP-1-interacting protein involved in post-Golgi trafficking. *EMBO J.* 24, 1122-1133

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

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

Young, J., Stauber, T., del Nery, E., Vernos, I., Pepperkok, R. & Nilsson, T. (2005). Regulation of microtubule-dependent recycling at the trans-Golgi network by Rab6A and Rab6A’. *Mol. Biol. Cell* 16, 162-177


## Ernst H.K. Stelzer

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


# Modern life sciences require the third dimension: leaving flatland behind

## Previous and current research

The light microscopy group has been through many changes during the past few years. For almost two decadest it has performed biological research, developed sophisticated microscopes and equipment and developed many methods and devices that are now in common use. The many collaborations with scientists all over the world has assured that various laser-based microscopy techniques are used for biological, biophysical and physical research.

Scientists working in the life sciences are employed to apply and guide the developments, alongside physicists, mathematicians and engineers, who devote their efforts to the physical and technical challenges. The group currently maintains a healthy mix of basic and applied research, technical development and applications in various fields of the modern life sciences.

The usage of the various available lasers for research in the life sciences is certainly a central topic. However, laser scanning for imaging purposes has become less important and has largely (but not entirely) been replaced by the use of lasers for manipulating very tiny as well as very large objects.

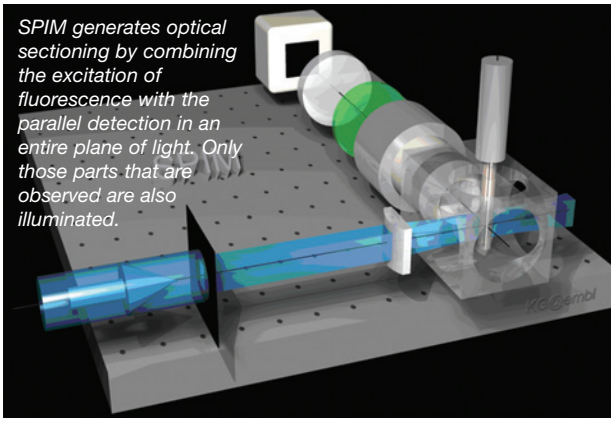
The Photonic Force Microscope (PFM) can be used to determine the location of small beads attached to single molecules or cytoskeletal elements. It has a precision in the low nm range and can operate at MHz rates. The diffraction limited pulsed UV laser cutter (laser nanoscalpel) has turned out to be a versatile tool in the manipulation of single cells as well as early embryos. It is particularly useful for *in vivo* studies of the cytoskeleton. The Single Plane Illumination Microscope (SPIM) has proved a great success and will become the basis for numerous new developments and experiments.

Particularly gratifying successes are our projects in quantitative biology, most recently the quite detailed description of the molecular basis of yeast spore formation. In our biological projects we place a special emphasis on a three-dimensional cell biology or biophysics. We try to avoid flat and hard surfaces (e.g. cover slips) since they usually do not exist in tissue and tend to generate artefacts while providing an un-physiological environment.

## Future projects and goals

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

We are currently witnessing the transition from a "flat" 2D cell biology to a more complex 3D cell biology. The evidence that cells cultivated on cover slips differ dramatically from their 3D natural counterparts, which depend on other cell types to support them, is extremely strong and should be addressed. For these reasons we have continued our work with various model embryos and put a lot of effort into establishing conditions that allow us to work with cells that are cultivated in 3D. We plan not only to develop tools such as 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.



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

## Selected references


Colombelli, J., Reynaud, E.G., Rietdorf, J., Pepperkok, R. & Stelzer, E.H.K. (2005). *In vivo* selective cytoskeleton dynamics quantification in interphase cells induced by pulsed ultraviolet laser nanosurgery. *Traffic* 6, 1093-1102

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

Keller, P.J., Pampaloni, F. & Stelzer, E.H.K. (2006). Life sciences require the third dimension. *Curr. Opin. Cell Biol.* 18, 117-124

Taxis, C., Keller, P., Kavagiou, Z., Jensen, L.J., Colombelli, J., Bork, P., Stelzer, E.H.K. & Knop, M. (2005). Spore number control and breeding in *Saccharomyces cerevisiae*: a key role for a self-organizing system. *J. Cell Biol.* 171, 627-640

## 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

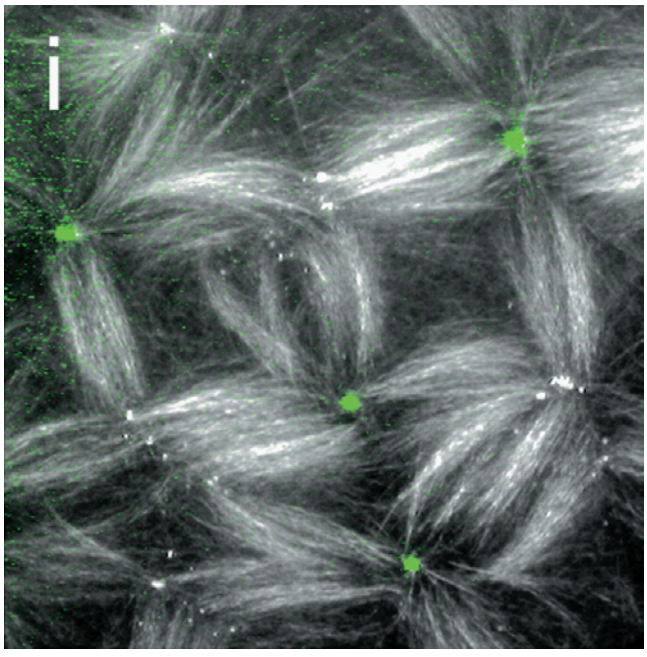
Motor proteins are key determinants for the spatial organisation of eukaryotic cells. They are thermodynamic non-equilibrium machines playing a crucial role for the dynamic nature of cellular order. In fact, they provide a paradigm for the concept of intracellular order depending on molecular dynamics. 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

Seitz, A. & Surrey, T. (2006). Processive movement of single kinesins on crowded microtubules visualized using quantum dots. *EMBO J.* 25, 267-277

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

# Developmental Biology Unit

Groups in the Developmental Biology (DB) Unit use a variety of organisms to address fundamental problems in animal development. A common theme is the concern with biological processes at the organismal level. Within that context, our interests range from the level of individual cells – e.g. sub-cellular localisation of RNA, polarized cellular responses to signalling gradients – to the organismal – e.g. control of tissue growth and metabolism.

One highlight from the DB unit in 2005 has been the "evo-devo" work by the Arendt lab. Evolution of developmental mechanisms is an area of keen interest to developmental biologists. The field has undergone a transition over the years from a primarily descriptive science, to one based in the "hard fact" of DNA sequence comparison. Whole genome sequences are becoming available for a growing number of animals, covering a great range of evolutionary diversity. Comparative genomics has begun to shed new light on evolution. The Arendt lab, in collaboration with the Bork lab from Structural and Computational Biology (SCB), have provided important new insights into animal evolution through their recent studies on the organisation of genes in the marine polychaete *Platynereis* (Raible et al., 2005, *Science*). At the level of RNA splicing, *Platynereis* looks a lot more like humans than it looks like other popular invertebrate model organisms like *Drosophila* and *C. elegans*. This contrasts with what might have been expected from its morphology.

The molecular basis for cell fate specification and differentiation are core interests of developmental biologists. An emerging part of this picture involves whole genome expression analysis. Various groups in the DB Unit are engaged in generating data-sets and databases that permit the use of expression pattern data as a computable resource. Quantitative data from expression profiling is one part of the picture. Another is qualitative assessment of the spatial and temporal patterns of gene expression that are rendered useful for computation by describing them with a controlled, hierarchical annotation vocabulary, similar to GO terms. Computational analysis based on spatial expression data has already proven useful in understanding microRNA function (a collaboration between the Cohen (DB) and Russell (SCB) labs). Expression data from zebrafish, medaka and *Platynereis* are being integrated in a project coordinated in the context of the Computational Biology Centre and with the participation of the Wittbrodt, Furlong, Steinmetz and Arendt labs. An important part of this effort involves collaboration with colleagues in the SCB Unit in Heidelberg and at the EBI to develop new tools for analysis of these large data sets. Integration of large-scale expression data with analysis of transcription factor binding through whole genome "chromatin immunoprecipitation" and with computational tools to predict gene-regulatory elements (developed in a collaborative effort of the Wittbrodt and Birney labs) will be important in developing a system-level view of the control of cell identity in animal development. The Arendt lab has begun to build on this type of thinking to view evolution from a cell-type perspective: suites of genes define a cell type. Considering evolution at the level of cell-type-specific gene expression "signatures", not only at the level of molecules, is emerging as a powerful new paradigm. Together these approaches will provide the opportunity to undertake ambitious longer-term systems biology goals, such as modelling regulatory networks controlling morphogenesis and organogenesis.

Several groups in the DB Unit have interests in disease mechanism. Recent work in the Treier lab has generated new mouse models for endocrine cancer, premature ovarian failure, polycystic kidney disease and obesity. Expression profiling has begun to provide important insights into the molecular basis of these diseases. Studies in other organisms on control of cell proliferation, epithelial mesenchymal transition, apoptosis and metabolism show promise of leading to the development of new mouse disease models in years to come.

## Steve 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 organise 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 have taken molecular and genetic approaches to understanding how these morphogen gradients are established and how they control tissue growth during development.

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.

The mechanisms that control growth of cells and tissues during embryonic life are closely linked to control of metabolism. Recent work in the lab has identified a number of genes implicated in control of metabolism as well as tissue growth. Among these are novel modulators of the Insulin/TOR signalling pathway. Ongoing work is aimed toward assessing their functions in mammalian systems.

## 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

Stark, A., Brennecke, J., Bushati, N. Russell, R.B. & Cohen, S.M. (2005) Animal microRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123,1133-1146

Teleman, A.A., Chen, Y.W. & Cohen, S.M. (2005a). *Drosophila* Melted modulates FOXO and TOR activity. *Dev. Cell* 9, 271-281

Teleman, A.A., Chen, Y.-W. & Cohen, S.M. (2005b) 4E-BP functions as a metabolic brake used under stress conditions but not during normal growth. *Genes Dev.* 19, 1844-1848

Thompson, B.J., Mathieu, J., Sung, H.H., Loeser, E., Rorth, P. & Cohen, S.M. (2005). Tumor suppressor properties of the ESCRT-II complex component Vps25 in *Drosophila*. *Dev. Cell* 9, 711-720


## Detlev Arendt

PhD 1999, Albert-Ludwigs-Universität, 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 in coastal regions.

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.

## 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)*

## Selected references


Arendt, D. (2005). Genes and homology in nervous system evolution: Comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory in Biosciences* 124, 185-197

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

Raible, F., Tessmar-Raible, K., Osoegawa, K., Wincker, P., Jubin, C., Balavoine, G., Ferrier, D., Benes, V., de Jong, P., Weissenbach, J., Bork, P. & Arendt, D. (2005). Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* 310, 1325-1326

## Anne Ephrussi

PhD 1985, Massachusetts Institute of Technology, USA.
Postdoctoral work at Harvard University and the Whitehead Institute, USA.
At EMBL since 1992.
Dean of Graduate Studies. EICAT Coordinator.


# 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.

In *Drosophila*, establishment of the body axes of the embryo relies on asymmetric localisation of cell fate determinants within the oocyte, before fertilisation. *oskar* is localised as an mRNA at the oocyte posterior pole. Oskar directs abdomen and germline formation. Tight restriction of Oskar activity to the posterior is critical and is achieved by mRNA localisation-dependent translation: *oskar* is translationally repressed during transport and activated once the mRNA reaches the posterior pole. Assembly of the *oskar* mRNA localisation/translational repression 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. Coassembled with the EJC and proteins bound to its 3'UTR, *oskar* mRNA is transported in particles containing multiple *oskar* mRNA molecules, by a mechanism involving kinesin heavy chain to the posterior of the oocyte, where it is translated and anchored. *oskar* mRNA translation is repressed at initiation by dual mechanisms, one cap-dependent and one cap-independent, that inhibit small ribosomal subunit recruitment. The mechanistically novel, cap-independent mechanism involves formation of large RNP particles ("silencing particles") containing known translational repressors of oskar, several of which are also P-body components.

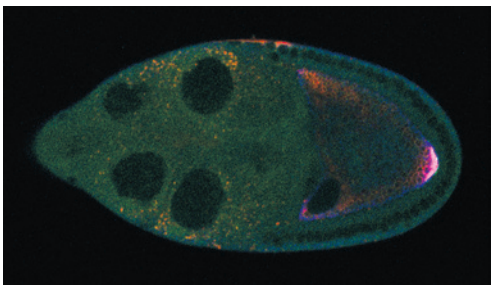
As oocyte polarity is critical for correct *oskar* mRNA localisation, this system is ideally suited to genetic, molecular, and cell biological investigation of the processes of cell polarisation, mRNA localisation and translational control. We make use of the polarised *Drosophila* oocyte as a model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

- Oskar protein and its assembly of the polar granules, the germline granules of *Drosophila*;
- the mechanisms underlying determination and polarisation of the *Drosophila* oocyte;
- 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;
- the relationship between *oskar* mRNA "silencing particles" and P-bodies.

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



*A* Drosophila *egg-chamber, showing colocalisation of oskar mRNA, Staufen protein and a microtubule polarity marker at the posterior of the oocyte.*

## Selected references


Chekulaeva, M., Hentze, M.W. & Ephrussi A. (2006). Bruno acts as a dual repressor of *oskar* translation, promoting mRNA oligomerization and formation of Silencing Particles. *Cell* 124, 521-533

Hachet, O. & Ephrussi, A. (2004). Splicing of *oskar* RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* 28, 959-963

Riechmann, V. & Ephrussi, A. (2004). Par-1 regulates *bicoid* mRNA localization by phosphorylating Exuperantia. *Development*, 131, 5897-5907

Vaccari, T., Rabouille, C. & Ephrussi, A. (2005). The *Drosophila* PAR-1 spacer domain is required for lateral membrane association and for polarization of follicular epithelial cells. *Curr. Biol.* 15, 255-261

Yano, T., López de Quinto, S., Matsui, Y., Shevchenko, A., Shevchenko, A. & Ephrussi, A. (2004). Hrp48, a *Drosophila* HnRNP A/B homolog, binds and regulates translation of *oskar* mRNA. *Developmental Cell* 6, 637-648

Yoshida, S., Müller, H.-A. J., Wodarz, A. & Ephrussi, A. (2004). PKA-RI 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.


# Regulatory networks required for tissue development – understanding the logic

One of the major challenges in developmental biology is to understand how cells become specified and organised to form complex tissues. Tissue and organ development requires precise integration of transcriptional and signalling networks in both space and time. While our understanding of the input signals required to select different cell fates is high, we still have a very poor understanding of how these regulators generate the desired output. Our work aims to bridge this gap by combining genomic, genetic and computational approaches to understand firstly what genes are directly regulated by key transcription factors during specific steps of tissue development and secondly how these genes work together to form functional modules that drive developmental progression. Of particular interest is the regulation of mesoderm subdivision into different muscle primordial and subsequent myoblast specification.

The progression of an undifferentiated mesodermal cell towards a mature muscle requires the activity of a number of well-characterised transcription factors. Despite their conserved role, very little is known about the molecular mechanism by which these transcription factors regulate these processes. Few of their direct target genes or effecter molecules are known. We are applying Chromatin ImmunoPrecipitation followed by microarray analysis (ChIP-on-chip) in combination with expression profiling to gain a global understanding of how transcription factors execute the myogenic program. ChIP-on-chip experiments in a developmental time-course have allowed us to identify temporally regulated groups of target genes. This work is ongoing, but some exciting patterns are starting to emerge. Subgroups of transcription factors selectively co-bind to sets of enhancers in a temporally regulated manner. The large groups of active developmental enhancer regions identified have enabled computational approaches to search for other transcription factors that are likely to also occupy these enhancers. We have also identified a number of new muscle specific transcription factors, which will further enrich the network.

## Future projects and goals

A number of experimental (including genetics, genomics, immuno-staining, biochemistry, tissue culture) and computational approaches will be used to address the following:

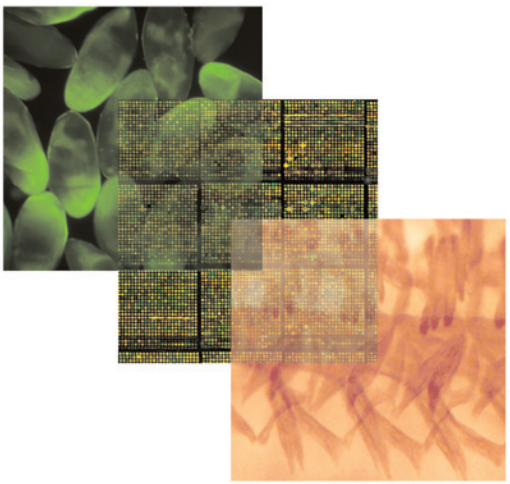
1) To dissect the regulatory code that provides precise temporal and spatial expression of overlapping subsets of genes during specific stages of myogenesis and in different types of myoblasts.

2) To identify new functions for these well characterised transcription factors, which may have been masked by the complexity of their mutant phenotypes.

3) To integrate systematic data from our ChIP-on-chip experiments, expression profiling and computational analysis to construct a regulatory network that describes key events in myogenesis.

4) To examine functional redundancy and robustness within the network.

5) To use both the topology and dynamics of the network to decipher the logic of how these integrated circuits regulate specific aspects of tissue development.



## Selected references


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

Artero, R., Furlong, E.E., Beckett, K., Scott, M.P. & Baylies, M. (2003). Notch and Ras signaling pathway effector genes expressed in fusion competent and founder cells during *Drosophila myogenesis*. *Development* 130, 6257-6272

Furlong, E.E. (2004). Integrating transcriptional and signalling networks during muscle development. *Curr. Opin. Genet. Dev.* 14, 343-350

Furlong, E.E. (2004). Whole genome approaches to studying *Drosophila* muscle development. In "Muscle Development in *Drosophila*", Sink, H. (ed.), Landes Bioscience Publishing

Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P. & Furlong, E.E. (2006). A temporal map of transcription factor activity: Mef2 directly regulates target genes at all stages of muscle development. *Dev Cell*, 10, 797-807

## Carl Neumann

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


# The genetic control of vertebrate organogenesis – eye and limb development in zebrafish

## Previous and current research

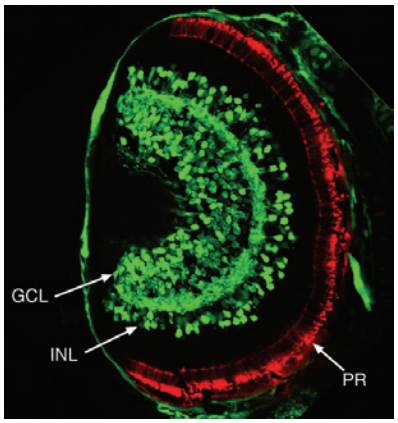
Organs are specialised tissues required for the survival and integrity of multi-cellular organisms. The first step in organogenesis is the specification of an organ primordium – the establishment of a group of cells committed to form an organ. This step is critical, since it determines where an organ develops in the embryo. Secondly, the cells in the organ primordium proliferate to an appropriate degree, and positional information is generated within the developing tissue, leading to distinct cellular identities in different regions of the organ. Finally, in order to complete organogenesis, precursor cells must undergo cell cycle exit at the correct time and location, and differentiate as cell types specialised for their context.

We are studying the genetic control of several steps in organogenesis, using the zebrafish, *Danio rerio*, as a model system, and focusing mainly on two organ systems: the paired appendages and the eyes. The paired appendages of zebrafish are homologous to the limbs of land vertebrates, and just like these, they develop from buds arising from the lateral plate mesoderm. In this context, we are interested in the process whereby organogenesis is initiated, and are analysing a group of zebrafish genes important for initiation of limb development. This group includes genes affecting retinoic acid, Wnt and Fgf signalling, as well as transcription factors such as Tbx5 and Prdm1. Our analysis of these genes has clarified their regulatory hierarchy, and has revealed how positional information present in the trunk of the embryo is transferred to the lateral plate mesoderm to trigger limb formation.

The vertebrate retina develops from an evagination of the neural tube. It is initially a non-neuronal, un-patterned epithelium, which develops into a highly organised three-dimensional array of neurons. Differentiation in the retina proceeds in a wave from the centre to the periphery, and is preceded by cell-cycle withdrawal of retinal precursor cells. In the context of the retina, our aim is to understand how cell-cycle exit is coordinated with differentiation, and how cell-cycle exit and cell fate specification are controlled by cell-cell signalling. We have recently shown that the Sonic hedgehog signalling protein functions as a short-range signal to trigger cell-cycle exit and differentiation in the zebrafish retina, and that this effect of Sonic hedgehog is mediated by transcriptional activation of the cyclin-kinase inhibitor p57Kip2. p57Kip2 inhibits activity of CDKs that drive G1 progression, and thereby promotes cell-cycle exit. These results shed light on the mechanism whereby cell-cycle exit is controlled by cell-cell signalling during eye development.

## Future projects and goals

In order to systematically characterise the genetic network controlling limb induction, we plan to perform expression profiling experiments on zebrafish mutants affecting limb induction. We also plan to examine the transcriptional regulation of genes important for initiation of limb development. To gain further insight into the mechanisms guiding cell-cycle progression and cell-cycle exit during development, we are analysing the role of several signalling pathways in directing the cell-cycle during organogenesis. Finally, since nearly all aspects of organogenesis are controlled by cell-cell signalling, we are also examining the role of heparan sulphate proteoglycans in regulating signalling proteins in the extracellular space.




*Organogenesis generates organs with differentiated cells in a precise pattern. In the retina, ganglion cells and amacrine cells (green) are found in the inner layers, whereas photoreceptors (red) are in the outer layer. GCL: ganglion cell layer. INL: inner nuclear layer.*

## Selected references


Fischer, S., Draper, B.W. & Neumann, C.J. (2003). The zebrafish fgf24 mutant identifies an additional level of Fgf signaling involved in vertebrate forelimb initiation. *Development* 130, 3515-3524

Norton, W.H., Ledin, J., Grandel, H. & Neumann, C.J. (2005). HSPG synthesis by zebrafish Ext2 and Extl3 is required for Fgf10 signalling during limb development. *Development* 132, 4963-4973

Shkumatava, A. & Neumann, C.J. (2005). Shh directs cell-cycle exit by activating p57Kip2 in the zebrafish retina. *EMBO Rep.* 6, 563-569

Stadler, J.A., Shkumatava, A., Norton, W.H., Rau, M.J., Geisler, R., Fischer, S. & Neumann, C.J. (2005). Histone deacetylase 1 is required for cell cycle exit and differentiation in the zebrafish retina. *Dev. Dyn.* 233, 883-889


## Pernille Rørth

PhD 1993, University of Copenhagen, Denmark.
Staff Associate at the Carnegie Institute 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

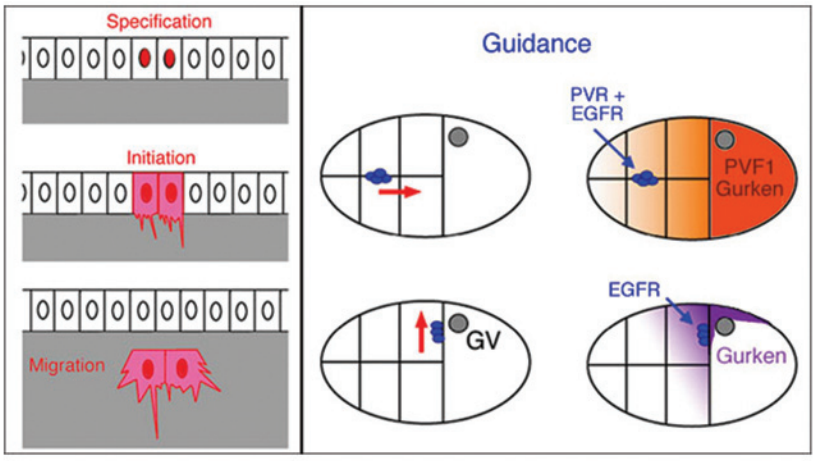
During animal development, many cells migrate from one place to another in order to perform their biological functions. Cell migration is a complex process that involves dynamic interactions between migrating cells and the tissue through which they migrate. In order to migrate, cells change their shape 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.

## 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.




## Selected references


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

Jekely, G., Sung, H.H., Luque, C.M. & Rørth, P. (2005). Regulators of endocytosis maintain localized receptor tyrosine kinase signaling in guided migration. *Dev. Cell* 9, 197-207

Rørth, P. (2002). Initiating and guiding migration: lessons from border cells. *Trends Cell Biol.* 12, 325-331

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

## Mathias Treier

PhD 1994, 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

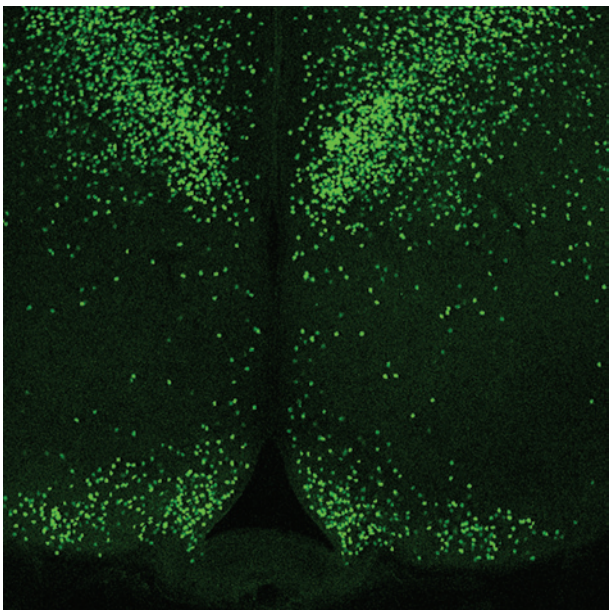
The development of specific cell types during mammalian organogenesis has been studied intensively in the last decade. The challenge now is to understand how these different cell types function in the context of a whole organ, and the ultimate goal is to determine how all the different organs function 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; the thyroid gland, the adrenal gland and the 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 FP6 consortium we are studying the development and function of hypothalamic circuitries that are implicated in this process (www.eurodiabesity.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. Firstly, 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. Secondly, 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.pkdcure.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 signaling is required for pituitary gland development. *Development* 128, 377-386

Treier, M., Gleiberman, A.S., O'Connell, S.M., Szeto, D.P., McMahon, J.A., McMahon, A.P. & Rosenfeld, M.G. (1998). Multistep signaling requirements for pituitary organogenesis in vivo. *Genes Dev.* 12, 1691-1704


## 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*.

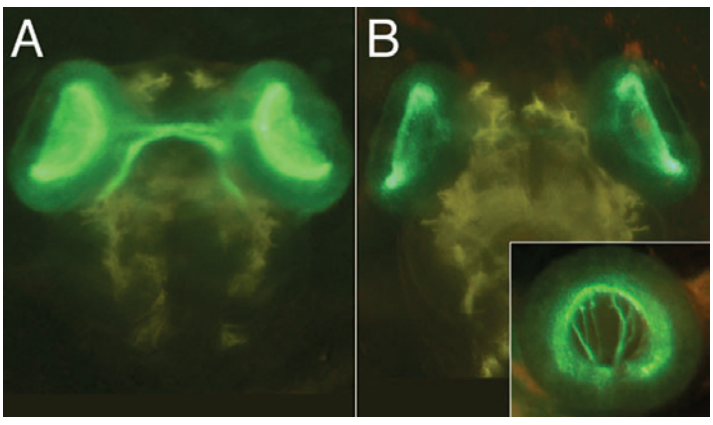
## Selected references


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. & Wittbrodt, J. (2005). Cell cycle control by homeobox genes in development and disease. Semin. *Cell Dev. Biol.* 16, 449-460

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

Martinez-Morales, J.R., Del Bene, F., Nica, G., Hammerschmidt, M., Bovolenta, P. & Wittbrodt, J. (2005). Differentiation of the vertebrate retina is coordinated by an FGF signaling center. *Dev. Cell* 8, 565-574


*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 training at EMBL, Germany.
Research Group Leader at the University of Geneva.
Group Leader at EMBL since 1999.
Ad interim coordinator of the Gene Expression Unit since May 2005.


# Post-transcriptional regulation of messenger RNA

## Previous and current research

For close to three decades, gene expression was thought to be mainly regulated at the transcriptional level. The discovery of RNA silencing pathways, alongside the realisation that post-transcriptional control provides conserved mechanisms by which cells can rapidly change gene expression patterns, have led to a renaissance in the field of post-transcriptional regulation. Post-transcriptional processes (e.g. mRNA processing, export, surveillance, silencing and turnover) are interlinked by the use of common factors and constitute a complex regulatory network that contributes to cell-type and organism specific gene expression patterns.

Our long-standing goal is the elucidation of the molecular mechanisms that regulate gene expression at the post-transcriptional level, using a combination of small-scale and functional genomic approaches.

## Future project and goals

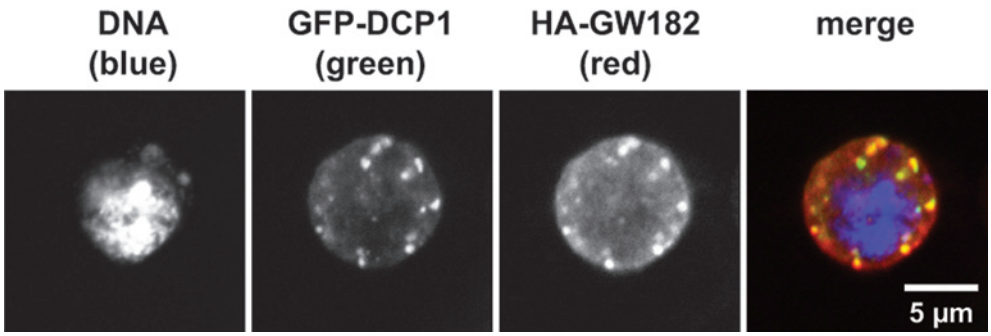
During its entire lifespan, the mRNA acts as a platform for the binding of numerous proteins, and exists in the cell as a ribonucleoprotein particle (mRNP). It is the mRNP which is the actual substrate of post-transcriptional processes. Consequently, many RNA binding proteins play roles at multiple steps of the post-transcriptional pathway. In addition, enzymes involved in general mRNA degradation and proteins involved in mRNA surveillance, RNA silencing and translational repression, co-localise in discrete cytoplasmic foci known as mRNA processing bodies or P-bodies (see figure), suggesting that these processes are interlinked.

Our aims are:

1) to build a comprehensive list of effectors and interaction networks for post-transcriptional pathways in order to understand how these pathways are connected with each other and with different cellular processes;

2) to determine the contribution of these post-transcriptional processes to gene expression on a global scale and to identify endogenous targets;

3) to understand how the regulation of the endogenous targets leads to the complex phenotypes observed at the cellular and organism level when these pathways are perturbed.




*P-bodies in* Drosophila *Schneider cells. P-bodies are discrete cytoplasmic domains where proteins required for bulk mRNA degradation, mRNA surveillance, RNA silencing and translational repression co-localise. The image shows the co-localisation of the mRNA decapping protein DCP1 with the GW182 antigen, a P-body marker in multicellular organisms.*

## Selected references


Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. (2005). The structure of the flock house virus B2 protein, a viral suppressor of RNA interference, shows a novel mode of double-stranded RNA recognition. *EMBO Rep.* 6, 1149-1155

Orban, T.I. & Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* 11, 459-69

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D. & Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* 11, 1640-1647

## 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

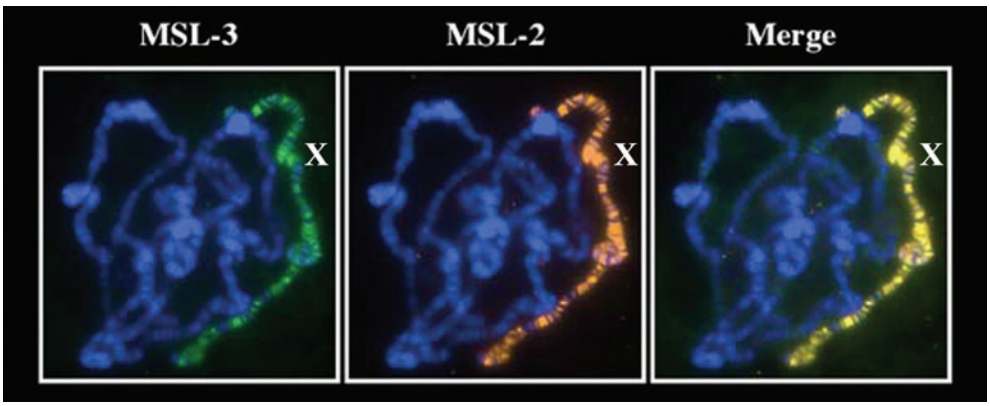
DNA tightly packed together with histones into nucleosomes is not easily accessible to the enzymes that use it as a template for transcription or replication. Consequently, remodelling of chromatin structure may play an essential role in the regulation of gene expression. Structural changes in chromatin 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.

Using biochemical approach we have recently revealed an unexpected physical and functional connection between nuclear pore components and chromatin regulation by MSL proteins, highlighting the role of nucleoporins in gene regulation in higher eukaryotes.

## 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. Moreover, we will be working towards understanding the mechanism by which nuclear pore components influence X-linked gene expression.

It is intriguing that there is a remarkable evolutionary conservation of the MSL complex 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


Akhtar, A. & Matera, A.G. (2006). In and around the nucleus. *Nat Cell Biol.* 8, 3-6

Buscaino, A., Legube, G. & Akhtar, A. (2006). X-chromosome targeting and dosage compensation are mediated by distinct domains in MSL-3. *EMBO Rep.* 7, 531-8

Legube, G., McWeeney, S.K., Lercher, M.J. & Akhtar, A. (2006). X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in *Drosophila*. *Genes Dev.* 20, 871-83

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., Wilm, M., Stunnenberg, H.G., Saumweber, H. & Akhtar, A. (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol Cell.* 21, 811-23

Niggeweg, R., Kocher, T., Gentzel, M., Buscaino, A., Taipale, M., Akhtar, A. & Wilm, M. (2006). A general precursor ion-like scanning mode on quadrupole-TOF instruments compatible with chromatographic separation. *Proteomics* 6, 41-53

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, 6798-810


## 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.


# Functional dynamics of nuclear structure during the cell cycle

## Previous and current research

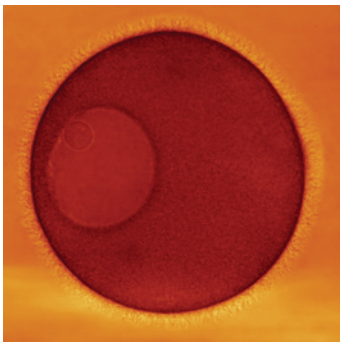
The genome of eukaryotic cells is compartmentalised inside the nucleus, delimited by the nuclear envelope (NE) whose membranes are continuous with the ER and stabilised by the nuclear lamina filament meshwork. The double nuclear membrane is perforated by nuclear pore complexes (NPCs), which allow selective traffic between nucleus and cytoplasm. Attached to the nuclear periphery are heterochromatic regions of chromosomes. These four major structural components of the nucleus are stably connected by many molecular interactions between their resident proteins. In M-phase, most metazoan cells dismantle the highly ordered structure of the NE. 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 mitotic 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 mechanistic models.

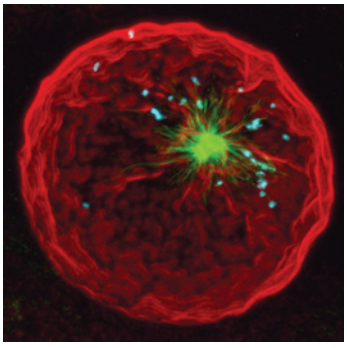
In the past, we showed that NPCs and lamins form a stably interlinked protein network in interphase. We could define the ER as the reservoir and means of partitioning for nuclear membrane proteins in mitosis and found that nuclear breakdown is facilitated by microtubule mediated tearing of the nuclear lamina. During meiotic maturation of starfish 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

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. To rapidly obtain quantitative data from intact cells, we aim to automate and standardise advanced fluorescence microscopy assays as much as possible. This enables us to apply them in higher throughput to all relevant proteins and achieve a systems level understanding of the transformations in nuclear structure during cell division. Because for many of these processes not all required proteins are known, we are also using high-throughput live cell imaging to identify novel genes that function in the above mitotic processes by RNAi screening.



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



*Right: DNA (blue) microtubules (green) and actin (red) in the nucleus of a starfish oocyte entering the first meiotic division.*

## 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* 436, 812-818

Rabut, G., Doye, V. & Ellenberg, J. (2004). Mapping the dynamic organization 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

Lénart, P., Rabut, G., Daigle, N., Hand, A.R., Terasaki, M. & Ellenberg, J. (2003). Nuclear envelope breakdown in starfish oocytes proceeds by partial NPC disassembly followed by a rapidly spreading fenestration of nuclear membranes. *J. Cell Biol.* 160, 1055-1068

## Matthias W. Hentze

MD 1984 University of Münster, Germany.
Postdoctoral training at the NIH, Bethesda, USA.
Group Leader at EMBL since 1989; Senior Scientist since 1998. Associate Director since 2005.
Co-Director of the EMBL/University of Heidelberg Molecular Medicine Partnership Unit since 2002.


# Cytoplasmic gene regulation and molecular medicine

## Previous and current research

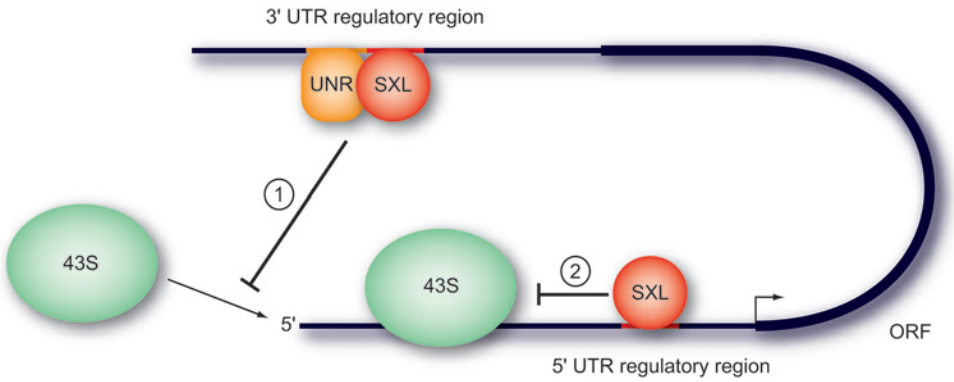
Important steps in the control of gene expression are executed in the cytoplasm: the regulation of mRNA translation and stability. We are elucidating these regulatory mechanisms, including the mode of function of miRNAs, using mostly biochemical approaches and mammalian, yeast and *Drosophila* model 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. We also study the role of miRNAs in cancer and other 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 help elucidate the role of RNA metabolism in disease, and to develop novel diagnostic and therapeutic strategies based on this knowledge.
- 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.




*Fail-safe mechanism regulating msl-2 mRNA translation (see Beckmann et al., 2005).*

## 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* 122, 529-540

Chekulaeva, M., Hentze, M.W. & Ephrussi, A. (2006). Bruno acts as a dual repressor of *oskar* translation, promoting mRNA oligomerization and formation of silencing particles. *Cell* 124, 521-533

Duncan, K., Grskovic, M., Strein, C., Beckmann, K., Niggeweg, R., Abaza, I., Gebauer, F., Wilm, M. & Hentze, M.W. (2006). Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation. *Genes & Dev.* 20, 368-379

Galy, B., Ferring, D., Minana, B., Bell, O., Janser, H.G., Muckenthaler, M., Schümann, K. & Hentze, M.W. (2005). Altered body iron distribution and microcytosis in mice deficient for iron regulatory protein 2 (IRP2). *Blood* 106, 2580-2589

Hentze, M.W., Muckenthaler, M. & Andrews, N.C. (2004). Balancing acts: molecular control of mammalian iron metabolism. *Cell* 117, 285-297


## 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 cellular DNA, protects it from damage and ensures access by the right machines at the right time. The smallest unit of chromatin is the nucleosome, a tightly-knit and stable assembly of histones and DNA, which is nevertheless a dynamic and versatile ("plastic") substrate. We know that it regulates gene expression and the inheritance of our genome. Our lab focuses on discovering novel mechanisms employed by this chromatin-encoded, "epigenetic" process of cellular 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 binds NAD metabolites through its macro domain. Our data suggest that chromosomes may be under the control of NAD metabolites, providing a molecular link between metabolism and human chromatin structure.

## Future project and goals

- Human macroH2A histones as receptors for NAD metabolites.
- Cellular roles of macroH2A histones and mammalian Sir2 deacetylases, which have been linked to obesity, cancer and longevity.
- Identification of heterochromatin effector targeting to fission yeast and human centromeres, including the role of RNAi in transcriptional gene silencing.
- Determination 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


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

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

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

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

## 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 in Tübingen, Germany.
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 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 bind to and modify 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 and trxG proteins using *in vitro* and *in vivo* approaches with the aim to elucidate the molecular mechanisms by which these proteins modify chromatin to generate transcriptional OFF and ON states of target genes.

## Future projects and goals

The regulation of HOX and other target genes in *Drosophila* will continue to provide the 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 study PcG/trxG regulation of novel target genes that we have recently identified by functional genomics approaches.

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


Klymenko, T., Papp, B., Fischle, W., Kocher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M. & Müller, J. (2006). A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev.* 20, 1110-1122

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

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


## 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* (Moody et al., 2005). We have developed chemical methods to convert highly polar signalling molecules like cyclic nucleotides, inositol phosphates and phosphoinositides to membrane-permeant, bioactivatable derivatives (Schultz, 2003). These compounds ("prodrugs" 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. As a new member of the Molecular Medicine Partnership Unit (MMPU) of EMBL and the University of Heidelberg, we are joining forces with Marcus Mall at the Medical School to test these compounds in CF mouse models.

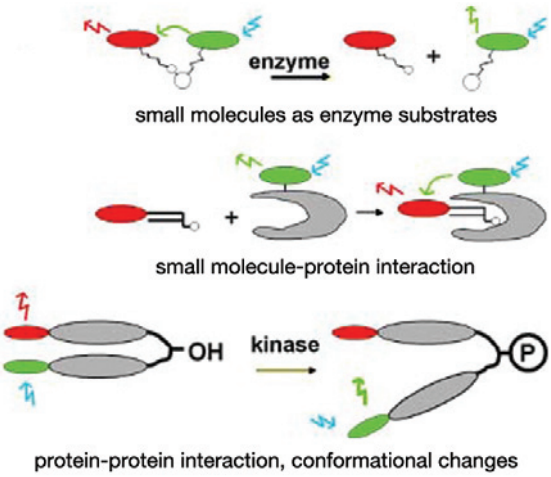
A central goal is to visualise the signalling pathways involved in chloride secretion by fluorescent probes. By doing so, we hope to provide a more complete picture of the signalling network and to create novel ways to screen compounds that might be beneficial for CF patients. The function of the probes is based on fluorescent resonance energy transfer (FRET) and is suitable for ratio imaging and confocal microscopy. The three approaches we pursue are depicted below. They are currently combined in an approach called Multiparameter Imaging, where 5-6 cellular events are monitored simultaneously (Schultz et al., 2005). 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 (Wichmann et al., 2006). 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).

To examine the interaction of phospholipids with lipid-binding proteins we are preparing fluorescently labelled phospholipids that serve as FRET partners for fluorescent fusion proteins. The labelled lipids are membrane-permeant and allow lipid translocation events to be studied. Furthermore, the effect of phospholipids on vesicle trafficking is investigated in collaboration with Rainer Pepperkok.

A third approach is based on conformational changes of double fluorescently labelled proteins that lead to changes in FRET. With these reporter probes we monitor several phosphorylation and aggregation events in living cells (Schleifenbaum et al., 2004; Brumbaugh et al., 2006; Piljic & Schultz, 2006). These projects are performed in collaboration with the groups of E. Conti, M. Sattler and F. Gannon (EMBL) and M. Gautel (London).

In 2005 a new HFSP-funded collaboration started to focus on visualising protein metabolism in living cells. We are currently offering several postdoc positions for candidates with an expertise in carbohydrate or lipid chemistry or in molecular biology and/or imaging.




## Selected references


Brumbaugh, J., Schleifenbaum, A., Gasch, A., Sattler, M. & Schultz, C. (2006). A dual parameter FRET probe for measuring PKC and PKA activity in living cells. *J. Am. Chem. Soc.* 128, 24-25

Schultz, C., Schleifenbaum, A., Goedhart, J. & Gadella T.W., Jr (2005). Multiparameter imaging for the analysis of intracellular signaling. *ChemBioChem.* 6, 1323-1330

Wichmann, O., Wittbrodt, J. & Schultz, C. (2006). A small molecule FRET probe to monitor PLA2 activity in cells and organisms. *Angew Chem. Int. Ed.* 45, 508-512

## 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 in yeast to identify the genetic variants that underlie complex traits and pathways, 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 Hemizygosity 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. We have also begun studies to dissect the genetic basis of sensitivity and resistance to malaria parasites in the mosquito, *Anopheles gambiae*, with a goal to apply the high-throughput methods developed in yeast to this medically relevant system.

## 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, D.G. 2nd, Xiao, W., David, L., Andreoli, C., Monroe, M.E., Moore, R.J., Gritsenko, M.A., Kozany, C., Hixson, K.K., Mottaz, H.M., Zischka, H., Ueffing, M., Herman, Z.S., Davis, R.W., Meitinger, T., Oefner, P.J., Smith, R.D. & Steinmetz, L.M. (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., 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.


# Parallel sequencing as key technology for systems biology and medical applications in proteomics

## Previous and current research

Proteomics, the research to increase our understanding of biological processes by the direct investigation of biological material, is still fueled by the technical development of mass spectrometric analysis techniques. The ultimate goal of proteomic research is to allow a chip-like analysis for any biological sample analysed in a mass spectrometer. This means for close to all proteins their sequence and their abundance should be measured. The total analysis of a proteomic sample will remain unrealistic simply because the protein abundance within a cell spans ten orders of magnitude. However, with the recent introduction of the parallel sequencing approach (http://www.narrador.embl-heidelberg.de/GroupPages/PageLink/activities/ParallelSequ.html) the mass spectrometric analysis can cover a higher percentage of the proteome than ever before. The ability to analyse every sample simultaneously in a quantitative and qualitative way will finally deliver the experimental data to support research in systems biology, quantitative biology and medical diagnosis.

## Future projects and goals

Our current and future activity focuses on realising quantitative proteomic analysis on our mass spectrometers. For this purpose we have installed new HPLCs that allow truly low flow chromatographic separations on 100 nL/min and less, and we have started to develop our own software program to analyse and visualise large chromatographic data sets. The program is based on Apple's Core Data libraries that allow the embedding of mass spectrometric data in an SQL database. This infrastructure is critical to allow multiple threads to access the data simultaneously without programming locks and allows the application to make use of multiple processors and multiple cores in a very symmetric fashion. High computational load is to be expected in proteomic applications, in particular when the parallel sequencing approach is used.

We will develop this application to a state that many chromatographic runs can be easily analysed for the quantitative time course of proteins and apply this technology to various biological and medical research projects.

## Selected references


Niggeweg, R., Kocher, T., Gentzel, M., Buscaino, A., Taipale, M., Akhtar, A. & Wilm, M. (2006). A general precursor ion-like scanning mode on quadrupole-TOF instruments compatible with chromatographic separation. *Proteomics* 6, 41-53

Franz, C., Askjaer, P., Antonin, W., Iglesias, C.L., Haselmann, U., Schelder, M., de Marco, A., Wilm, M., Antony, C. & Mattaj, I.W. (2005). Nup155 regulates nuclear envelope and nuclear pore complex formation in nematodes and vertebrates. *EMBO J.* 24, 3519-3531

Shin, H.W., Hayashi, M., Christoforidis, S., Lacas-Gervais, S., Hoepfner, S., Wenk, M.R., Modregger, J., Uttenweiler-Joseph, S., Wilm, M., Nystuen, A., Frankel, W.N., Solimena, M., De Camilli, P. & Zerial, M. (2005). An enzymatic cascade of Rab5 effectors regulates phosphoinositide turnover in the endocytic pathway. *J. Cell Biol.* 170, 607-618

# Structural and Computational Biology Unit

The Structural and Computational Biology Unit pursues a few common large projects, such as trying to obtain a cell at molecular resolution, that require the input of the different skill sets in the Unit. At the same time it advances science in the different science disciplines it unites.

Currently, the unit consists of twelve groups with another EM group leader hopefully arriving soon. The wide range of methodological experience covers Electron Microscopy (two groups), X-ray (two groups, one team), NMR (one group), a group at the interface between Computational Biology and Structural Biology, a team at the interface between Chemistry and Structural Biology, and Computational Biology (one group, three teams). In addition, two groups based in different units have shared appointments with the SCB unit (the Ladurner group from Gene Expression and the Nédélec team from Cell Biology). Finally, we are temporarily hosting two small visiting/associated teams with external funding, Claudia Muhle-Goll (NMR) and Martin Lercher (cellular modelling), to strengthen areas with little critical mass.

The SCB 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; also facilities for electron cryo-microscopy, cryo-3D tomography crystallography, CD and fluorescence spectroscopy, analytical ultracentrifugation as well as 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.

There is a continuing interplay between the different groups with expertise in 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 and that computational biology is essential to integrate the variety of tools and heterogeneous data.

Thus, beyond the individual research efforts of each group, the SCB groups are currently working together on a project that aims to bridge different biological levels: determining or modelling all protein complexes in a cell and map them to tomograms of entire cell. First proof of principle results have recently been obtained.


## 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.


# Designing biological systems: from proteins to networks

## Previous and current research

Our group is interested in understanding how proteins acquire a 3D structure and the relationship between structure and function. The idea is to infer general rules and develop algorithms that have predictive power and could be used for rational design, as well as to address biological problems. During the last few years we have moved from the analysis of protein folding to the development of computer algorithms that could be used for protein design (Agadir, Perla, FoldX). Currently, we are mainly focused on a particular aspect of the folding problem: aggregation and amyloid formation. We continue to develop our automatic protein design algorithm, FoldX, and have recently modified it so that it can predict protein–DNA interactions. Regarding the misfolding studies and based on the experimental analysis of several hundred peptides (variants of those we previously designed to be amyloidogenic), we have developed two algorithms, Tango and AmyScan, that predict aggregation and amyloid tendency in protein sequences. Currently we are using FoldX in combination with structures of protein complexes to predict protein-protein interaction networks.

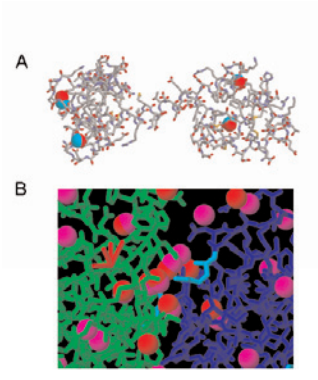
Some years ago we started a new line of research centred on systems biology. We believe that the combination of theoretical and experimental approaches is the most fruitful way to address such a complicated field, and so on the one hand we designed and tested small gene networks while on the other we developed a software tool, SmartCell, that could be used to study the dynamics of biological networks. Recently we also developed a system that allows the engineering and analysis of special networks in a controlled *in vitro* system, and we use it to study patterning. Similarly, we have developed a system to engineer spatial distributions of modified cells to study cell–cell interactions.

## Future projects and goals

In the near future we will like to expand the capabilities of FoldX to be able to predict binding of small molecules to protein. This contributes to the interplay between chemistry and biology to study biological systems. Also, backbone moves will be included which will increase the accuracy and uses of the algorithm. We will also expand the prediction activity on protein–protein and protein–DNA interactions and we will try to make an interface between SmartCell and FoldX so that the SmartCell user wanting to design a network could use the protein design capabilities of FoldX to make new connections, as well as to predict binding constants.

Regarding the misfolding, we will centre all our efforts on the study the toxicity effects of our designed amyloidogenic systems in cell cultures and in organisms like *Drosophila*.

Finally, in systems biology we want to develop SmartCell further by introducing collections of cell tomograms that the user could use to make spatial realistic simulations of cells. New mathematical tools will be implemented – ODEs, pDEs, Boolean, etc. – so that the user could choose, in the same graphic interface, a suit of different algorithms depending on the problem to be analysed. Also, we would like to create a module that could be used to design new networks that will be connected to FoldX and to the parts repository of the MIT. Finally, we will connect FoldX to the major databases of protein and gene networks and we would aim to simulate a full small cell (around 600 genes) in the code and re-engineer it based on the simulations and our design tools.




*Prediction of metal binding sites and water bridges (Water molecules that make two or more H-bonds to the protein) in proteins. A) X-tal structure of calmodulin showing the 4 calcium ions (cyan) and the predicted positions by FoldX (red). B) X-tal structure of the complex between Ras and a RBD domain showing the water molecules in the interphase of the complex. Red: Predicted water bridges. Magenta: Crystal waters. There are more magenta molecules because some of the waters in the complex are only making one H-bond or none.*

## Selected references


Beltrao, P. & Serrano, L. (2005). Comparative Genomics and disorder prediction identify biologically relevant SH3 protein interactions. *PLoS Comput. Biol.* 1, e26

Esteras-Chopo, A., Serrano, L. & Lopez de la Paz, M. (2005). The amyloid stretch hypothesis: recruiting proteins toward the dark side. *Proc. Natl. Acad. Sci. USA* 102, 16672-16677

Isalan, M., Lemerle, C. & Serrano, L. (2005). Engineering gene networks to emulate *Drosophila* embryonic pattern formation. *PLoS Biol.* 3, e64

Schymkowitz, J.W., Rousseau, F., Martins, I.C., Ferkinghoff-Borg, J., Stricher, F. & Serrano, L. (2005). Prediction of water and metal binding sites and their affinities by using the Fold-X force field. *Proc. Natl. Acad. Sci. USA* 102, 10147-10152

## Peer Bork

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


# Deciphering function and evolution of biological systems

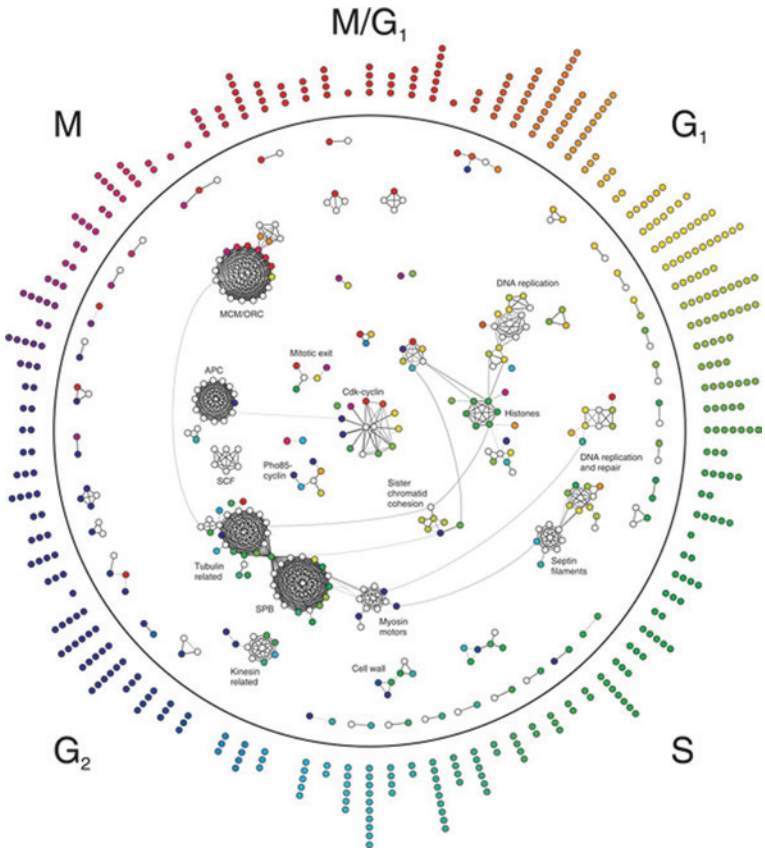
## Previous and current research

The main focus of the Bork Group – Computational Biology – is to predict function and to gain insights into evolution by comparative analysis of complex molecular data. The group currently works on three different scales:

- genes and proteins;
- protein networks and cellular processes;
- phenotypes and environments.

They require both tool development and applications. Some selected projects include comparative gene, genome and metagenome analysis, mapping interactions to proteins and pathways as well as the study of temporal and spatial protein network aspects. All are geared towards the bridging of genotype and phenotype through a better understanding of molecular and cellular processes.

The group is partially associated with the Max Delbrück Center for Molecular Medicine (MDC) in Berlin.




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

## Selected references


Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B. & Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283-7

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

Jensen, L.J., Saric, J. & Bork, P. (2006). Literature mining for the biologist: from information retrieval to biological discovery. *Nat Rev Genet.* 7, 119-29

de Lichtenberg, U., Jensen, L.J., Brunak, S. & Bork, P. (2005). Dynamic complex formation during the yeast cell cycle. *Science* 30, 724-7

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.Comparative metagenomics of microbial communities. *Science* 308, 554-7


## Bettina Böttcher

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


# Understanding the structural organisation of biological complexes

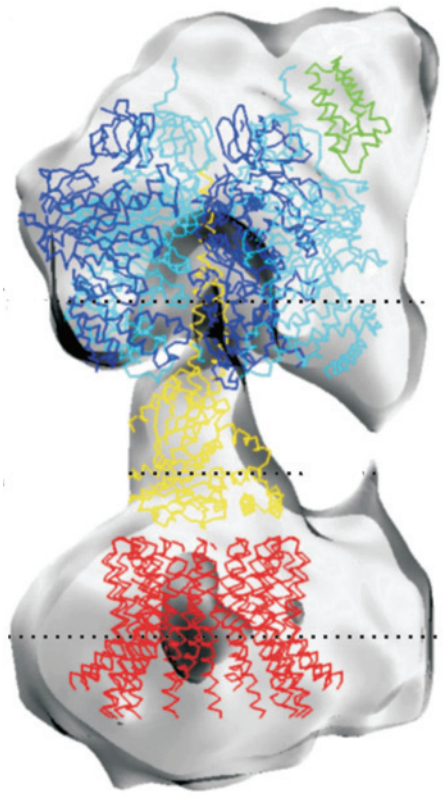
## Previous and current research

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 the the EU Integrated Project 3D Repertoire. 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


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

Kronenberg, S., Böttcher, B., von der Lieth, C.W., Bleker, S. & Kleinschmidt, J.A. (2005). A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. *J. Virol.* 79, 5296-5303

Mellwig, C. & Bottcher, B. (2003). A unique resting position of the ATP-synthase from chloroplasts. *J. Biol. Chem.* 278, 18544-18549

Venzke, D., Domgall, I., Kocher, T., Fethiere, J., Fischer, S. & Böttcher, B. (2005). Elucidation of the stator organization in the V-ATPase of *Neurospora crassa*. *J. Mol. Biol.* 349, 659-669

## 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.


# Structural studies of nucleo-cytoplasmic transport and mRNA metabolism

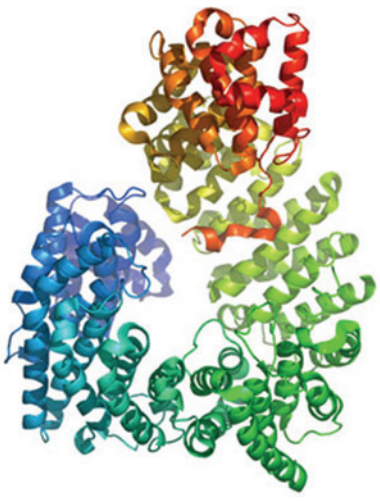
## Previous and current research

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. In eukaryotes this process allows communication to occur between cytosolic translation and nuclear replication and transcription. 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. 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. The bulk of nuclear import/export is mediated by the karyopherin family of proteins (for example Cse1). These proteins can carry both proteins and RNA cargoes such as tRNAs or pre-miRNAs. The transport of mRNA is instead mediated by a different system unrelated to the karyopherins. Nuclear export of mRNAs is connected to a complex series of events ensuring that only correct and properly spliced mRNAs are translated into proteins. The exon-exon junction complex (EJC) plays a crucial role in this context. In humans, the EJC is deposited on mRNA upon splicing and functions together with SMG proteins to mediate nonsense-mediated mRNA decay (NMD). NMD is a surveillance mechanism that recognises mRNAs with premature stop codons and targets them for rapid degradation, thus avoiding the synthesis of truncated and potentially harmful protein products. Degradation of nonsense transcripts is mediated by the same cellular machinery that participates in general mRNA turnover. In particular, degradation of mRNAs in the 3' to 5' direction is catalysed by the exosome, a protein complex that is also involved in the maturation of structural RNAs in the nucleus.

## Future projects and goals

Our goal is to understand the mechanisms by 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 and targeting (i.e. the exon junction complex and the SMG proteins) 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 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.*

## Selected references


Conti, E. & Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* 17, 316-325

Cook, A., Fernandez, 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

Lorentzen, E. & Conti, E. (2005). Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol. Cell* 20, 473-481


## 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

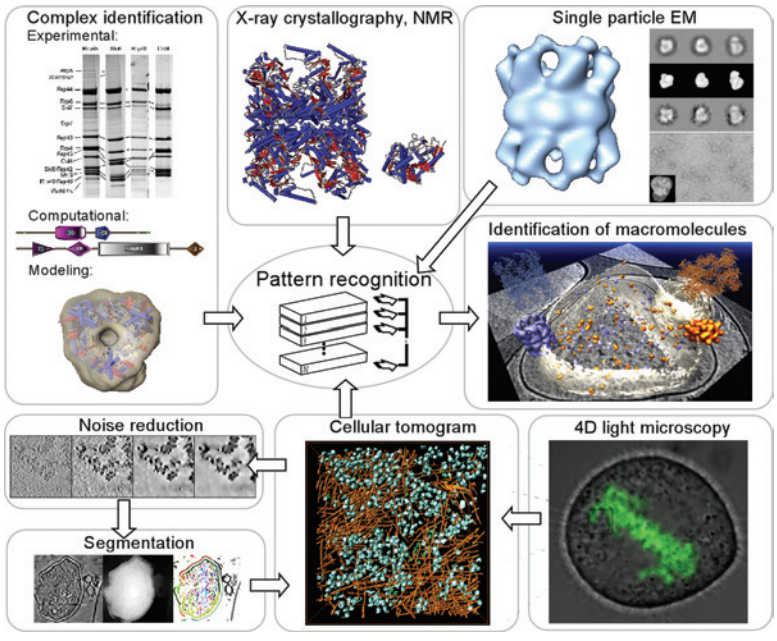
## Previous and current research

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


Frangakis, A.S. & Forster, F. (2004). Computational exploration of structural information from cryo-electron tomograms. *Curr. Opin. Struct. Biol.* 14, 325-331

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

Kurner, J., Frangakis, A.S. & Baumeister, W. (2005). Cryo-electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. *Science* 307, 436-438

Medalia, O., Weber, I., Frangakis, A.S., Nicastro, D., Gerisch, G. & Baumeister, W. (2002). Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science* 298, 1209-1213

## Anne-Claude Gavin

PhD 1992, University of Geneva, Switzerland.
Postdoctoral research at EMBL.
Director, Molecular and Cell Biology, Cellzome AG, Heidelberg.
Team Leader at EMBL since 2005.


# Biochemical and chemical approaches to biomolecular networks

## Previous and current research

How is biological matter organised? Can the protein and chemical worlds be matched to understand the cell's inner works? As our knowledge on the basic building blocks of eukaryotic genomes, proteomes and metabolomes grows, the challenge resides in understanding how these parts relate to each other. At cellular levels, gene products very rarely act alone. The orchestration of complex biological functions is the result of networks of molecules. Traditional approaches have typically focused on a few, selected gene products and their interactions in a particular physiological context. We are proponents and pioneers of more general strategies aiming at understanding complex biological systems. We are interested in understanding the principles that govern the assembly of these networks and follow three main lines of research.

**The charting of protein-protein interaction networks:** Our knowledge on protein-protein interaction is still anecdotic; current estimations reveal that probably less than 10% have been characterised so far. We adopted the Tandem-Affinity purification/Mass Spectrometry (TAP/MS) technology to perform a genome-wide analysis of protein complexes in the yeast *S. cerevisiae*. More than 400 different protein complexes, more than half entirely novel, were characterised. The approach was particularly successful in further extensive collaborations within the program. These collaborations aimed at the structural characterisation of protein complexes through integration of electron microscopy data and *in silico* approximations.

**The study of protein complexes and network order of assembly and dynamics:** Generally, the use of protein interaction networks to predict the behaviour of whole systems has been relatively limited. Protein networks usually fail to capture the dynamic aspect of protein interactions that is essential for the whole cell functioning. The charting and modelling of the highly dynamic assembly and reorganisation of protein complexes following cell perturbation represents one of the major current research interests of the group.

**The extension of interaction networks from proteins to other cell's building blocks; metabolites-on-proteomes networks:** Metabolites account for about half of the cell's volume and represent important class of biomolecules. They have long been considered simple building blocks for the assembly of more complex macromolecules. It is however becoming evident that the interactions between the metabolites' and the proteins' worlds are not limited to substrate/product relationships. Metabolites can have well known signalling functions and many proteins are allosterically modulated by metabolites. These bindings are sometimes mediated by a variety of specialised domains. Every time it has been possible to chart such interactions they turned out to have profound functional implications. The interactions taking place between the cell's chemical world and proteomes are still poorly defined and have certainly not yet been studied in a comprehensive way. It represents the second major research interest of the group.

## Future projects and goals

- Analysis of the order of assembly and dynamic nature of yeast protein complexes, in a pathway-oriented approach.
- Further development and improvement of existing chemical biology methods, based on affinity purification ("metabolite pull-down") to monitor protein-metabolites interaction.
- Global screen aiming at the systematic charting of the interactions taking place between the proteome and the metabolome in the model organism *Saccharomyces cerevisiae*.
- Develop new and existing collaborations with computational and structural biology groups at EMBL and outside to tackle the structural and functional aspects of biomolecular recognition.

## Selected references


Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M.A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A.M., Schelder, M.,Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B. & Superti-Furga, G. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631-6

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, 2026-9

Gavin, A.C. & Superti-Furga, G. (2003). Protein complexes and proteome organization from yeast to man. *Curr. Opin. Chem. Biol.* 7, 21-7


## 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 protein 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.

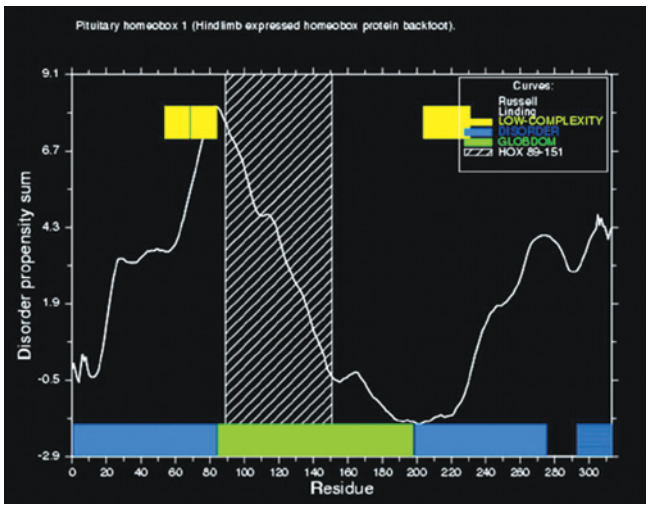
We are probably best known for our involvement with the Clustal W and Clustal X programs that are widely used for multiple sequence alignment. We work closely with Julie Thompson (Strasbourg) and Des Higgins (Dublin) to maintain and develop these programs. We also maintain several public web servers 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 developed the 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 α2-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 may 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 α2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? *Genome Biol.* 5, R38

Diella, F., Cameron, S., Gemund, C., Linding, R., Via, A., Kuster, B., Sicheritz-Ponten, T., Blom, N. & Gibson, T.J. (2004). Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. BMC *Bioinformatics* 5, 79

Linding, R., Russell, R.B., Neduva, V. & Gibson, T.J. (2003). GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* 31, 3701-3708

Puntervoll, P., Linding, R., Gemund, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D.M., Ausiello, G., Brannetti, B., Costantini, A., Ferre, F., Maselli, V., Via, A., Cesareni, G., Diella, F., Superti-Furga, G., Wyrwicz, L., Ramu, C., McGuigan, C., Gudavalli, R., Letunic, I., Bork, P., Rychlewski, L., Kuster, B., Helmer-Citterich, M., Hunter, W.N., Aasland, R. & Gibson, T.J. (2003). ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res.* 31, 3625-3630

## Rob Russell

PhD 1993, Oxford University, UK.
Postdoctoral research at Imperial Cancer Resarch 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 research is focused broadly on the relationships between protein three-dimensional (3D) structure, function and evolution. A major interest over the past four years has been to develop methods to understand and predict the molecular details for macromolecular interactions. These efforts can be currently classified into three main areas.

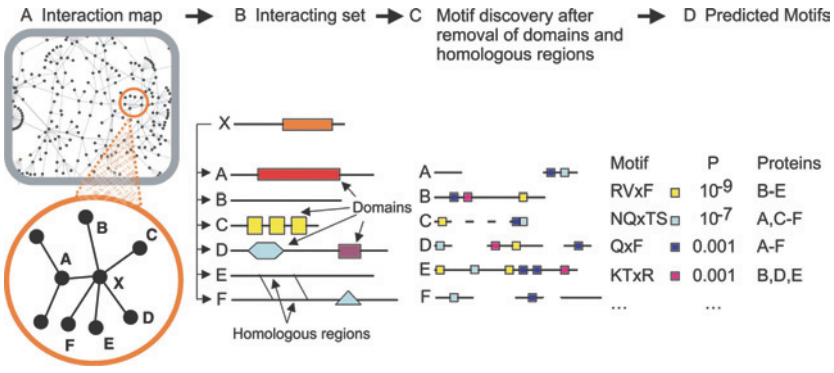
**Protein interactions and complexes:** Protein interaction networks are central to any understanding of cellular processes, and though many thousands are now known, few initiatives to uncover them pay much attention to one of the best sources of data available: complexes of known 3D structure. We thus 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.

**Protein and DNA sequence motifs mediating interactions:** 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. We are also actively involved in experimental efforts to test or validate our predictions, either doing them ourselves or working in close collaboration with experimental groups.

**Chemical biology: linking chemicals to pathways:** We are now actively developing a database of chemicals linked to proteins in order to interrogate the likely effects of disrupting pathways with chemicals. This will be complemented by new and existing methods for the design of small-molecule or peptide inhibitors with the aim of working closely with experimentalists.

## Future projects and goals

- To develop the "one-stop-shop" for structural information on currently available information (predicted or experimental) on large protein complexes. This will involve development and application of new methods for complex structure prediction.
- To devise techniques to predict peptide binding sites on protein surfaces to complement our predictions of interacting peptides.
- To derive methods to predict the outcome of chemical treatment on particular pathways by cross-referencing chemicals with pathway components.




*Schematic outlining our approach for finding protein linear motifs that mediate protein-protein interactions. Sets of proteins (A-F) sharing an interaction partner (X) are grouped and domains and homologous sequences are removed. We then search for 3-8 residue motifs that are over-represented in the remaining sequence, and score these by a binomial probability to give a ranked set of candidate motifs mediated the interaction with protein X.*

## Selected references


Aloy, P. & Russell, R.B. (2006). Structural systems biology: modelling protein interactions. *Nat Rev Mol Cell Biol.* 7, 188-97

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M.A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A.M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B. & Superti-Furga, G. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631-6

Neduva, V., Linding, R., Su-Angrand, I., Stark, A., de Masi, F., Gibson, T.J., Lewis, J., Serrano, L. & Russell, R.B. (2005). Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol.* 3, e405

Stark, A., Brennecke, J., Bushati, N., Russell, R.B. & Cohen, S.M. (2005). Animal microRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123, 1133-46


## 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

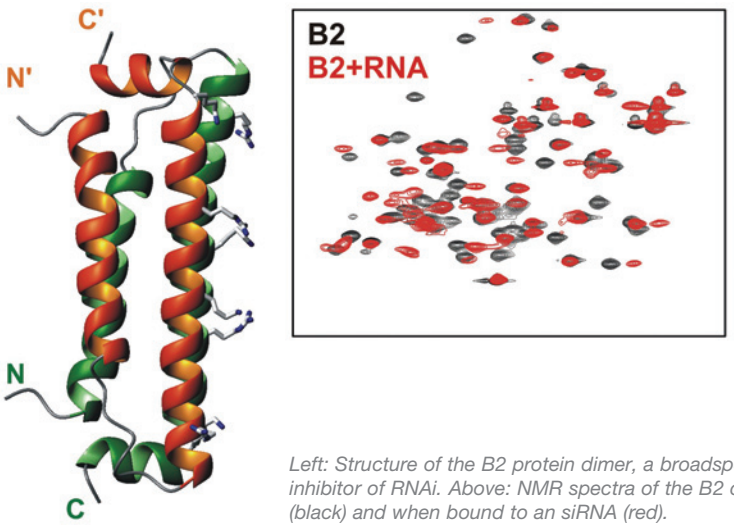
We study the structure, molecular recognition and dynamics of biomolecules in solution. Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique not only for determining 3D structures but also for the characterisation of interactions and dynamics of biological molecules in solution. Recent advances in NMR methodology and instrumentation allow multi-domain proteins and complexes of 100 kDa molecular weight and higher to be studied.

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. Larger proteins are modular and consist of structural domains of 50-150 amino acids in size, which can be efficiently studied by NMR techniques. NMR methods can also be used to characterise the structure of multidomain proteins and changes in the domain orientation and dynamics, for example, upon ligand binding or phosphorylation. We are currently studying domains and multidomain proteins which function in the regulation of gene expression and signal transduction.

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, including (alternative) splicing and RNA interference. The spliceosome is a highly dynamic machinery, which involves numerous protein-RNA interactions. 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 study the structures and molecular interactions of biomolecules in solution to understand the molecular basis of their biological function. A main focus of our research lies in protein-protein and protein-RNA interactions with important roles in RNA metabolism. We are optimising NMR methods to characterise the structure and dynamics of larger multi-domain protein-protein and protein-RNA complexes.




*Left: Structure of the B2 protein dimer, a broadspectrum inhibitor of RNAi. Above: NMR spectra of the B2 dimer alone (black) and when bound to an siRNA (red).*

## Selected references


Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. (2005). The structure of the flock house virus B2 protein, a viral suppressor of RNA interference, shows a novel mode of double-stranded RNA recognition. *EMBO Rep.* 6, 1149-1155

Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat. Struct. Mol. Biol.* 11, 576-577

Liu, Z., Luyten, I., Bottomley, M.J., Messias, A.C., Houngninou-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

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

## Klaus Scheffzek

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


# Signal transduction – disease proteins

## Previous and current research

Defects in signalling pathways are often associated with the occurrence of severe diseases, with cancer a common example. We are interested in understanding the mechanisms of pathogenesis associated with cancer related diseases. Previously, we have characterised the regulation of Ras – a GTP binding protein mutated in 30% of human tumours – and the related Rho proteins. Ras 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.

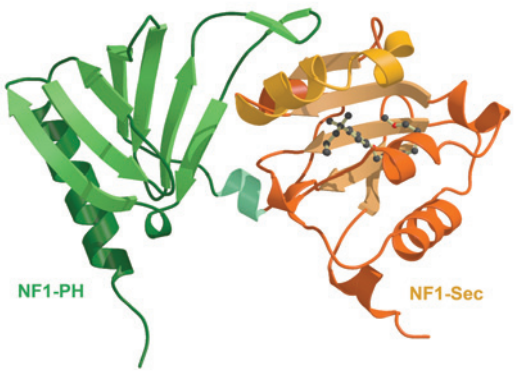
Currently our major focus is on neurofibromatosis type 1 (NF1), a genetic disease with an incidence of 1 in 3500 worldwide. NF1 patients have an increased risk of developing certain types of tumours and frequently show learning disabilities. The NF1 gene encodes a huge protein of about 2,800 amino acids, termed neurofibromin, and when mutated 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. Within ten years of research after 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 or by bound ligands obtain ideas for functional/biochemical experiments. Work on this project offers the opportunity to contribute to a challenging and physiologically exciting research topic. Our main technique is X-ray crystallography but other methods are increasingly employed. Using the approach described above we have recently discovered a novel bipartite module containing a lipid binding Sec14-homology and a previously undetected pleckstrin homology (PH) like domain the function of which we are currently investigating. Our approach has 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 3D 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. We are increasingly including automated strategies to identify soluble protein fragments that are accessible to biochemical/structural analysis. In addition we will continue searching for interaction partners of neurofibromin and study their role in the protein function.

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




*Structure of a bipartite module from neurofibromin composed of a lipid binding Sec14 homologous and a pleckstrin homology-like domain.*

## Selected references


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

D'Angelo, I., Welti, S., Bonneau, F. & Scheffzek, K. (2006). A novel bipartite phospholipid-binding module in the neurofibromatosis type 1 protein. *EMBO Rep.* 7, 174-179

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

Reynaud, E.G., Andrade, M.A., Bonneau, F., Ly, T.B., Knop, M., Scheffzek, K. & Pepperkok, R. (2005). Human Lsg1 defines a family of essential GTPases that correlates with the evolution of compartmentalization. *BMC Biol.* 3, 21


## 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

Nowadays 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 such that it can be 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 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 developing data-mining tools (e.g. text-mining);
- visualisation of heterogeneous data mining results.

## Future projects and goals

Our goal is to develop a knowledge platform for the life sciences. We will first focus on the biology-driven research areas, but will also extend into chemistry-related fields, primarily 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). A bioinformatics perspective on proteomics: data storage, analysis, and integration. *Biosci. Rep.* 25, 95-106

Ofran, Y., Punta, M., Schneider, R. & Rost, B. (2005). Beyond annotation transfer by homology: novel protein-function prediction methods to assist drug discovery. *Drug Discov. Today* 10, 1475-1482

## 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 widely varying specificities, the formation of multi-component protein-nucleic acid complexes and the catalytic mechanisms involved in nucleic acid modifications.

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 DNA 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). Currently we are working on HJ resolvases from phage T4 endonuclease VII (Endo VII), whose intrinsic conformational flexibility is thought to be of functional significance for its broad substrate specificity, archaeal Hjc's and yeast Cce1, which in contrast to Endo VII do specifically recognise and cleave 4-way DNA junctions. Crystals of several HJ complexes have been obtained.

A major focus of our current work are the multi-component Arc1p and Ccr4-Not complexes from *S. cerevisiae*. Arc1p forms a stable complex with Met and Glu aminoacyl-tRNA synthetases (MetRS, GluRS) and stabilises their interaction with the cognate tRNAs. The yeast Arc1p complex functionally links tRNA nuclear export and aminoacylation and represents an evolutionary intermediate of the mammalian, multi-subunit aminoacyl-tRNA synthetase complex. The X-ray structure of a complex between the interacting N-terminal domains of Arc1p and GluRS reveals, that both domains adopt a GST-like fold, but that their interaction is completely different from that found in classical GST homodimers. Analysis of the complex interface and sequence comparisons suggests a structural model for the heteromerisation between different components of the tRNA channeling cycle.

The Ccr4-Not core complex, a global gene expression regulator playing a major role in mRNA deadenylation, has a MW of ~1MDa and consists of 9 proteins. It contains two nucleases, Pop2 (also known as Caf1) and Ccr4, both of which are essential for deadenylation *in vivo*. We have solved the crystal structure of the Pop2p non-canonical nulease domain revealing its structural homology with the 3'-5'exonuclease domains of *E. coli* DNA polymerases I and III. Several constructs of the Ccr4p, Pop2p and Not1p proteins are currently being expressed/co-expressed to study their interactions.

Another project deals with Sm-like proteins from archaea and bacteria. Our X-ray structures of the *E. coli* Hfq protein and RNA complexes of Sm-related proteins from Archaea (*A. fulgidus*, *P. abyssi* and *S. solfataricus*) have provided a high resolution picture of RNA-binding in an Sm core domain and revealed the close evolutionary relationship to eukaryotic snRNPs. Based on immunoprecipitation experiments we proposed a possible role of the *A. fulgidus* Sm proteins in tRNA maturation. To learn more about the function of archaeal Sm proteins, we are trying to identify their *in vivo* partners by expressing tagged proteins in *S. solfataricus*. *E. coli* Hfq acts as an RNA chaperone and post-transcriptional regulator by promoting base-pairing between regulatory RNAs and their mRNA targets. To understand the underlying mechanisms, we have started a structural investigation of several binary and ternary complexes of Hfq with full length and truncated RNA targets.

## Future projects and goals

Future projects include the X-ray structure determinations of substrate complexes of Endo VII and other resolvases, of the multi-component Arc1p-tRNA synthetase and Ccr4-Not complexes and structure-function studies of Sm-related proteins.

By determining the crystal structures of these complexes, we want to reveal the structural basis for their formation, their selectivity and function.

## Selected references


Kilic, T., Thore, S. & Suck, D. (2005). Crystal structure of an archaeal Sm protein from *Sulfolobus solfataricus*. *Proteins* 61, 689-693

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., Mauxion, F., Seraphin, 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

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 specific 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 estrogen. The physiological effects of estrogen are transduced through specific nuclear proteins, the Estrogen Receptors, which are estrogen-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 destabilising the ER-α messenger RNA. More recently, they have comprehensively portrayed the sequence of events, instigated by Estrogen Receptor, that achieve and then limit transcription of estrogen responsive promoters. Ultimately, the group's aim is to understand how estrogen 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.

## Iain Mattaj

PhD 1979, University of Leeds, UK.
Postdoctoral work at the Friedrich Miescher Institute and the Biocenter, Basel, Switzerland.
Group Leader at EMBL since 1985. Programme Coordinator since 1990.
Scientific Director of EMBL 1999–2005. Director General since 2005.


# The RanGTPase as a spatial regulator

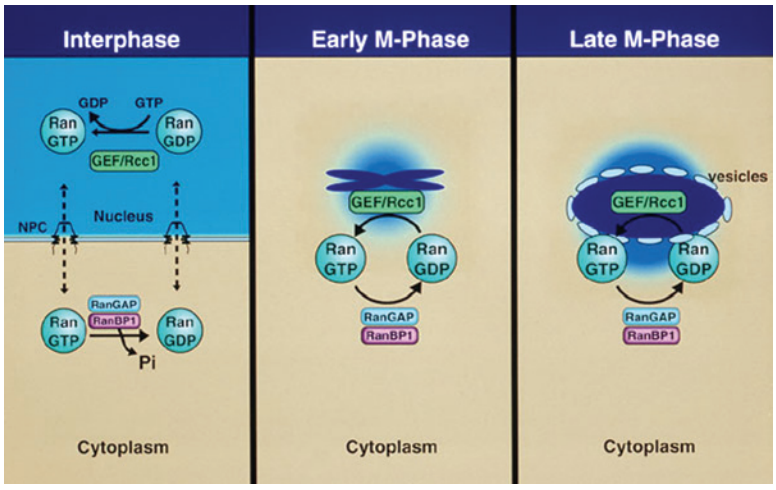
## Previous and current research

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 production 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

Hetzer, M.W., Walther, T.C. & Mattaj, I.W. (2005). Pushing the envelope: structure, function, and dynamics of the nuclear periphery. Annu. Rev. *Cell Dev. Biol.* 21, 347-380

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


## 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 of EMBO; Secretary General of EMBC. At EMBL since 1994.


# The expression and functional regulation of Estrogen Receptor-α

## Previous and current research

A major focus of my laboratory's work is the control of the expression of genes by the Estrogen 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 estrogen 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 estrogen 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 the Cell Biology and Biophysics, Gene Expression and Structural and Computational Biology Units, as well as with several of the core facilities, particularly with the Genomics and Chemical Biology facilities.

## Selected references


Metivier, R., Reid, G. & Gannon, F. (2006). Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Rep.* 7, 161-167

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-estrogen 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). Estrogen receptor-α directs ordered, cyclical and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 115, 751-763

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 estrogen receptor-α, in response to deacetylase inhibition by valproic acid and trichostatin A. *Oncogene* 24, 4894-4907

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 estrogen signalling. *Mol. Cell* 11, 695-707

# 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, work 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.


## 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

Graf, R., Rietdorf, J. & Zimmermann, T. (2005). Live cell spinning disk microscopy. Adv. *Biochem. Eng. Biotechnol.* 95, 57-75

Moita, L.F., Wang-Sattler, R., Michel, K., Zimmermann, T., Blandin, S., Levashina, E.A. & Kafatos, F.C. (2005). *In vivo* identification of novel regulators and conserved pathways of phagocytosis in *A. gambiae*. *Immunity* 23, 65-73

Zimmermann, T. (2005). Spectral imaging and linear unmixing in light microscopy. Adv. *Biochem. Eng. Biotechnol.* 95, 245-265

## 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

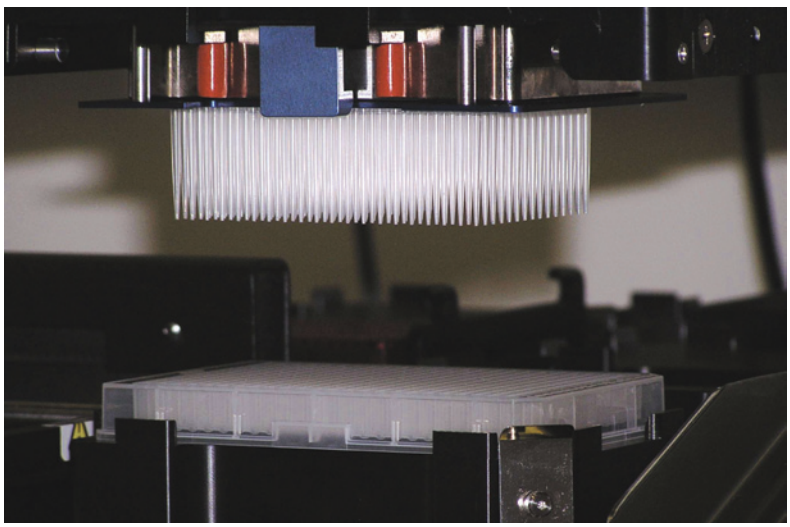
The facility was established at the beginning of 2004. Through a large team effort, it was fully functional in August 2004. We have a very strong pipeline of projects from EMBL and DKFZ, several of which have completed screening and are now in the lead optimisation phase.

## 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.*

## Selected references


Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J. D. & and Stark, H. (2005). Structure of the hepatitis C virus IRES bound to the human 80S ribosome: Remodelling of the HCV IHRES. *Structure* 13, 1695-1706

Neduva, V., Linding, R., Su-Angrand, I., Stark, A., de Masi, F., Gibson, T.J., Lewis, J., Serrano, L. & Russell, R.B. (2005). Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol.* 3, e405

Ostareck-Lederer, A., Clauder-Münster, S., Thermann, R., Polycarpou-Schwartz, M., Gentzel, M., Wilm, M. & Lewis, J.D. (2005). The role of RNA interference in drug target validation: Application to Hepatitis C. In "RNA Interference Technology", Appasani, K. (ed.), Cambridge University Press, 318-330


## Claude Antony

PhD 1984, Université 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. EM Core Facility techniques contribute to various projects across the units, and the facility also trains new users to make best use of our advanced equipment. We also develop new approaches and methods in EM application to cellular and developmental biology.

## Major projects and accomplishments

- 3D reconstruction of the whole *E. coli* bacteria volume calculated from original serial sections that we acquired on the EM scope. We have used the same approach with a yeast cell, *Saccharomyces cerevisiae*, and obtained a 3D volume model of a full cell and its various subcellular compartments (with Raul Gomez, Serrano lab).
- The study concerned the nuclear envelope and pore complex assembly in an *in vitro* system based on *Xenopus* egg extracts using the expression of mutant proteins. The EM work allows us to assess the exact contribution of each protein tested (with Wolfram Antonin, Mattaj lab).
- The study of alterations of the heart muscle induced by the expression of a mutant form of human troponin in rat hearts. The human mutated troponinT, which conveys a high risk of heart attack, is best analysed this way in detail (including the EM approach) in transgenic animal systems. We detected and measured a significant reduction of the mitochondriome volume, as well as an increase in muscle fibre striation with quite disorganised cellular contents (collaboration with Mark Luedde, Medizinische Klinik der Universität, Heidelberg).

## Services provided

- Providing up-to-date know-how 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

**FEI Company:** supplier of advanced electron microscopes. They have been our partner for many years for our EM scopes (previously Philips) to be maintained and fixed.

**Leica-microsystems:** we recently acquired the newly released Leica HPFreezer EMPACT2. The EMPACT2 is a portable machine and has an optional attachment, the Rapid Transfer System (RTS), which permits easy and reproducible loading of the sample and allows correlative light and electron microscopy with high time resolution. Up to now this machine gave very satisfactory results both for cell biology-EM (freeze substituted and plastic embedded samples) as well as when used for frozen hydrated specimen followed by cryosectioning (performed in the SBC Unit).

**Soft Imaging System:** we have acquired a new CCD camera (KeenView) which is set on the Philips Biotwin CM120 EM scope. The quality of the images is very good, and we no longer need to use negatives and prints for posters and publications.

## 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

Carazo-Salas, R.E., Antony, C. & Nurse, P. (2005). The kinesin Klp2 mediates polarization of interphase microtubules in fission yeast. *Science* 309, 297-300

Franz, C., Askjaer, P., Antonin, W., Iglesias, C.L., Haselmann, U., Schelder, M., de Marco, A., Wilm, M., Antony, C. & Mattaj, I.W. (2005). Nup155 regulates nuclear envelope and nuclear pore complex formation in nematodes and vertebrates. *EMBO J.* 24, 3519-3531

Tanaka, K., Mukae, N., Dewar, H., van Breugel, M., James, E.K., Prescott, A.R., Antony, C. & Tanaka, T.U. (2005). Molecular mechanisms of kinetochore capture by spindle microtubules. *Nature* 434, 987-994

## Andrew Riddell

BSc Hons 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 Core 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

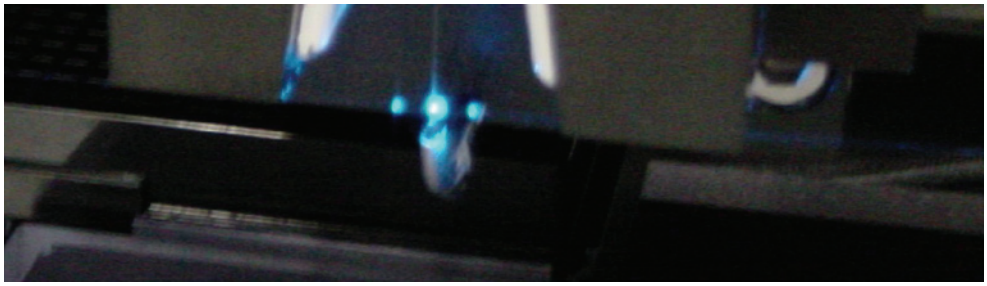
- The analysis of algae life cycle project relied on the intrinsic fluorophores in algae to identify life cycle stages. It required photosaturation of the photosynthetic units in the algae.
- Cell cloning by Darwinian selection required a series of single cell sorts of a target population into 96-well plate in order to select a stable integrated gene into a cell line.
- There is a project investigating a bi-stable state of a reworked bacterial signalling cascades that requires precise and accurate instrument measurements of the bacteria in order for them to be identified.
- The apoptosis project utilises a novel FRET assay in order to identify apoptosis in a cell line.
- An ongoing project, in collaboration with the University of Heidelberg's Chemistry Department, investigates the flow cytometric analysis of cellular uptake of novel synthetically produced probes.

## 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 the EMBL's scientific activities.

## Technology partners

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



*Deflection illumination for calibrating droplet break-off point.*

## Selected references


Corcoran, A.E., Riddell, A., Krooshoop, D. & Venkitaraman, A.R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* 391, 904-907

Fairley, E.A., Riddell, A., Ellis, J.A. & Kendrick-Jones, J. (2002). The cell cycle dependent mislocalisation of emerin may contribute to the Emery-Dreifuss muscular dystrophy phenotype. *J. Cell Sci.* 115, 341-354

Füssl, A., Schleifenbaum, A., Goritz, M., Riddell, A., Schultz, C., Kramer, R. (2006). Cellular uptake of PNA-terpyridine conjugates and its enhancement by Zn2+ ions. J. Am. Chem. Soc., 128, 5986-5987

Goyenechea, B., Klix, N., Yelamos, J., Williams, G.T., Riddell, A., Neuberger, M.S. & Milstein, C. (1997). Cells strongly expressing Ig(κ) transgenes show clonal recruitment of hypermutation: a role for both MAR and the enhancers. *EMBO J.* 16, 3987-3994


## Vladimír Beneš

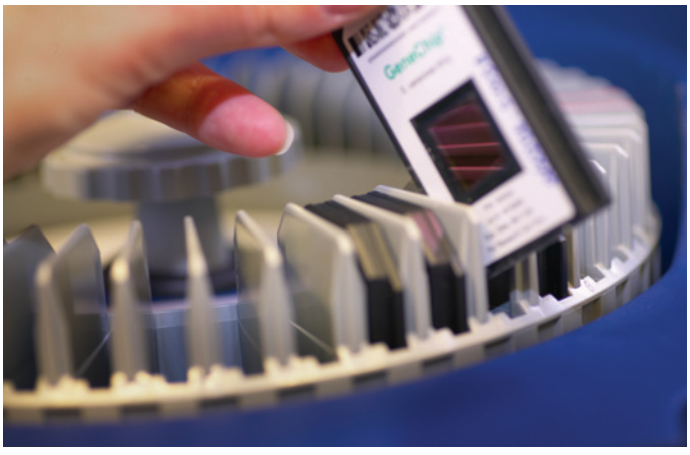
PhD 1994, Czech Academy of Sciences,
Prague, Czech Republic.
Posdoctoral research at EMBL.
Facility Head since 2001.


# 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 35,000 reactions during 2005 on samples from almost all the "wet lab" groups in Heidelberg and Grenoble and some from Hamburg. 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*.



*High density microarrays are a key component of the GeneCore gene expression profiling services.*

## Selected references


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

Raible, F., Tessmar-Raible, K., Osoegawa, K., Wincker, P., Jubin, C., Balavoine, G., Ferrier, D., Benes, V., de Jong, P., Weissenbach, J., Bork, P. & Arendt, D. (2005). Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* 310, 1325-1326

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 estrogen receptor-α, in response to deacetylase inhibition by valproic acid and trichostatin A. *Oncogene* 24, 4894-4907

Relogio, 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

## 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 facility has an ongoing 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 facility's technologies.

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

## Selected references


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. *J. Cell Biol.* 127, 1289-1299 (published erratum appears in *J. Cell Biol.* 128, following 988)

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., Massimi, M., Bullard, B., Ansorge, W. & Sawyer, A. (2005). High throughput production of mouse monoclonal antibodies using antigen microarrays. *Proteomics* 5, 4070-4081


## Ario de Marco (until March 2006)

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.


# The development of the services offered by the Protein Expression and Purification Core Facility

Core Facilities provide the scientific community with services that speed up the achievement of final results and render them more reliable, as technically demanding experiments are performed by specialists. In this context there is a bidirectional interaction between research groups and facility members, who provide updated technology while also promoting attention to neglected but relevant aspects of protein manipulation.

During the past year the Protein Expression and Purification Core Facility has tried to exceed users' expectations – by reactivating a biophysical analysis service and making available a yeast fermentation service, for instance – to anticipate the interest for tools and protocols in the field of antibody technology. It also improved knowledge of underestimated biotechnological processes, like the formation of soluble aggregates or expression at very low temperatures.

## Major projects and accomplishments

The facility was involved in few development projects. The complexity of the protein aggregation during the recombinant expression has been studied using innovative protocols and the use of a 3D-geometry gel. Such a device has also been successfully used for parallel SDS-PAGE separation and analysis of 386 protein samples.

The possibility of using chemical chaperones, alone or in combination with molecular chaperones, to improve the protein solubility has been investigated and molecular reporters developed to follow *in vivo* the effect of the expression parameters on the host cell physiology. Cold adapted bacteria have been tested in comparison with *E. coli* to challenge the folding efficiency of recombinant proteins expressed at 4°C.

We developed protocols for immortalising polyclonal antibodies, for selecting recombinant antibody expressing cells using permeable membranes, for separating single lines of polyclonals, identifying single epitopes in the antigens, and rescuing binders from recombinant antibody libraries.

## Services provided

All the facility members are committed to offering their suggestions and explanations for optimising the work of the EMBL researchers and of any external scientist who asks for advice. We also distribute, inside the Laboratory and externally, several hundreds of our home-developed vectors and maintain the insect and bacteria strain collections. Furthermore, we produce proteins of general use, like Pfu, Taq and T7 polymerases, TEV and 3C proteases, LIF and, from this year, Cre recombinase.

The screening for the identification of optimal expression conditions for the production of recombinant soluble proteins from bacteria and the large scale expression in *E. coli* and insect cells hasbeen integrated with the production of yeast in a 42L fermenter, even though this latter service is, for the moment, restricted to the members of the the EU Integrated Project 3D Repertoire, which funds it.

The new service of biophysical analysis (ITC and AUC) has been enthusiastically welcome and need new investments for being able to keep the pace with the constant increase of requests. All together, during the year the Facility increased the number of managed projects to more than 400.

## Technology partners

Vivascience has been an appreciated partner for the establishment of a purification protocol based on a double-tag vector. The chemical chaperone experiments have been developed in collaboration with Dr. P. Goloubinoff (University of Lausanne); the cold-adapted bacteria project was performed together with Dr. L. Tutino (University of Napoli); the characterisation of protein aggregates by FT-IR is done in collaboration with Drs. M. Lotti and S. Doglia (University Milano-Bicocca).

## Selected references


de Marco, A., Vigh, L., Diamant, S. & Goloubinoff, P. (2005). Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. *Cell Stress & Chaperones* 10, 329-339

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

Schrodel, A., Volz, J. & de Marco, A. (2005). Fusion tags and chaperone co-expression modulate both the solubility and the inclusion body features of the recombinant CLIPB14 serine protease. *J. Biotechnol.* 120, 2-10

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

## 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


Funk, J., Li, X. & Franz, T. (2005). Threshold values for detergents in protein and peptide samples for mass spectrometry. *Rapid Commun. Mass Spectrom.* 19, 2986-2988

Li, X., Wilm, M. & Franz, T. (2005). Silicone/graphite coating for on-target desalting and improved peptide mapping performance of matrix-assisted laser desorption/ionization-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). Characterization 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


# 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 beamlines for macromolecular crystallography, in developing the associated instrumentation and techniques and in providing biochemical laboratory facilities and expertise to help external visitors making measurements. The ESRF beamlines are now highly automated and all are equipped with EMBL-designed diffractometers and frozen crystal sample changers.

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 and complexes involved in RNA processing, transport and degradation. The analysis of protein–DNA interactions and mechanisms of transcriptional regulation is another important topic. 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 transcription (e.g. yeast polymerase III) and transcriptional regulation (e.g. chromatin remodelling factors).

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 and complexes 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 closely connected to the Outstation's involvement in the EU-funded SPINE project (Structural Proteomics in Europe) and its successor (SPINE2-Complexes). A very successful high-throughput robotic system for crystallisation has been implemented, and a novel, high-throughput selection method has been developed for finding soluble protein fragments from otherwise badly expressed proteins. These platforms form part of the Partnership for Structural Biology (PSB), which has been established with the neighbouring ESRF, ILL and l'Institut de Biologie Structurale (IBS). The PSB is now housed in a brand new building adjacent to the Outstation, together with the CNRS-Grenoble University Institute for Structural Virology (IVMS).

## Stephen Cusack

PhD 1976 Imperial College, London, UK. Postdoctoral work at EMBL Grenoble, France.
Staff Scientist at EMBL Grenoble 1980-1989.
Group Leader and Head of Outstation since 1989.
Joint appointment with Gene Expression Programme.


# Structural biology of RNA-protein complexes in gene expression and host-pathogen interactions

## 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 metabolism and translation. Additionally we work on the structure of adenovirus capsid proteins involved in host-cell entry, influenza virus polymerase and innate immune system receptors.

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 focused 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 respectively primary and secondary receptor binding activity 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 cell adhesion protein known as CAR) to which the fibre binds. We have determined the structure of the penton base and shown how it interacts with the N-terminal tails of the trimeric fibre. By combining the known crystal structures of the major adenovirus capsid proteins with a 10A cryo-EM reconstruction of the entire virus, our collaborators and us have obtained a quasi-atomic structure of the whole virus.

## 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 enzymes, proteins involved in nonsense mediated decay (NMD) such as Upf1, Upf2 and Upf3, and transport proteins involved in snoRNP assembly and snRNA export. 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 organisms and tropical disease producing nematodes. The focus in the adenovirus work is now on the fibre proteins of serotypes which have alternative receptors to CAR. Major new projects have started on the structure of the influenza virus RNA-dependent polymerase and on intracellular pattern recognition receptors of the innate immune system.

## Selected references


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

Mazza, C., Segref, A., Mattaj, I. & Cusack, S. (2002). Large scale induced fit recognition of an m7GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.* 21, 5548-5557

Tukalo, M., Yaremchuk, A., Fukunaga, R., Yokoyama, S. & Cusack, S. (2005). The crystal structure of leucyl-tRNA synthetase complexed with tRNA(Leu) in the post-transfer-editing conformation. *Nat. Struct. Mol. Biol.* 12, 923-930

Zubieta, C., Schoehn, G., Chroboczek, J. & Cusack, S. (2005). The structure of the human Adenovirus 2 penton. *Mol. Cell* 17, 121-35


## 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 at the best level the MX beamlines of the EMBL/ESRF Joint Structural Biology Group. In collaboration with the Synchrotron Crystallography Team (opposite), 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 Institut Laue-Langevin. 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, diffractometers (MD2x models), air bearing axis, on-beam axis video microscopes, kappa goniometers and automatic sample changers (SC3 model developed and constructed in collaboration with the ESRF and MRC France).

At the ESRF, seven MX beamlines and the BM14 UK CRG beamline have been equipped with the MD2x diffractometers and with SC3 sample changers (see figure, opposite). The c3d crystal alignment library is now implemented on the MD2x diffractometers, allowing thus fully automated crystal processing. In addition, a complementary method based on UV induced fluorescence is under evaluation at the ID23-1 beamline. The team is also actively involved in the test of the automated data collection pipeline that makes extensive use of the MD2x goniometers, SC3 sample changers and C3D crystal centring software.

Crystal Reorientation is a method we want to use routinely in a close future to optimise the data collection; a MiniKappa goniometer head designed to limit the risk of collision with external devices has been developed. Four units have been installed on the MX MAD beamlines for evaluation. First tests with users showed that new "smart data collection strategies" can be developed around this device.

Monitoring the crystal radiation damage is also essential to guarantee the quality of data collected. This is now possible using the new version of our spectrophotometer. The device that can be easily installed on ESRF MX beamline equipped with MD2M diffractometers is available for users experiments. Changes in the crystal absorption spectrum can be observed during data collection without the limitations due to the size of the prototype.

## Future projects and goals

Our major goals are the improvement of data quality and the automation of the MX beamlines. In addition to supporting and improving the existing instruments for optimal performance and reliability, several projects will be continued and new ideas explored in collaboration with the Synchrotron Crystallography Team and the Grenoble HTX crystallisation team.

- A finalised version of the Minikappa and associated comprehensive software package that includes modules for calibration, crystal re-orientation calculation and automatic re-centring, as well as smart multi pass strategy, will be developed.
- A future project is underway to develop an instrument for monitoring the crystals' humidity and which freezes them at the beamlines. This is a way to optimise the crystal quality at freezing time, and to improve the quality of the data collected.
- If successful, the crystal alignment method based on UV fluorescence will be implemented on the beamlines to increase the success rate of the automatic crystal alignment in visible light.
- Automatic detection of crystals in crystallisation trays is still the weak point of high-throughput crystallisation facilities. The feasibility of a technique based on UV fluorescence should be evaluated.

## Selected references


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

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. and Meth.* A392, 471-474

Perrakis, A., Cipriani, F., Castagna, J.C., Claustre, L., Burghammer, M., Riekel, C. & Cusack, S. (1999). Protein microcrystals and the design of a microdiffractometer: current experience and plans at EMBL and ESRF/ID13. *Acta Crystallogr. D Biol. Crystallogr.* 55, 1765-1770

Vernede, X., Lavault, B., Ohana, J., Nurizzo, D., Joly, J., Jacquamet, L., Felisaz, F., Cipriani, F. & Bourgeois, D. (2006). UV laser-excited fluorescence as a tool for the visualization of protein crystals mounted in loops. *Acta Crystallogr. D Biol. Crystallogr.* 62, 253-261

Raimond Ravelli

PhD 1998, Chemistry, University of Utrecht, The Netherlands.
At EMBL-Grenoble since 1998.

# Synchrotron Crystallography Team

## Previous and current research

Our interests lie in the improvement of synchrotron X-ray macromolecular (MX) data collection techniques to aid structural biologists in their quest for detailed atomic structures used for understanding biological functions of macromolecules. The Instrumentation Group comprises two teams led by Cipriani (opposite) and Ravelli. The group develops hardware, software and novel methodologies for data collection and phasing. Members of the group are involved in collaborations on the cell cytoskeleton and leucine-rich repeat proteins.

The instrumentation group in Grenoble works in close collaboration with the European Synchrotron Radiation Facility (ESRF) on the development, operation and maintenance of the joint structural biology group (JSBG) beamlines ID14-[1-4], ID23 and ID29. In addition, the group provides scientific and technical support for the UK CRG beamline BM14 at the ESRF.

An ongoing development of methods in crystallography has been the characterisation, mitigation and utilisation of radiation damage in MX. The intense undulator synchrotron radiation of third generation synchrotrons rapidly damages the fragile, cryo-cooled, crystalline macromolecules. Part of our research aims to get an improved understanding as well as a better crystallographic treatment of radiation damage.

## Future projects and goals

The present and future work of the group includes:

- development and commissioning of the fully MX-dedicated JSBG microfocus beamline ID23-2;
- exploring "smart data collection strategies" using an inhouse-developed Minikappa device;
- exploring the use of radiation damage for phasing;
- complementary use of XANES, MX and UV/Vis spectroscopy;
- study of protein dynamics.


*MD2M diffractometer and SC3 sample changer.*

## Selected references

Gigant, B., Wang, C., Ravelli, R.B.G., Roussi, F., Steinmetz, M.O., Curmi, P.A., Sobel, A. & Knossow, M. (2005) Structural basis for the regulation of tubulin by vinblastine. *Nature* 435, 519-522

Nanao, M.H., Sheldrick, G.M. & Ravelli, R.B.G. (2005). Improving radiation-damage substructures for RIP. *Acta Crystallogr. D Biol. Crystallogr.* 61, 1227-1237

Singh, S.K., Hora, R., Belrhali, H., Chitnis, C.E. & Sharma, A. (2006). Structural basis for Duffy recognition by the malaria parasite Duffy-binding-like domain. *Nature* 439, 741-4


## 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

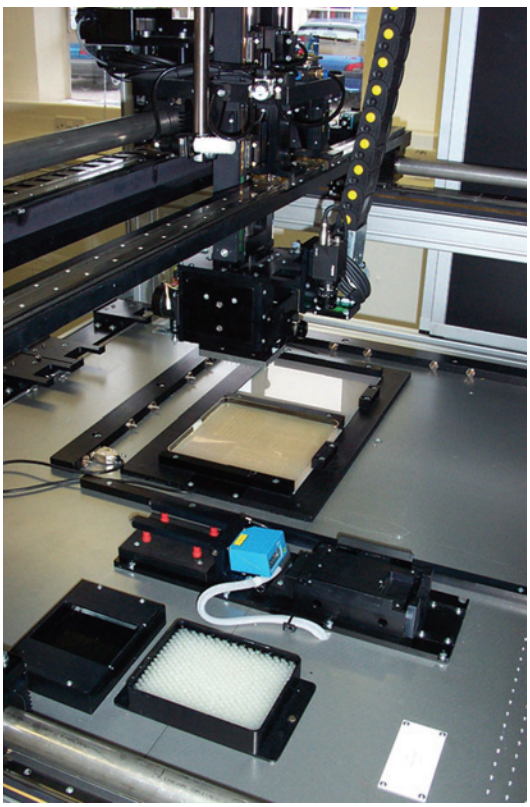
## Previous and current research

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. In addition, as part of the EU Integrated Project 3D Repertoire, 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.



*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.*

## Selected references


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

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

Hart, D.J. & Tarendeau, F. (2006). Combinatorial library approaches for improving soluble protein expression in *Escherichia coli*. *Acta Crystallogr. D Biol. Crystallogr.* 62, 19-26

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

José Márquez

PhD 1997, University of Valencia, Spain.
Postdoctoral research at EMBL.
Staff Scientist at EMBL Grenoble since 2003.
Team Leader since 2005.

# 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 EMBL Hamburg 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

Hothorn, M., D'Angelo, I., Márquez, J.A., Greiner, S. & Scheffzek, K. (2004). The invertase inhibitor Nt-CIF from tobacco: a highly thermostable four-helix bundle with an unusual N-terminal extension. *J. Mol. Biol.* 335, 987-995

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., 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, Albert-Ludwigs-Universität, Freiburg, Germany.
Postdoctoral work at Harvard University, Cambridge, USA.
Group Leader at EMBL Grenoble since July 1995.
Senior Scientist and Deputy Head of the Grenoble Outstation.
Joint appointment with the Gene Expression Unit.


# 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 mostly obtained by X-ray crystallography combined with biochemical and other biophysical techniques to gain insight into these complex cellular processes. The biochemical characterisation of proteins, protein/DNA or multi-protein complexes is followed by their structural analysis using X-ray crystallography or electron microscopy. The resulting structural information often provides valuable insight into the function of the investigated systems but also asks new questions, which are than followed up using molecular biology and biochemical techniques.

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/NFkB proteins, T-box transcription factors, STAT proteins, GCM protein and the viral transcription factor ZEBRA. Our research aims to obtain a detailed understanding of the global architecture of these transcription factors, their interaction with DNA and with other cellular factors.

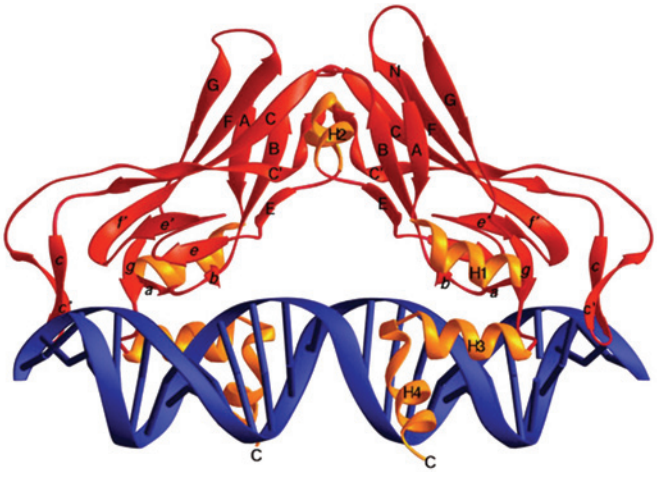
Many eukaryotic transcription factors are found in latent form in the cytosol and upon activation are imported into the nucleus, where they bind to their DNA target sequences. Regulated export from the nucleus is one important mechanism for their subsequent inactivation. Currently our research is concentrating 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 increasingly 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 TFIIIC from yeast and the nucleosome remodelling complex CHRAC.

## Future projects and goals

Our work on transcription factor/DNA complexes, transport complexes 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.

Work on RNA polymerase III transcription and the structural and functional analysis of chromatin modifying enzymes will be further extended.




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

## Selected references


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

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

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

Petosa, C., Schoehn, G., Askjaer, P., Bauer, U., Moulin, M., Steuerwald, U., Soler López, M., Baudin, F., Mattaj, I.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

Petosa, C., Patrice M., Baudin, F., Moulin, M. Artero, J.B., Müller, C.W. (2006). Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein. *Mol. Cell* 21, 565-572

## Neutron Team

L-r: François Dauvergne, Matthew Blakeley, Monika Budayova-Spano, Marianne Weidenhaupt, Marie-Thérèse Dauvergne.

# Deuteration Isotope Labelling Facility

## Previous and current research

Neutron protein crystallography can provide a powerful complement to X-ray crystallography by enabling key hydrogen atoms to be located in biological structures that cannot be seen by X-ray analysis alone. The neutron Laue diffractometer LADI, run jointly by EMBL and ILL at the ILL high flux reactor in Grenoble, is a dedicated facility for neutron protein crystallography at high-resolution (1.5 Å) and provides 10-100 fold gains in efficiency compared with conventional neutron diffractometers. The availability of fully deuterated protein eliminates the hydrogen incoherent scattering contribution to the background and brings further ~10-fold improvements in the signal to noise ratios.

The production of D-labelled macromolecules is done using bacterial expression systems. Growth on deuterated media is slow and produces lower yields of biomass and recombinant proteins. The molecular response mechanism(s) underlying bacterial adaptation to deuterium are unknown. In order to understand and improve the adaptation process, we investigate the deuterium effect on *E. coli* in a comparative proteomic approach using 2D gel electrophoresis. Alternatively, an *E. coli* expression library is used to identify key proteins that could trigger more efficient bacterial adaptation.

A major hurdle to neutron protein crystallography is that unusually large crystals (~1mm$^3$) are required to compensate for the weak flux of available neutron beams. A method and a device for the promotion of crystal growth by keeping the growth solution metastable during the growth process have been developed. This works by regulating the temperature of the growth solution using control parameters determined *in situ* during the growth process.

Whilst the X-ray structures of perdeuterated and hydrogenated proteins are essentially indistinguishable to near atomic resolutions, the subtle differences in fixed-point properties of $D_2O$ and $H_2O$ affect protein solubility, intermolecular interactions and association between H/D macromolecules in a small but significant way. We are characterising the effects of $D_2O$ and deuteration on the physico-chemical properties of bio-macromolecules in order to help optimise the crystallisation of deuterium-labelled biological macromolecules. Solubility measurements with several diverse H/D protein systems were measured and indicate that the replacement of $H_2O$ by $D_2O$ in general decreases the solubility due to an increase in the attractive intermolecular interactions in solution.

## Future projects and goals

- Development and commissioning of LADI-III; diffractometer upgrade (neutron image plates and readout system); H142 guide refit; improvement in focusing optics; relocation to higher intensity beam position.
- Improvement of Laue data analysis software; new processing tools for weak and/or spatially overlapped reflections.
- Development of neutron protein cryo-crystallography; study of the effects of cryo-cooling on large protein crystals.
- Development and use of novel techniques and strategies for optimised large protein crystal growth for neutrons and X-rays.
- Exploring the solubility and intermolecular interactions in solutions of hydrogenated as well as perdeuterated proteins from the standpoints of surface hydrophobicity, surface charge, secondary structure and protein hydration.
- Neutron structural studies addressing questions of broad biological significance concerning enzymatic mechanism, ligand-binding interactions, solvent effects, structure dynamics and their implications.
- Identification of key molecules of *E. coli* involved in adaptation to D.
- Engineering of more efficient expression hosts for D-labelled biopolymers.

## Selected references


Budayova-Spano, M., Bonneté, F., Ferté, N., El Hajji, M., Meilleur, F., Blakeley, M.P. & Castro, B. (2006). A preliminary neutron diffraction study of Rasburicase, recombinant urate oxidase enzyme, complexed with 8-azaxanthin. *Acta Crystallograph. F Struct. Biol. Cryst. Commun.* 62, 306-9

Budayova-Spano, M., Fisher, S.Z., Dauvergne, M.-T., Agbandje-McKenna, M., Silverman, D.N., Myles, D.A.A. & McKenna, R. (2006). Production and X-ray crystallographic analysis of fully deuterated human carbonic anhydrase II. *Acta Crystallograph. F Struct. Biol. Cryst. Commun.* 62, 6-9

Hazemann, I., Dauvergne, M.T., Blakeley, M.P., Meilleur, F., Haertlein, M., Van Dorsselaer, A., Mitschler, A., Myles, D.A.A. & Podjarny, A. (2005). High-resolution neutron protein crystallography with radically small crystal volumes: application of perdeuteration to human aldose reductase. *Acta Cryst. D.* 61, 1413-1417

Blakeley, M.P., Kalb, A.J., Habash, J., Helliwell, J.R., Myles, D.A.A. & Raftery, J. (2005). Neutron protein crystallography: technical aspects and some case studies at current capabilities and beyond, 62-72. In "Neutron Scattering in Biology: Techniques and Applications". Fitter, J., Gutberlet, T. & Katsaras, J. (eds.), Springer, Berlin, Germany


## 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

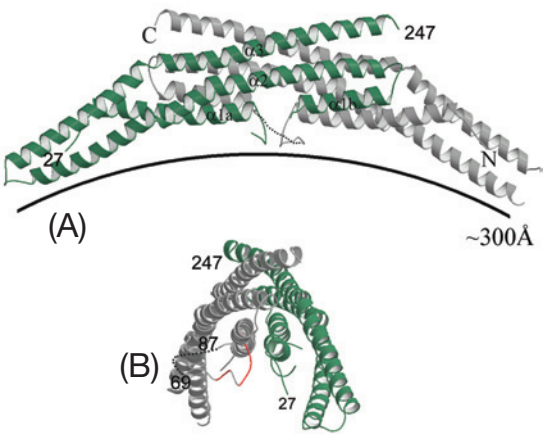
Our research concentrates on the structural basis of protein-mediated changes in membrane structures with a specific focus on enveloped virus entry and egress. We have previously determined the crystal structures of core domains from the class I membrane fusion glycoproteins of HIV-1 and Ebola virus and we are currently working towards the understanding of intermediate conformations of class I membrane fusion protein subunits. This should shed light on how to use this information for interfering with the membrane fusion process of enveloped viruses such as HIV-1.

Enveloped virus replication involves the spatial and temporal coordination of an intricate interplay between viral structural proteins and cellular factors to assemble and release infectious virions. Although 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, other components, such as the genome containing nucleocapsid, have to be formed and delivered to the correct site of assembly. Previous crystal structures of the Ebola virus matrix protein VP40 in different conformations, together with functional studies has shed light on the assembly process of filoviruses. We have now extended our structural work to the understanding of the assembly process of negative strand RNA virus nucleocapsids with the aim to link nucleocapsid formation to the function of the matrix protein during assembly.

Viral matrix proteins such as Ebola VP40 or HIV-1 Gag interact with cellular factors, which are part of a protein machinery that 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, which is composed of at least 43 different protein-protein interactions to the site of budding. Of particular interest is the structure and function of proteins that bend membranes such as BAR domain containing proteins. In this context we have recently solved the structure of endophilin, which is mainly active in endocytotic processes and has been also implicated in virus budding. The structure Endo-BAR folds into a crescent-shaped dimer, composed of two elongated three helical bundles, with two additional domains of 30 residues each inserted into helix 1 at the centre of the concave side of the dimer, which we suggest to interfere with the proposed mode of BAR domain membrane interaction. We have found similar membrane deformation functions in proteins active in MVB sorting and thus virus budding processes. Our aim is to link these functions to either vesicle formation and/or to the final membrane fission process that releases vesicles or viruses from host cell membranes.

## 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.




*General architecture of the endophilin-A1 BAR domain. (A) Ribbon representation of endo-BAR. The two monomers are shown in green and grey. Secondary structure elements and residue numbers are indicated. The regions connecting the disordered parts of the inserted domains of each monomer are drawn as dotted lines. The concave surface of the dimer would fit a curved membrane of a diameter of ~300 Å, if the disordered parts of the inserted domains fold towards each side of the dimer. (B) Endo-BAR rotated by ~70° indicates the path of the inserted domains. The sequence of the highly conserved inserted domain is shown in red (for one monomer). The disordered part is indicated by dotted lines.*

## 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

Weissenhorn, W. (2005). Crystal structure of the endophilin-A1 BAR domain. *J. Mol. Biol.* 351, 653-661


# EMBL Hamburg, Germany

The Hamburg Outstation is situated in one of the most beautiful areas of the Hanseatic city of Hamburg in northern 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 through the DORIS positron storage ring and the PETRA-3 ring within the near future (expected start of operation 2009). At present, the EMBL Hamburg Unit operates seven SR beamlines with applications in life sciences, ranging from biocrystallography to small angle X-ray scattering of biological samples and X-ray absorption spectroscopy (EXAFS). A high-throughput crystallisation facility has been added recently. The presently available facilities will be complemented and, in part, replaced by two beamlines for applications in biocrystallography and one beamline for small angle X-ray scattering experiments in biology at the future PETRA-3 storage ring with top-notch optical parameters.

EMBL Hamburg has a well-established record for the development of novel, innovative technologies in structural biology. World-leading software packages for the automation of data interpretation have been developed and are used in a large number of projects across the world's research community. 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. It is integrated into an automated software pipeline, named Auto-Rickshaw, which has been compiled at EMBL Hamburg as well. Another package is called ATSAS and allows the automatic interpretation of small angle X-ray scattering data for structural shape determination.

EMBL Hamburg offers additional facilities in molecular biology, heterologous expression in prokaryotic and eukaryotic hosts, protein purification, biophysical characterisation and crystallisation for demanding and ambitious structural biology projects. Our biological interests include regulation of transcription and translation, viral replication, protein-ligand interactions in signalling proteins and protein kinases, metal-containing proteins, giant muscle proteins and protein translocation into peroxisomes. One further common field of interest is one protein targets from the pathogenic mycobacteria tuberculosis. Our present research focus is on promising targets for the development of new antibiotics.

EMBL Hamburg either co-ordinates or contributes to several large scale projects funded by the European Commission. BIOXHIT concentrates on the development of state-of-the-art techniques in biological X-ray crystallography. SAXIER has a similar focus in the field of small angle X-ray scattering. SPINE-II, 3D Repertoire and VIZIER are directed towards targets or target complexes within specific research fields.


## 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.


# Structural biology and proteomics with synchrotron radiation

## Previous and current research

The central focus of my research group is on the structural characterisation of interactions in networks of biological molecules. We use synchrotron radiation to determine molecular structures at high resolution by X-ray crystallography, complemented by other structural biology methods such as NMR spectroscopy, Small Angle X-ray Scattering, *in vitro* FRET for distance measurements and *in vivo* imaging methods.

Current highlights include:

**Structural proteomics:** Our first structural proteomics project is on the 3D proteome of SH3 domains from yeast. We have been able to determine the 3D structures of half of all available SH3 domains, in the absence or presence of PXXP sequence motif containing peptide ligands, and we use the data generated to predict novel protein-protein interactions in yeast. Our future focus will be on the determination of protein-protein complexes involved some of these SH3 domains. The second structural proteomics project (www.xmtb.org) is concerned with potential drug targets from *Mycobacterium tuberculosis* (figure 1).

**The molecular architecture of muscle sarcomeres:** We have been able to determine the structure of the N-terminal assembly complex of titin, the largest gene product of the human genome, which comprises up to 38,000 residues in its largest isoform. In this complex, two titin filaments are found in an unprecedented antiparallel arrangement, which is mediated by the Z-disk protein telethonin (figure 2). Titin 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.

**Protein kinases involved in apoptosis and tumour suppression:** Recently we have determined several structures of protein kinases involved in apoptosis and tumour suppression in different conformational states.

**Protein import into peroxisomes:** We have been able to determine the structure of the peroxisome import receptor Pex5p in the presence and absence of a cargo protein (SCP2).

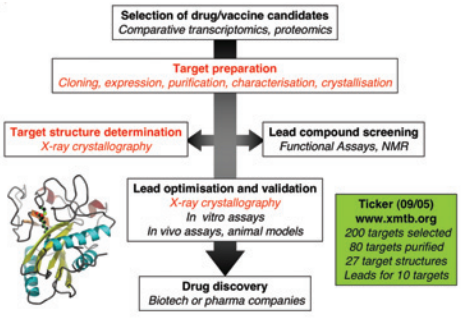



*Figure 1: Work flow of the X-MTB structural proteomics project (www.xmtb.org).*

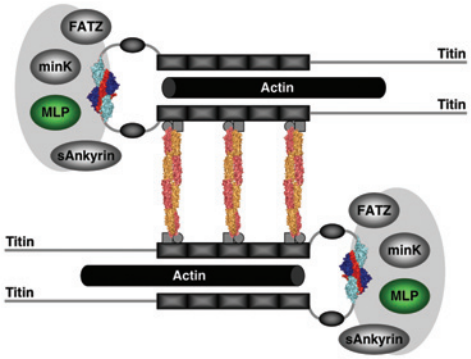



*Figure 2: Model outlining the architecture of the sarcomeric Z-disk. Titin filaments are assembled by a dual Z-disk bridging system, by a-actinin rods on a variable number of titin Z-repeats (three bridges are shown), and by telethonin via the N-terminal IG domains.*

## Selected references


Kim, C., Gajendran, N., Mittrucker, H.W., Weiwad, M., Song, Y.H., Hurwitz, R., Wilmanns, M., Fischer, G. & Kaufmann, S.H. (2005). Human alpha-defensins neutralize anthrax lethal toxin and protect against its fatal consequences. *Proc. Natl. Acad. Sci. USA* 102, 4830-4835

Kuper, J., Doenges, C. & Wilmanns, M. (2005). Two-fold repeated (βα)(4) half-barrels may provide a molecular tool for dual substrate specificity. *EMBO Rep.* 6, 134-139

Remenyi, A., Scholer, H.R. & Wilmanns, M. (2004). Combinatorial control of gene expression. *Nat. Struct. Mol. Biol.* 11, 812-815

Stanley, W.A., Sokolova, A., Brown, A., Clarke, D.T., Wilmanns, M. & Svergun, D.I. (2004). Synergistic use of synchrotron radiation techniques for biological samples in solution: a case study on protein-ligand recognition by the peroxisomal import receptor Pex5p. *J Synchrotron Radiat.* 11, 490-496

Christoph Hermes

PhD 1981, Technical University,
Munich, Germany.
At EMBL Hamburg since 1988.

# 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 3D 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.

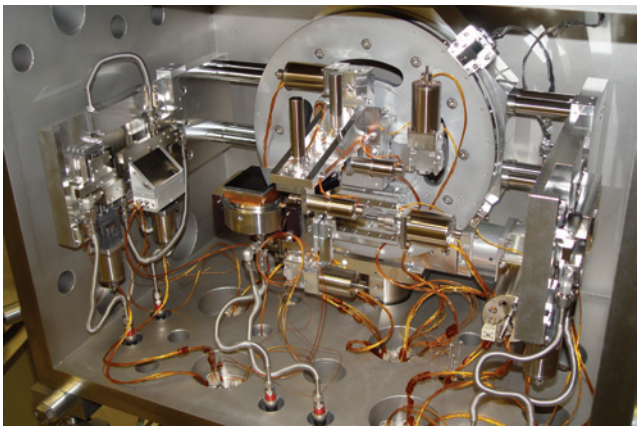
As a pilot project we have installed a sample mounting robot on one of our beamlines following the general trend to increase the level of automation of experiments leading to high-throughput facilities.

## 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 on the horizon.

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



*The double monochromator set-up of the MAD beamline X12.*

## Selected references

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., 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


## 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.


# Protein crystallography

A major focus of the group research activity is the development of underlying methodology for biological macromolecular crystallography. Targeted towards automation, high-throughput and making the most of the experiment, such developments are now given a main emphasis in the crystal structure determination and interpretation.

## Previous and current research

An ongoing development is the ARP/wARP software suite for refinement and modelling of protein structures. Owing to its leading performance and continuous progress, ARP/wARP has become the method of choice in macromolecular crystallography if the data extend to sufficient resolution. Highlights of 2005 included the advancement in modelling of secondary structural elements at low resolution, provision of web-based services for remote execution of model building tasks, the collection of a large database of computational protocols and the incorporation of ARP/wARP modules into other software pipelines.

The automation and the enhancement of the crystallographic experiment continues to be of special interest to the group. A rapid development of the XREC software for automatic centring of the crystals in the X-ray beam is particularly valuable for experiments at synchrotrons. A related development is the direct radiation damage protection of the crystals in the X-ray beam with the use of radical scavengers. Good progress has been made with the optimisation of the diffraction properties of crystals by a rational mutagenesis of surface residues that lead to the first crystal structure of a dehydrogenase used in biotechnological industry. Furthermore, we have been succeeded in the exploitation of atomic resolution protein structures for the extraction of their functionally relevant motion.

Together with S. Panjikar, P. Tucker and M. Weiss we are developing the automated structure determination pipeline Auto-Rickshaw. With the help of ARP/wARP and other software it is able to produce an interpretable electron density map and a partial structure within minutes after the completion of data acquisition. A synchrotron beamline user can receive immediate feedback as to whether the measured data are of good enough quality for successful structure determination.

## Future projects and goals

Future projects will be centred on development of novel methods and software tools for efficient 3D structure determination and analysis of biological macromolecules. We want to understand the mysteries of crystallisation of protein molecules, the structural basis of their functional motion and involvement in biological processes at the atomic level. The development of the ARP/wARP, BEST, XREC and Auto-Rickshaw software towards a unified system will aim at extension of their applicability to most cases of macromolecular structure determination and enable structural biologists, rather than dealing with crystallographic technicalities, to dedicate their efforts to the analysis of scientific information provided by X-ray structures.

## Selected references


Meijers, R., Morris, R.J., Adolph, H.W., Merli, A., Lamzin, V.S. & Cedergren-Zeppezauer, E.S. (2001). On the enzymatic activation of NADH. *J. Biol. Chem.* 276, 9316-9321

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. & Lamzin, V.S. (2005). Extraction of functional motion in trypsin crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 61, 1132-1139

Zwart, P.H., Langer, G.G. & Lamzin, V.S. (2004). Modelling bound ligands in protein crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2230-2239

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

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. Our research deals with metal functionality, metal binding and metal selectivity in biological systems. The group's projects combine structural techniques with molecular biology, biochemistry and further methods aiming at a complete understanding of metal-related biological processes. Apart from methods development (Korbas, 2006) current research includes:

**Proteins of the metallo-β-lactamase superfamily.** This superfamily, with an active site capable of binding up to two metal ions, catalyses 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:

- 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.

**An exosite unique to the metallo-β-lactamase superfamily is required for substrate binding in the Zinc Phosphodiesterase ZiPD:** *Escherichia* coli ZiPD is the best characterised protein encoded by the elaC gene family and is a model for the 3'-pre-tRNA processing endoribonucleases (tRNase Z). A metal ligand-based sequence alignment with crystallographically characterised members of the superfamily identifies a ZiPD-specific sequence insertion of ~50 residues, the ZiPD exosite. Homologs from *Bacillus subtilis*, *Methanococcus janaschii* and human share the presence of the ZiPD exosite, which is also present in the amino-terminal, but not in the carboxyl-terminal, domain of ElaC2 proteins. Another class of functionally characterised tRNase Z enzymes from *Thermotoga maritima* and *Arabidopsis thaliana* lack characteristic motifs in the exosite but possess a sequence segment with clustered basic amino acid residues.

To investigate the function of the exosite we constructed a ZiPD variant that lacks this module, ZiPDΔ. This has almost wild-type-like catalytic properties for hydrolysis of the small, chromogenic substrate bis(p-nitrophenyl) phosphate. Hence, the exosite is not required for the intrinsic phosphodiesterase activity of ZiPD. In contrast to the wild-type enzyme, ZiPDΔ does not process pre-tRNA, and gel shift assays demonstrate that only the wild-type enzyme, but not ZiPDΔ, binds mature tRNA. These findings show that the exosite is essential for pre-tRNA recognition. In conclusion, we identify a ZiPD exosite that guides physiological substrate recognition in the ZiPD/ElaC protein family (Schilling 2005).

**New metal binding motifs:** *Azotobacter vinelandii* is a diazotrophic bacterium characterised by the outstanding capability of storing Mo in a special storage protein, which guarantees Mo-dependent nitrogen fixation even under growth conditions of extreme Mo starvation. Each protein molecule can store at least 90 Mo atoms. The corresponding genes were unequivocally identified; thereby we revealed that the Mo storage protein is not related to any other known molybdoprotein. The protein's shape has been determined by small-angle X-ray scattering (SAXS). Extended X-ray absorption fine-structure spectroscopy (EXAFS) identified a metal-oxygen cluster bound to the Mo storage protein. Thus this protein is the only known non-iron metal storage system in the biosphere containing a metal-oxygen cluster (Gnida, 2005).

## Future projects and goals

Future projects will make use of the research results achieved in recent years and, in addition to the metal specificity of proteins, focus on the regulation of metal concentrations in cells. We continue to widen the range of techniques available for research on proteins isolated by the group. At present we combine structural analysis (e.g. XAS, SAXS, protein crystallography) and biochemical methods (element analysis, isothermal calorimetry, enzyme kinetics).

## Selected references

Fenske, D., Gnida, M., Schneider, K., Meyer-Klaucke, W., Schemberg, J., Henschel, V., Meyer, A.K., Knochel, A. & Müller, A. (2005). A new type of metalloprotein: The Mo storage protein from *azotobacter vinelandii* contains a polynuclear molybdenum-oxide cluster. *Chembiochem.* 6, 405-413

Korbas, M., Fulla Marsa, D. & Meyer-Klaucke, W. (2006). Kemp – a program script for automated biological X-ray absorption spectroscopy data reduction. *Rev. Scient. Instr.*, in press

Schilling, O., Vogel, A., Kostelecky, B., Natal da Luz, H., Spemann, D., Spath, B., Marchfelder, A., Troger, W. & Meyer-Klaucke, W. (2005). Zinc- and iron-dependent cytosolic metallo-beta-lactamase domain proteins exhibit similar zinc-binding affinities, independent of an atypical glutamate at the metal-binding site. *Biochem. J.* 385, 145-153

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 *Carboxydothermus hydrogenoformans*. *Proc. Natl. Acad. Sci. USA* 101, 446-451


## Jochen Müller-Dieckmann

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


# High-throughput crystallisation X-ray crystallography

## Previous and current research

**High-throughput crystallisation:** Since neither the ability to crystallise nor specific crystallisation conditions for biological macromolecules can be predicted, a large number of different molecular constructs have to be scrutinised with respect to a large number of different crystallisation solutions in order to identify appropriate conditions. Setting up numerous crystallisation experiments (300–600 per construct and temperature) is a very repetitive and time-intensive task and, as such, is predestined for automation. As a consequence, the processes of crystallisation, crystal visualisation and handling rely heavily on automated methods.

We have established a state-of-the-art high-throughput crystallisation facility at EMBL Hamburg which is the largest in Europe, with open access to the general user community. The facility has the capacity to generate more than 100 crystallisation plates (over 10,000 experiments) in eight hours. During 2005 we set up more than 540 samples from 82 users in 2,985 crystallisation plates; about 50% of the users came from non-EMBL institutions.

We have developed a new crystallisation format in collaboration with the Weiss team, M. Rössle from the Svergun group and the Institut für Mikrosystemtechnik (IMTEK) at Albert-Ludwigs-Universität in Freiburg. The main objective of this collaboration is the development of an easy to operate and fast device that is affordable to a wide range of academic laboratories. We were able to demonstrate in a proof of principle experiment that the new device has the capability for the crystallisation of sample proteins.

**Proteins from *Mycobacterium tuberculosis*:** Tuberculosis (TB) is the most prevalent infectious disease, affecting about a third of the world's population. Every year, more than eight million new cases are diagnosed 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. Our team contributes to the American TB Structural Genomics Consortium (http://www.doe-mbi.ucla.edu/TB) as well as the German Structural Proteomics Consortium (http://www.xmtb.org). We have chosen 10 proteins from a list of targets which have been identified in a variety of assays as being differentially expressed during various life cycles of *M. tuberculosis*. Inhibition of these proteins should therefore be detrimental for the bacterium. Several of these proteins have no clearly assigned function and are classified as hypothetical proteins or proteins of probable or putative function. Structural elucidation of these hypothetical targets should help to assign functions to these proteins and reveal their role in the bacterium. For the targets with know function attempts are made to establish functional assays in collaboration with Jens-Peter von Kries at the Forschungsinstitut für Molekulare Pharmakologie (FMP) in Berlin. They will then be subject to high-throughput screening for small molecule inhibitors to generate lead compounds for structure based drug design.

## Future projects and goals

In early 2006 we officially opened the high-throughput crystallisation facility to the general public. This requires further development of our web-based access to the facility and the crystallisation database. This is being done within the general context of developing a LIMS and automated X-ray crystallography pipeline. We expect the demand for crystallisation services to increase by 50–100% in 2006. This requires the development of more efficient (yet still user-friendly) data management and administrative tools, like a common scoring and data mining interface, to the facility.

We are currently working on ten proteins of *M. tuberculosis* that have been validated by the MPI for Infection Biology in Berlin as potential drug targets or that constitute anti-targets. Two of the targets have been crystallised, and their structural elucidation is an important goal for the future. We plan to co-crystallise compounds that have been identified as binders with their corresponding targets.

## Selected references


Atwell, S., Adams, J.M., Badger, J., Buchanan, M.D., Feil, I.K., Froning, K.J., Gao, X., Hendle, J., Keegan, K., Leon, B.C., Müller-Dieckmann, J., Nienaber, V.L., Noland, B.W., Post, K., Rajashankar, K.R., Ramos, A., Russell, M., Burley, S.K. & Buchanan, S.G. (2004). A novel mode of Gleevec binding is revealed by the structure of spleen tyrosine kinase. *J. Biol. Chem.* 279, 55827-55832

Müller-Dieckmann, J., Weiss, M.S. & Wilmanns, M. (2005). Protein crystallization for X-ray. In "Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine", Ruckpaul, K. & Ganten, D. (eds.), Springer Verlag, Heidelberg

## Dmitri Svergun

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


# 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.

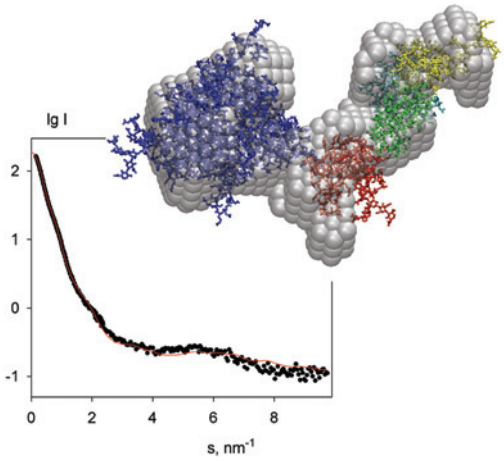
Small-angle X-ray scattering (SAXS) allows us to study native biological macromolecules, from individual proteins to large complexes, in solution under nearly physiological conditions. SAXS not only provides low resolution three-dimensional models of particle shapes but yields answers to important functional 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.




*The X-ray scattering curve on the left (intensity versus scattering angle) was used to create a low-resolution model of receptor tyrosine kinase MET, shown in grey. Superimposed on this (in colour) are atomic models of separate domains obtained from crystallography and positioned to fit the SAXS data. The protein molecule is about 10 nm across. This example illustrates how SAXS can help to assemble together high resolution models of individual domains into the model of the entire macromolecular complex.*

## Selected references


Gherardi, E., Sandin, S., Petoukhov, M.V., Finch, J,, Youles, M.E., Ofverstedt, L.G., Miguel, R.N., Blundell, T.L., Vande Woude, G.F., Skoglund, U. & Svergun, D.I. (2006). Structural basis of hepatocyte growth factor/scatter factor and MET signalling. *Proc. Natl. Acad. Sci. USA* 103, 4046-4051.

Konarev, P. V., Petoukhov, M.V., Volkov, V. V. & Svergun, D.I. (2006). ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr*. 39, 277–286.

Petoukhov, M.V. & Svergun, D.I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J.* 89, 1237-1250.

Vestergaard, B, Sanyal, S., Roessle, M., Mora, L., Buckingham, R. H., Kastrup, J. S., Gajhede, M., Svergun, D.I. & Ehrenberg, M. (2005). The SAXS solution structure of RF1 differs from its crystal structure and is similar to its ribosome-bound cryo-EM structure. *Mol. Cell* 20, 929–938.


## 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.
At EMBL Heidelberg 1985-1998; at EMBL Hamburg since 1998.


# Structural studies of proteins from pathogens

## Previous and current research

One major focus has been trying to understand the mechanism by which classical two-component systems regulate gene expression by responding to the environment. We have been 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 *M. tuberculosis* that acts at the level of transcriptional anti-termination rather than transcriptional activation/repression.

The second area of study has been on proteins that are involved in replication of viral genomes. We are continuing our work on the interaction of single-stranded DNA binding proteins of the dsDNA viruses with other components of the replication machinery. The major emphasis has, however, shifted to proteins that are important for replication of RNA viruses, especially for members of the calciviridae and flaviviridae.

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 wish to develop the preliminary structural work on two-component systems and viral replication proteins towards the identification and structural characterisation of specific inhibitors for the processes in which these proteins are involved.

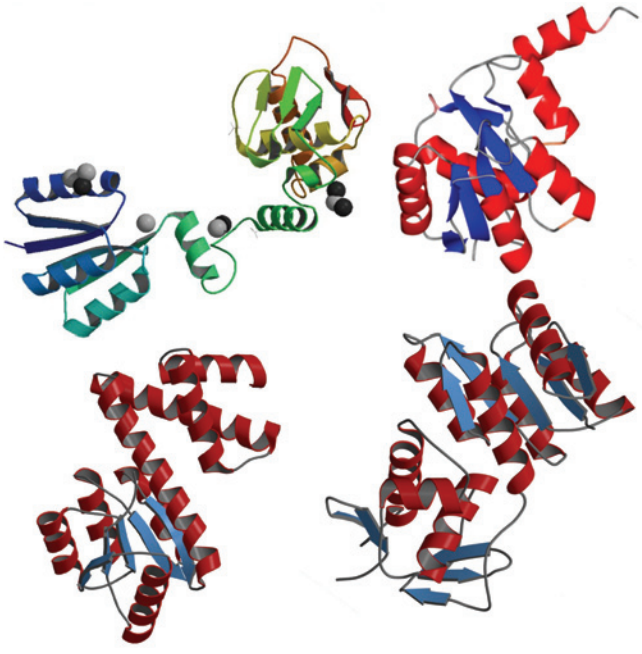
## Selected references


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, 2990-29

Morth, J.P., Gosmann, S., Nowak, E. & Tucker, P.A. (2005). A novel two-component system found in *Mycobacterium tuberculosis*. *FEBS Lett.* 579, 4145-4148

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

Sallai, L. & Tucker, P.A. (2005). Crystal structure of the central and C-terminal domain of the $\sigma^{54}$-activator ZraR. *J. Struct. Biol.* 151, 160-170

*The structures of the two-component system response regulators of* Mycobacterium tuberculosis *that were solved recently in the group. Clockwise from top left: proteins with genomic location tags Rv0491, Rv0844, Rv0903c and Rv1626.*

## Manfred Weiss

PhD 1992, Albert-Ludwigs-Universität, 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.


# X-ray crystallography of biological macromolecules

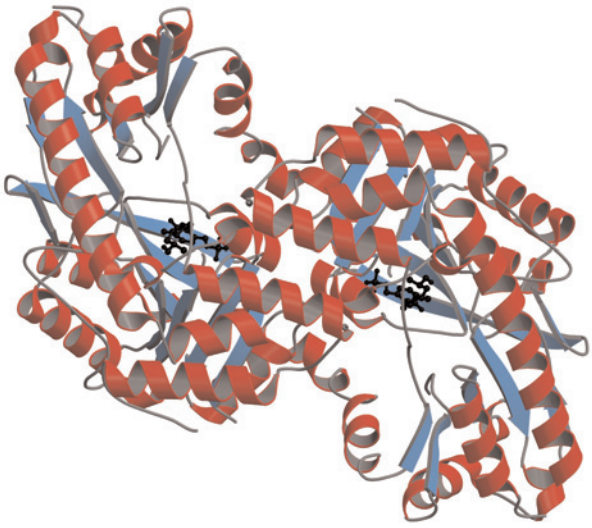
## Previous and current research

**Structural analysis of proteins from *Mycobacterium tuberculosis*.** Within the Hamburg group of the X-Mtb consortium (www.xmtb.org) a total of about 240 proteins from *M. tuberculosis* are currently under investigation. In our team the focus has been placed on proteins from the leucine and lysine biosynthesis pathways of this organism. Recently, we have been able to determine the structures of five enzymes from *M. tuberculosis*: one enzyme from the leucine biosynthesis and four from the lysine biosynthesis pathway. As an example the dimeric structure of N-succinyldiaminopimelate aminotransferase (Rv0858c) is depicted below. 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. In addition, we have determined the structure of a hypothetical protein from *M. tuberculosis*, the function of which is currently under investigation.

**Development of new methods for macromolecular structure determination.** In collaboration with the group of Paul Tucker and within the frame of a project funded by the DFG we are investigating the use of longer X-ray wavelengths in macromolecular crystallography. More specifically, we are interested in determining phases based on the very small anomalous signal provided by sulphur and phosphorus atoms, which are naturally present in proteins and nucleic acids. Recently, we have been able to establish that almost independent of the protein and the anomalous substructure the wavelength at which the largest anomalous signal-to-ratio can be obtained is about 2.0 Å. We are now in the process of extending this work towards a general strategy of phase determination from weak anomalous signals.

## Future projects and goals

The structure analysis of proteins from *M. tuberculosis* will be continued. We hope to be able to completely structurally characterise the two pathways. Also, we will lay further emphasis on complex formation between enzymes catalysing adjacent reactions in the pathways. In cases where protein-protein complexes exist, we will try to isolate and crystallise them. In cases where only transient complex formations occurs, we will try to characterise them by different means. Another important aspect in the *M. tuberculosis* project is the search for small molecules which bind to the proteins and which, in the case of enzymes, have the capacity to inhibit them. This work will be carried out in collaboration with the screening facility led by Jens-Peter von Kries at the Forschungsinstitut für Molekulare Pharmakologie (FMP) in Berlin.



*Schematic representation of the dimer of Rv0858c from* M. tuberculosis. *The view is down the twofold axis. α-helices are represented in red and β-strands in light blue. The PLP-moiety bound to Lys232 is shown as a ball-and-stick model in black.*

## Selected references


Carugo, K.D., Helliwell, J.R., Stuhrmann, H. & Weiss, M.S. (2005). Softer and soft X-rays in macromolecular crystallography. *J. Synchrotron Radiat.* 12, 410-9

Müller-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.* 61, 1263-1272

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., Müller-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

# EMBL-EBI, Hinxton, UK

The European Bioinformatics Institute (EMBL-EBI) lies in the 55 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 EMBL-EBI has a fourfold mission:

- to provide freely available data and bioinformatics services to all facets of the scientific community in ways that promote scientific progress;
- to contribute to the advancement of biology through basic investigator-driven research in bioinformatics;
- to provide advanced bioinformatics training to scientists at all levels, from PhD students to independent investigators;
- to help disseminate cutting-edge technologies to industry.

As a hub of bioinformatics in Europe, the EMBL-EBI provides data resources in all the major molecular domains. The EBI grew out of EMBL's pioneering work in providing public biological databases to the research community. Its comprehensive range of data resources now includes EMBL-Bank (DNA and RNA sequences), Ensembl (animal genomes), ArrayExpress (microarray data), UniProt (protein sequences and functional information), the Macromolecular Structure Database (protein structures), InterPro (protein families, motifs and domains) IntAct (molecular interactions) and Reactome (pathways). All of these resources are the products of international collaborations with other data providers.

We have a broad palette of research interests that complement our data resources, and these two strands of activity are mutually supportive, with many collaborations between research groups and service teams. Seven research groups aim to understand biology through the development of new approaches to interpreting biological data. These approaches are diverse, and include: classifying and understanding proteins and their interactions; mathematical analyses of evolutionary models; computational modelling of neuronal signaling; text mining; statistical approaches to functional genomics; and large scale analysis of regulatory systems and differentiation. In addition, our services teams perform extensive research to enhance existing data resources and develop new ones.

Almost all of our groups offer PhD places through EMBL's international PhD Programme. For a list of those groups with PhD places available for 2006/2007, please see the PhD studies section of our training web pages: www.ebi.ac.uk/training/Studentships/. Other positions are advertised through the EMBL jobs pages (www.embl.org/jobs).

## Janet Thornton

PhD 1973, King's College & National Inst. for Medical Research, London. Postdoctoral research at Oxford University, NIMR & Birkbeck College, London. Lecturer, Birkbeck College 1983-1989. Professor of Biomolecular Structure, University College London since 1990. Consultant at EMBL-EBI since 1994. Bernal Professor at Birkbeck College, 1996–2002. Director of the Centre for Structural Biology at Birkbeck and University College London, 1998–2001. Director of EMBL-EBI since 2001.


# Computational biology of proteins – structure, function and evolution

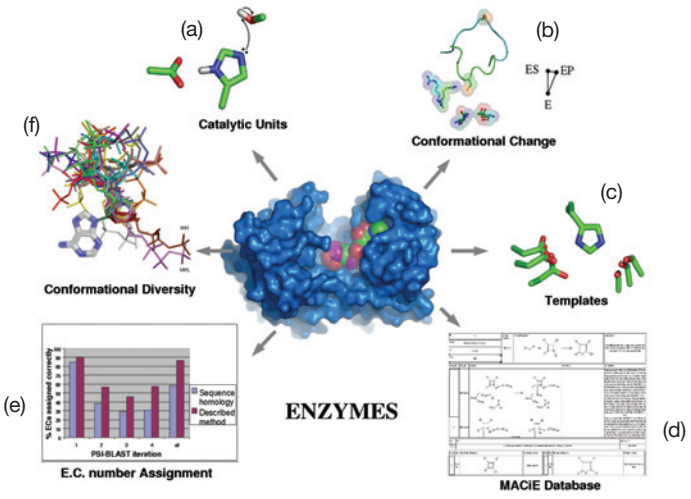
## Previous and current research

The goal of our research is to understand more about how biology works at the molecular level, how enzymes perform catalysis, how these molecules recognise one another and their cognate ligands, and how proteins and organisms have evolved to create life. We develop and use novel computational methods to analyse the available data, gathering data either from the literature or by mining the data resources, to answer specific questions. Much of our research is collaborative, involving either experimentalists or other computational biologists. In 2005 our major contributions were in the following five areas:

- enzyme structure and function;
- using structural data to predict protein function and to annotate genomes;
- evolution of organisms;
- functional genomics analysis of ageing;
- development of tools and web resources.

## Future projects and goals

Our goal is to elucidate fundamental principles relating the molecules of life to the biological processes that they perform by analysing and combining sequence, structure and expression information. Future research in the laboratory will address these key questions: can we relate function to structure, and thereby predict function from structure? Can we improve drug design? Can we understand more about catalytic mechanisms, gene duplications and phylogenetic history to elucidate the evolution of proteins, pathways and organisms? What are the molecular processes involved in ageing?




*Summary of enzyme research in the Thornton group. Clockwise from top left: (a) we have identified "catalytic units" – groups of interacting catalytic residues that recur in many unrelated catalytic sites. (b) Conformational changes can occur in enzymes upon binding of substrate and following production of product. We have analysed the extent and nature of these changes. (c) Structural templates can be used to represent catalytic sites. We have validated the performance of these templates in identifying homologues. (d) In collaboration with John Mitchell at Cambridge University, we maintain the MACiE database, which records details of enzyme mechanisms. (e) We have assigned enzymatic functions (classified by Enzyme Commission number) to protein sequences whose function is unknown, on the basis of sequence homology combined with conservation of catalytic residues. (f) Flexible substrates and cofactors can bind to different enzymes in a range of conformations. We have analysed their conformational diversity.*

## Selected references


Freilich, S. et al. (2005). Relationship between the tissue-specificity of mouse gene expression and the evolutionary origin and function of the proteins. *Genome Biol.* 6, R56

Gutteridge, A. & Thornton J.M. (2005). Understanding nature's catalytic toolkit. *Trends Biochem Sci.* 30, 622-629

Laskowski, R.A. et al. (2005). Protein function prediction using local 3D templates. *J. Mol Biol.* 351, 614-626

Nobeli, I. et al. (2005). A ligand-centric analysis of the diversity and evolution of protein-ligand relationships in *E. coli*. *J. Mol. Biol.* 347, 415-36


## Paul Bertone

PhD 2005, Yale University, USA.
At EMBL-EBI since 2005.


# Differentiation and development

## Previous and current research

Much of our previous work has entailed the development of new microarray platforms for genomic and proteomic analyses, extending these technologies beyond conventional array-based gene expression profiling and towards the global analysis of complex biological systems. Related platforms include chromosome- and genome-scale DNA tiling arrays for the discovery of functional elements in eukaryotes (Bertone, Gerstein & Snyder, 2005), and protein-based microarrays for the large-scale characterisation of biochemical activities (Bertone & Snyder, 2005). These enable the high-resolution mapping of transcripts and regulatory sequences in eukaryotes, as well as the characterisation of various biochemical activities on a proteome-wide scale.

DNA tiling arrays are designed to interrogate large regions of genomic sequence in an unbiased fashion, allowing the identification of novel transcribed RNAs and *cis*-regulatory elements. This approach represents a conceptual departure from traditional gene-based arrays, which focus exclusively on coding sequences, and its implementation has required significant technological advances in polymer synthesis and array fabrication, as well as the development of new array design algorithms and data analysis methods (Bertone et al., 2006).

Comprehensive tiling arrays have recently been constructed to represent the entire non-repetitive sequence of both strands of the human genome (Bertone et al., 2004). These systems are now being used for high-resolution analyses of gene transcription and regulation, thereby annotating the genomes determined through large-scale sequencing efforts. Hybridisation to various RNA sources has identified thousands of previously unannotated transcripts, and chromatin profiling experiments are revealing a vast network of transcription factor binding sites and chromatin modifications across the human genome.

## Future projects and goals

Our research group investigates the cellular and molecular processes underlying mammalian stem cell differentiation and lineage commitment. Stem cells maintain a unique state of self-renewal, whereby they are capable of producing identical daughter cells, and pluripotency, or the ability to differentiate into specialised progenitors whose fates are more restricted. These properties are fundamental to cellular diversification and growth patterning during embryogenesis and development, as well as the initiation of cellular repair processes throughout adult life.

A primary focus of our research is to map the global transcriptomic and proteomic activity of neural and embryonic stem cells during various stages of differentiation and lineage commitment in tissue culture. This work involves the application of sophisticated genomic, proteomic and computational analyses, combined with new advances in cell culture methods to establish pure undifferentiated stem cell lines for investigation. Experimental projects are realised in part through collaborative research programmes established with laboratories in the Cambridge area and elsewhere; our group performs computational analyses in the design of these experiments and the interpretation of results. The generation of large-scale data from microarray and functional genomic experiments will help to identify and characterise the regulatory influence of key transcription factors, signalling genes and non-coding RNAs involved in early developmental pathways, leading to a more detailed understanding of the molecular mechanisms of vertebrate embryogenesis.

We are also involved in large-scale analyses of microarray-based gene expression experiments comparing numerous stem cell lines with differentiated tissues, in collaboration with the Genomics and Regulatory Systems Group headed by Nick Luscombe. The analysis of these experiments is expected to be an ongoing project, as we incorporate results from disparate sources when new data become available. Another interest lies in the integration of data from related work on epigenetic imprinting and gene suppression studies, through our interactions with the EBI Epigenomics team led by Ewan Birney as a joint initiative between the EMBL-EBI and the Wellcome Trust Sanger Institute. This work is expected to provide key insights into the effects of chromatin modification on the determination of cell fate.

## Selected references


Bertone, P., et al. (2006). Design optimization methods for genomic DNA tiling arrays. *Genome Res.* 15, 674-680

Bertone, P., Gerstein, M. & Snyder, M. (2005). Applications of DNA tiling arrays to experimental genome annotation and regulatory pathway discovery. *Chromosome Res.* 15, 259-274

Bertone, P. & Snyder, M. (2005). Advances in functional protein microarray technology. *FEBS J.* 272, 5400-5411

Bertone, P. et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science* 306, 2242-2246

## Nick Goldman

PhD 1992, Cambridge University, UK.
Postdoctoral research at National Institute for Medical Research, London, and Cambridge University, UK.
Wellcome Trust Senior Fellow 1995-2006.
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 study of the theoretical foundations of phylogenetic analyses of sequence data, aiming to understand and improve inferential techniques and develop 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 a manner intended to add 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. Finally, methods and models are combined in biological applications designed to learn about living organisms' evolutionary history and about the processes of evolution that have shaped the diversity of life on Earth.

Increased understanding of the processes of molecular evolution has come from models that acknowledge the constraints placed on evolution by protein secondary structure, that relax the restriction that DNA substitutions occur only as point mutations (i.e. affect single sites), and that can represent the forces of natural selection acting on DNA 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. We have devised, coded and applied these mathematical models 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 (http://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 our group's 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 and large genomic regions 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 generate improved alignments of genomic and protein-coding DNA, and implement complex models of DNA and protein sequence evolution. We will be using comparative genomic data to study previously undetected functional elements, in particular looking at rates of evolution and the effects of natural selection, and heterogeneity of these processes both spatially throughout genomes and temporally over the evolution of species.

## Selected references


Loytynoja, A. & Goldman, N. (2005). An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl Acad. Sci. USA*, 102, 10557-10562

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

Pardi, F. & Goldman, N. (2005). Species choice for comparative genomics: Being greedy works. *PLoS Genet.* 1, 672-675

Whelan, S. et al. (2006). PANDIT: an evolution-centric database of Protein and Associated Nucleotide Domains with Inferred Trees. *Nucleic Acids Res.* 34, D327-D331


## Wolfgang Huber

PhD 1998, Albert-Ludwigs-Universität, Freiburg, Germany.
Postdoctoral Research at IBM Almaden Research Center, San José, California and German Cancer Research Center (DKFZ), Heidelberg, Germany.
At EMBL-EBI since 2004. Joint appointment with the Gene Expression Unit.


# Functional genomics

## Previous and current research

The Huber group develops advanced mathematical and statistical methods for analysing and understanding functional genomics data. This includes:

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

In addition, we regard the publication of high-quality scientific software as complementary to the publication of new methodical approaches and contribute 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 using automated microscopy. To make these fruitful for systematic models of biological processes, we aim to stay at the forefront of developments in experimental design, data analysis, statistical software and mathematical modelling. An emphasis is on project-oriented collaborations with experimenters.

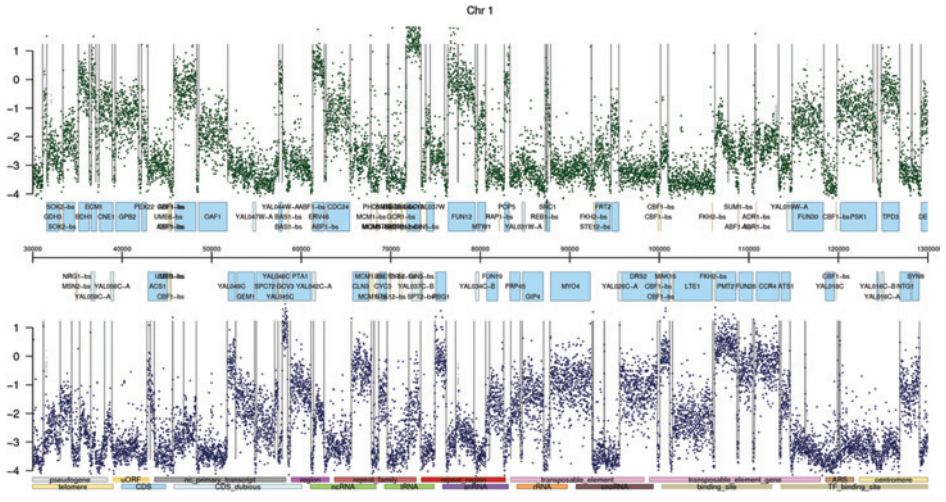
## Selected references


Arlt, D. et al. (2005). Functional profiling: From microarrays via cell-based assays to novel tumor relevant modulators of the cell cycle. *Cancer Res.* 65, 7733-7742

David, L. et al. (2006). A high-resolution map of transcription in the yeast genome. *Proc. Natl. Acad. Sci. USA* 103, 5320-5325

Gentleman, R. et al. (2005). Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer Verlag

Sultmann, H. et al. (2005). Gene expression in kidney cancer is associated with cytogenetic abnormalities, metastasis formation, and patient survival. *Clin. Cancer Res.* 11, 646-655


*Visualisation of yeast tiling array intensities along 100 kB of yeast chromosome 1, corresponding to about 1% of the complete genome, at one condition, exponential growth in YPD. Methods to analyse and model such data for multiple conditions, time courses, and different genetic backgrounds are being developed. The data and the analysis are part of a collaboration with Lars Steinmetz, EMBL Heidelberg.*

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 Le Novère group's interests revolve around signal transduction in neurons, ranging from the molecular structure of membrane proteins involved in neurotransmission to modelling signalling pathways. In particular, we focus our work on 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 computational 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.

Modelling requires various types of software, including design environments and simulators. The Systems Biology Markup Language (SBML; http://sbml.org) is designed to facilitate the exchange of biological models between different types of software. Involved in the development of SBML since the origin of the project, we are working on the language extensions, and are developing software that supports SBML. Moving from the form to the content, we are also developing standards for model curation and annotation (MIRIAM), and controlled vocabularies to increase the information content of models (the Systems Biology Ontology).

However well a model is designed, it is useful only if potential users can access it. The BioModels Database (http://www.ebi.ac.uk/biomodels/) is an effort to develop a data resource that allows biologists to store, search and retrieve published mathematical models of biological interest. The syntax and semantics of the models in the BioModels Database are verified. Furthermore, they are annotated and linked to relevant data resources, including publications, databases of compounds and pathways and controlled vocabularies.

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. E-MeP is a research platform that will focus on solving the bottlenecks that prevent the high-throughput determination of high-resolution structures of membrane proteins and membrane protein complexes. Most of the project's 300 pre-selected eukaryotic targets are involved in neurotransmission, many of them implicated in devastating pathologies. Our group maintains the computing infrastructure of the consortium (both communication support and data storage) and participates in the bioinformatics analysis of the targets.

## 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 $Ca^{2+}$ waves and phosphorylation gradients. Taking advantage of the growth of the BioModels Database, 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.

## Selected references

Donizelli, M. et al. (2006). LGICdb. A manually curated sequence database after the genomes. *Nucleic Acids Res.* 34, D267-269

Le Novère, N. et al. (2006). BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res.* 34, D689-691

Le Novère, N. et al. (2005). Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat. Biotechnol.* 23, 1509-1515

Schilstra, M.J. et al. (2006). CellML2SBML: Conversion of CellML into SBML. *Bioinformatics* 22, 1018-1020


## 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 the 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 so far 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 (see figure).

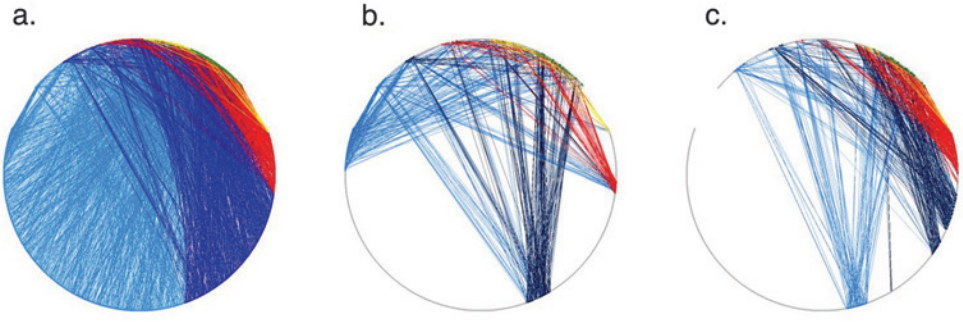
Some our projects have included:

- 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.

Our aim is to continue advancing our understanding of regulatory systems in comparatively simple organisms and to expand eventually towards more complex organisms. A major element of our research is to interact closely with research groups performing functional genomics experiments.

## Future projects and goals

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




*A network representation displays (a) the known yeast regulatory system and its usage during (b) the cell cycle and (c) the stress response. Distinct regions of the network are clearly used and this is accompanied by fundamental changes in the underlying network structure.*

## Selected references


Bertone, P. et al. (2006). An overview of regulatory pathways in pathogenic bacteria. *Curr. Opin. Microbiol.*, in press

Darot, J. et al. (2006). Transcription regulation in *Saccharomyces cerevisiae*: an overview from a genomic perspective. In "Yeast Gene Analysis", Stark, M & Stansfield, I. (eds.), Elsevier, in press

Luscombe, N.M. et al. (2004). Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* 431, 308-312

Yu, H. et al. (2003). Genomic analysis of gene expression relationships in transcriptional regulatory networks. *Trends Genet.* 19, 422-427

## 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.


# Information extraction from scientific literature

## Previous and current research

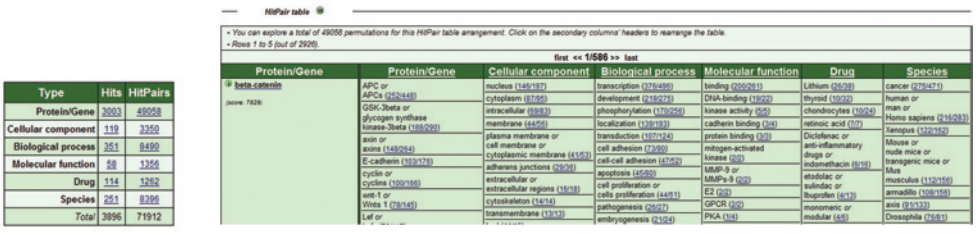
Research in the Rebholz group is focused on information extraction from the scientific literature. Our goal is to develop software that analyses a document, identifies all the biomedical terms in it and links them to relevant biomedical databases, identifies all events such as protein–protein interactions and recognises all biomedical nomenclature, including point mutations, karyotypes and dissociation constants. One important step in this analysis is retrieving the sense of ambiguous terms such as abbreviations (acronym disambiguation). Solutions to these text-mining challenges should allow authors and curators to submit documents that the software has not previously "seen", and should receive the annotations instantly.

We have made several significant advances over the past year. First, we have proved that acronyms in Medline can be disambiguated at 98% accuracy (98.2% precision, 98.9% recall). Second, we have developed two new services (EBIMed and Protein Corral) that combine information retrieval and information extraction. EBIMed (http://www.ebi.ac.uk/Rebholz-srv/ebimed/index.jsp) is the first public text-mining service that analyses Medline abstracts in real time to identify large sets of terms from public resources and links the findings to online data resources. Protein Corral (http://www.ebi.ac.uk/Rebholz-srv/pcorral/index.jsp) retrieves Medline abstracts and identifies protein–protein interactions with high precision using language patterns.

As part of our collaboration in the EU-funded Network of Excellence "Semantic Interoperability and Data Mining in Biomedicine" (SemanticMining; http://www.semanticmining.org/), we organised the First Symposium of Semantic Mining in Biomedicine, which took place on April 2005 at the EMBL-EBI. A new initiative to improve the interoperability of information-extraction and text-mining modules was launched as a result of the symposium. Further collaborations include the EU-funded Symbiomatics project (http://www.symbiomatics.org) and various groups at EMBL Heidelberg.

## 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 a range of text-mining techniques including co-occurrence, dependency or full parsing, hidden Markov models or support vector machines. Topics of particular interest include: mapping terms from text to ontological resources and vice versa; 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.



| Type | Hits | HitPairs |
|---|---|---|
| Protein/Gene | 3003 | 49058 |
| Cellular component | 119 | 3350 |
| Biological process | 351 | 8490 |
| Molecular function | 58 | 1356 |
| Drug | 114 | 1262 |
| Species | 251 | 8396 |
| Total | 3896 | 71912 |

HitPair table

• You can explore a total of 49058 permutations for this HitPair table arrangement. Click on the secondary columns' headers to rearrange the table.
• Rows 1 to 5 (out of 2926).

first << 1/586 >> last

| Protein/Gene | Protein/Gene | Cellular component | Biological process | Molecular function | Drug | Species |
|---|---|---|---|---|---|---|
| beta-catenin | APC or APCs (252/448) | nucleus (146/197) | transcription (379/496) | binding (200/261) | Lithium (26/38) | cancer (275/471) |
| | GSK-3beta or glycogen synthase kinase-3beta (188/290) | cytoplasm (87/95) | development (219/275) | DNA-binding (19/22) | thyroid (10/32) | human or man or Homo sapiens (216/283) |
| | axin or axins (149/264) | intracellular (68/83) | phosphorylation (170/256) | kinase activity (5/5) | chondrocytes (10/24) | Xenopus (122/162) |
| | E-cadherin (123/176) | membrane (44/56) | localization (139/193) | cadherin binding (3/4) | retinoic acid (7/7) | Mouse or nude mice or transgenic mice or Mus musculus (112/158) |
| | cyclin or cyclins (100/166) | plasma membrane or cell membrane or cytoplasmic membrane (41/53) | transduction (107/124) | protein binding (3/3) | Diclofenac or anti-inflammatory drugs or indomethacin (6/16) | |
| | wnt-1 or Wnts 1 (78/145) | adherens junctions (29/36) | cell adhesion (73/90) | mitogen-activated kinase (2/2) | | |
| | Lef or | extracellular or extracellular regions (16/18) | cell-cell adhesion (47/52) | MMP-9 or MMP-9 (2/2) | etodolac or sulindac or ibuprofen (4/13) | armadillo (109/156) |
| | | cytoskeleton (14/14) | apoptosis (45/80) | E2 (2/2) | | axis (91/133) |
| | | transmembrane (13/13) | cell proliferation or cells proliferation (44/51) | GPCR (2/2) | monomeric or modular (4/6) | Drosophila (76/81) |
| | | | pathogenesis (26/27) | PKA (3/4) | | |
| | | | embryogenesis (21/24) | | | |

*Results from an EBIMed summary page. The keyword query "Wnt" leads to the retrieval of 3,888 abstracts referring to 3,003 Uniprot proteins and covering 49,058 hit-pairs of UniprotKB/Swiss-Prot proteins, i.e. co-occurrences of UniprotKB/Swiss-Prot proteins in single sentences. The hit-pair table lists "β-catenin" in the left-most ("primary") column. All other columns are secondary columns and each of their entries refers to the entry in the primary column (hit-pair), e.g. "APC" in a secondary column next to "β-catenin" in the primary column refers to the co-occurrence of these two terms. Selecting the header of any secondary column induces it to become the primary column and rearranges all the hit-pairs accordingly. The two numbers after the hit-pair term refer to the number of abstracts and sentences, respectively, that contain this hit-pair. The numbers link to the list of sentences containing the hit pair, where each sentence is shown with the abstract's PubMed ID, its first author and year of publication, and a link to the abstract. Abstracts and sentences are marked up with identified terminology, which also links to the data resource from which the terms were taken.*

## Selected references


Gaudan, S. et al. (2005). Resolving abbreviations to their senses in Medline. *Bioinformatics* 21, 3658-3664

Hakenberg, J. et al. (2005). LLL'05 Challenge: Genic Interaction Extraction with Alignments and Finite State Automata. In 4th Learning Language in Logic Workshop (LLL05) at the 22nd Int. Conf. on Machine Learning. Cussens, J. & Nédellec, C. (eds.), LLL, in press

Kirsch, H. et al. (2005). Distributed modules for text annotation and IE applied to the biomedical domain. *Int. J. Med. Inform.*

Rebholz-Schuhmann, D. et al. (2005). Facts from text – is text mining ready to deliver? *PLoS Biol.* 3, 188-191

Rolf Apweiler

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

# The Sequence Database Group

## Previous and current research

The activities of the Sequence Database group are focused on the production of protein sequence, protein family and nucleotide sequence databases at the EMBL-EBI. We maintain and host EMBL-Bank (the European nucleotide sequence database), the UniProt protein resource, and a range of databases around these two core databases. Our main research activities are on automatic annotation of proteins, genes and genomes, as well as mining high-throughput genomics and proteomics datasets.

EMBL-Bank (Cochrane et al., 2006) is Europe's primary nucleotide sequence resource. It is produced as part of the International Nucleotide Sequence Database Collaboration (www.insdc.org). Each of the three groups in the collaboration collects a proportion of the total sequence data reported worldwide, and all new and updated database entries are exchanged between the groups on a daily basis. EMBL-Bank contains over 58 million DNA and RNA sequences, ranging from ten base pairs to entire genomes. Its sequences come from three main sources: individual research groups, genome-sequencing projects and patent applications.

UniProt (Bairoch et al., 2005) is produced by the UniProt Consortium, a collaboration among the EMBL-European Bioinformatics Institute (EMBL-EBI), the Swiss Institute of Bioinformatics (SIB) and Georgetown University Medical Center's Protein Information Resource (PIR). UniProt comprises three components: the UniProt KnowledgeBase (UniProtKB), which gives the user as much functional information as possible about a protein; the UniProt Archive (UniParc), which is the world's most comprehensive protein sequence archive, and the UniProt Reference Clusters (UniRef), which allow users to perform rapid searches of the sequence data in UniProt.

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 (Mulder et al., 2005) to recognise domains and to classify all protein sequences in the UniProt Knowledgebase (UniProtKB) into families and superfamilies, and various systems to transfer 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). Errors are checked by 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 (Wieser et al., 2004). Our 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 EMBL-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

Bairoch, A. et al. (2005). The Universal Protein Resource (UniProt). *Nucleic Acids Res.* 33, D154-D159

Kanz, C. et al. (2005). The EMBL Nucleotide Sequence Database. *Nucleic Acids Res.* 33, D29-33

Mulder, N.J. et al. (2005). InterPro, progress and status in 2005. *Nucleic Acids Res.* 33, D201-D205

Wieser, D. et al. (2004). Filtering erroneous annotation. *Bioinformatics* 20 (Suppl.), I342-I347

Ewan Birney

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

# Ensembl, Reactome and the Birney Research group

## Previous and current research

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

The Ensembl team is one of the leading groups for genome annotation, which is the process of finding functional elements in large genomes using computational approaches. Ensembl 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. It is now possible to view multiple alignments of selected mammalian genomes in Ensembl. Progress in the prediction of non-coding RNA genes has also contributed significantly to the richness of annotation in Ensembl.

Reactome (www.reactome.org), a curated database of biological processes in humans, 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. Reactome is produced in collaboration with Lincoln Stein's group at Cold Spring Harbor Laboratory (CSHL), NY, USA. 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 the EMBL-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.

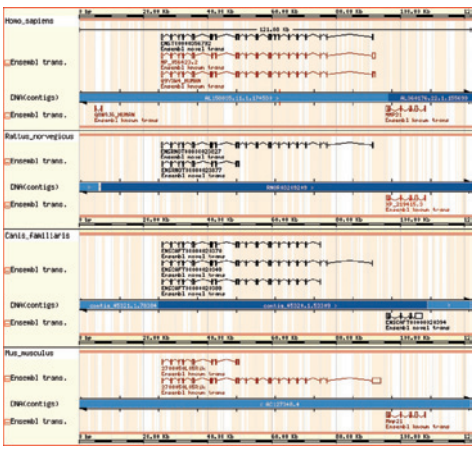
Reactome has grown steadily since its launch in 2004. New pathways include the complete insulin receptor pathway and the electron transport chain.

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 metozoans;
- development of novel hidden Markov models for gene prediction.



*Multiple alignment of four genomes – human, rat, dog and mouse. Gene annotations from each genome are mapped onto the multiple alignment. The shaded areas are aligned regions; the intervening white areas are large gaps in a particular genome.*

## Selected references

Ettwiller, L. et al. (2005). The discovery, positioning and verification of a set of transcription-associated motifs in vertebrates. *Genome Biol.* 6

Hubbard, T. et al. (2005). Ensembl 2005. *Nucleic Acids Res.* 33, D447-D453

Joshi-Tope, G. et al. (2005). Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.* 33, D428-D432

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


## Alvis Brazma

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

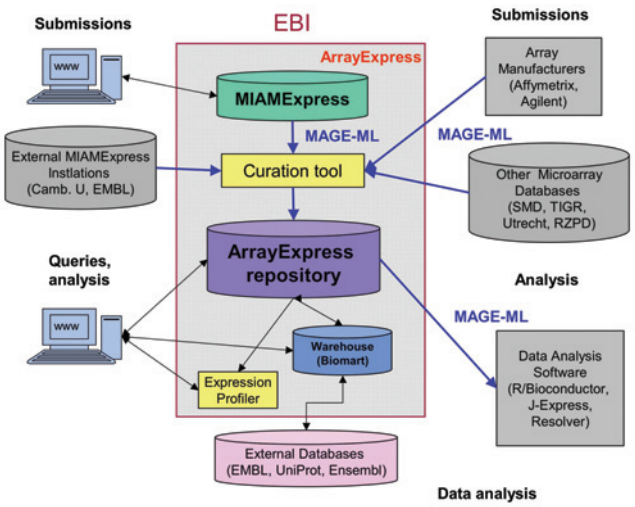

# The Microarray Informatics Team

## 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 began work to establish ArrayExpress – a public repository for microarray data. Over the past year the size of the ArrayExpress repository has more than doubled; as of May 2006 it holds data from over 40,000 microarrays. More recently, as a part of several large European projects, we have extended our work to other domains of high-throughput data, such as proteomics and metabonomics, which we are integrating with medical informatics data. ArrayExpress development is the major focus of many experienced software engineers and scientific database curators in the group, while our PhD students and postdocs focus mostly on analysing these data to build models 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 (Schlitt et al., 2005). 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 networks.

## Future projects and goals

The ongoing development of ArrayExpress from a prototype to a robust 24/7 service database will remain the main focus of the group. We will continue to improve the ease with which users can submit, access and mine data in ArrayExpress, and to populate ArrayExpress with high-quality data describing basic biological processes, disease and toxicity. 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. For instance. the study and modelling of the cell cycle will be an important collaborative project with the Wellcome Trust Sanger Institute and other partners. We will use 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 will build and analyse a four-dimensional gene expression atlas integrating expression data from different technologies, organisms, tissue types and developmental stages. Large collaborative projects, such as integration of transcriptomics, proteomics, and metabonomics data to understand the molecular mechanisms of Type 2 Diabetes and other disease will be an important activity related to medical bioinformatics.




*ArrayExpress components and MAGE-ML-based data sharing infrastructure.*

## Selected references


Parkinson, H. et al. (2005). ArrayExpress – a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.* 33, D553-D555

Sarkans, U. et al. (2005). The ArrayExpress gene expression database: a software engineering and implementation perspective. *Bioinformatics* 21, 1495-1501

Schlitt, T. & Brazma, A. (2005). Modelling gene networks at different organisational levels. *FEBS Lett.* 579, 1859-1866

Schlitt, T. & Brazma, A. (2006). Modelling in molecular biology: describing transcription regulatory networks at different scales. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361, 483-494

Kim Henrick

PhD 1974, University of Western Australia, Perth.
Postdoctoral research at the Polytechnic of North London, Imperial College of Science and Technology, SERC Darebury Laboratory, Cambridge Centre for Protein Engineering and National Institute for Medical Research. 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 Macromolecular Structure Database (MSD) (http://www.ebi.ac.uk/msd/) team is one of the three partners in the worldwide Protein Databank (wwPDB), the consortium entrusted with the collation, maintenance and distribution of the global repository of macromolecular structure data. The focus of the MSD team is to accept depositions to the Protein Data Bank, produce an integrated relational database of three-dimensional (3D) protein structures and to provide a service to the scientific community. Since its inception, the MSD group has worked with partners around the world to improve the quality of PDB data, through a clean-up programme that addresses inconsistencies and inaccuracies in the legacy archive and through a continued commitment to improving the tools and methods for the deposition of new data by the community at large. The improvements in data quality in the legacy archive have been achieved largely through the creation of a unified data archive, in the form of a relational database that stores all of the data in the wwPDB.

The rate of deposition of data to the wwPDB continues to grow, making it increasingly challenging to process submissions to the PDB. The MSD team has been active in a number of projects to improve tools and methodologies for data harvesting, validation and archival. These include: the UK BBSRC-funded eHTPX project, which aims to unify the procedures of protein structure determination into a single interface; PIMS (also BBSRC funded), which is developing a laboratory information management system for structural biologists; the EU-funded BIOXHIT project, which is defining dictionaries for the exchange of metadata; and 3DEM, an EU-funded Network of Excellence that is developing standardised procedures for 3D electron microscopy. In addition we are working with both CCP4 (for X-ray crysallography) and CCPN (for nuclear magnetic resonance spectroscopy (NMR)) data standards and software harvesting.

In 2005 the MSD team released three new services: (1) MSDmotif allows users to search the PDB for small structural or sequence-based motifs; (2) MSDtemplate allows users to search the PDB for structure-based templates such as catalytic triads, metal-binding sites and N-linked glycosolation sites; and (3) MSDpisa is an interactive tool for exploring protein interfaces, predicting probable quaternary structures and searching for structurally similar interfaces and assemblies. Significant upgrades have been made to the MSD database, the data warehouse, the database loader and database exporter.

## 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 visualization techniques.

## Selected references

Golovin, A. et al. (2005). MSDsite: A database search and retrieval system for the analysis and viewing of bound ligands and active sites. *Proteins-Structure Function and Bioinformatics*, 58, 190-199

Krissinel, E., & Henrick, K. (2005). Detection of protein assemblies in crystals. In "Computational Life Sciences: First International Symposium, CompLife 2005", Berthold, M.R. et al. (eds.), Springer Verlag, 163-174,

Pajon, A. et al. (2005). Design of a data model for developing laboratory information management and analysis systems for protein production. *Proteins-Structure Function and Bioinformatics*, 58, 278-284

Velankar, S. et al. (2005). E-MSD: an integrated data resource for bioinformatics. *Nucleic Acids Res.* 33, D262-D265


## Henning Hermjakob

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


# The Proteomics Services Team

## Previous and current research

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. Extending the "standards" theme, we provide the IntEnz and ChEBI reference resources for enzyme and chemical nomenclature.

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. 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.

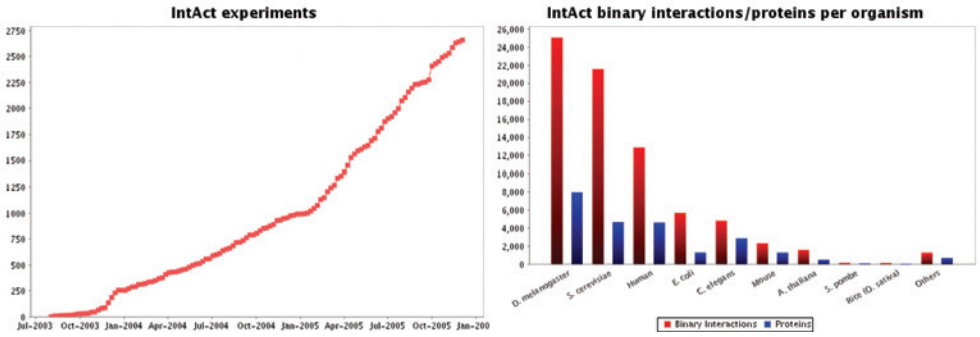
In the context of the EU-funded BioSapiens (www.biosapiens.info) and Transfog projects, the team contributes to the development of tools for proteomics data integration and distribution based on the lightweight DAS (Distributed Annotation System) protocol, in particular the DAS UniProt reference server (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 team contributes to the standardisation of proteomics-related nomenclature and data representation.

## Future projects and goals

For the coming year, we expect the IMEx consortium of interaction databases to move into production mode, implementing a regular data exchange between major interaction databases. While the political consensus seems to be reached now, the data exchange between complex, independent databases represents a major challenge. Our goal in this domain is to achieve a significant volume of regular data exchange with the IMEx partners by the end of 2006.

In the domain of mass spectrometry, we plan to follow the strategy of the IMEx consortium, and initiate a similar collaboration with other mass spectrometry databases. The PRIDE protein identification database is currently rapidly developing; the major challenge for the coming year is the establishment of an efficient submission and curation process.




*IntAct statistics. Left: the increasing number of IntAct experiments reflect the ongoing curation efforts. Right: the binary interactions/organism show the efforts to annotate interactions for a number of model organisms.*

## Selected references


Cote, R.G. et al. (2006). The Ontology Lookup Service, a lightweight cross-platform tool for controlled vocabulary queries. BMC *Bioinformatics* 7

Jones, P. et al. (2006). PRIDE: a public repository of protein and peptide identifications for the proteomics community. *Nucleic Acids Res.* 34, D659-663

Robbe, K. et al. (2005). IntAct: an open source database system and analysis tools for protein interaction data. *FEBS J.* 272, 113-113

Orchard, S. et al. (2005). Further steps in standardisation. Report of the second annual Proteomics Standards Initiative Spring Workshop (Siena, Italy 17-20 April 2005). *Proteomics* 5, 3552-3555

Nicola Mulder

PhD 1998, University of Cape Town, South Africa.
At EMBL-EBI since 1999.
Joint appointment between EMBL-EBI and the University of Cape Town since 2004.

# The InterPro Team

## Previous and current research

The InterPro team coordinates the InterPro, CluSTr and Gene Ontology Annotation (GOA) projects at EMBL-EBI.

InterPro is an integrated documentation resource for protein families, domains and functional sites, which aims to integrate signatures from the major protein signature databases. InterPro integrates data from Pfam, PRINTS, PROSITE, ProDom, SMART, TIGRFAMs, PIRSF, SUPERFAMILY, Gene3D and PANTHER.

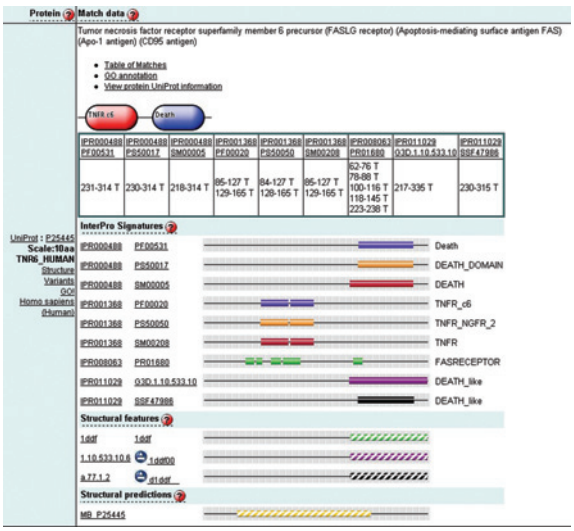
InterPro integrates signatures by rationalising where more than one protein signature is attempting to describe the same protein family/domain. These are united into single InterPro entries with biological annotation and links to external databases, e.g. GO, PDB, SCOP and CATH. InterPro precomputes matches of its signatures to UniProtKB proteins using the InterProScan software and displays the matches in various formats, including table and graphical views and the InterPro Domain Architectures view (see figure).

InterPro has important applications in the automatic annotation of proteins, the provision of statistical analysis for whole proteomes and to provide large-scale mapping of proteins to GO terms for the GO Annotation (GOA) project. The GOA and CluSTr projects now form part of the InterPro team. The CluSTr project aims to cluster all UniProtKB proteins and protein sets from complete genomes.

## Future projects and goals

Our future goals in InterPro are to integrate new signatures from member databases and to extend InterPro's existing functionality. The primary aim in integrating new signatures is coverage of new protein families. Currently, InterPro provides matches for all proteins in the UniProt Knowledgebase, however, this is being extended to all proteins in UniParc. InterPro also aims to develop new features, including browsing InterPro entries by functional categories and providing links to gene expression and protein-protein interaction data.

Future plans for CluSTr include producing a single-linkage clustering for the UniProt and IPI data sets, the Environmental samples section of UniProt (1.2. million sequences) and UniParc; developing better visualisation tools; refining CluSTr-based homology predictions, using InterPro Domain Architectures and genomic synteny; and incorporating CluSTr into EMBL-EBI's automatic annotation pipeline to predict annotation of TrEMBL entries.




*Graphical view formats of InterPro matches for protein P25445.*

The future plans of the GOA project are to increase the manual curation of proteins. The focus will continue to be on humans, but a new area, chicken gene product annotation, will be pursued. New tools for increasing coverage will be explored, such as the use of orthologues and their GO annotations, and splice variants will be annotated with GO terms.

## Selected references

Carninci, P. et al. (2005). The transcriptional landscape of the mammalian genome. Science, 309, 1559-1563

Mulder, N.J. et al. (2005). InterPro, progress and status in 2005. *Nucleic Acids Res.* 33, D201-D205

Quevillon, E. et al. (2005). InterProScan: protein domains identifier. *Nucleic Acids Res.* 33, W116-W120


## Peter Rice

BSc 1976, University of Liverpool, UK.
Previously at EMBL Heidelberg (1987-1994), The Sanger Centre (1994-2000) and LION bioscience AG (2000-2002).
At EMBL-EBI since 2003.


# Grid and e-Science R&D at the EMBL-EBI

## Previous and current research

The team'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 services provided by the group remain largely SOAP-based web services. These have proved themselves to be highly useful to prototype and develop service and metadata standards. We are looking, especially through the EMBRACE project, to migrate to true Grid services, but like many other groups we are waiting for the long-anticipated merging of web and grid service standards.

The EMBOSS project plans to expand in the coming few years to cover bioinformatics more generally, including genomics, protein structure, gene expression, proteomics, phylogenetics, genetics and biostatistics. In each area we will expect to build a service-based eScience architecture around the applications, with support and guidance from our community of users both in academia and in industry.

## Selected references


Alpdemir, M.N. et al. (2005). Contextualised workflow execution in MyGrid. In "Advances in Grid Computing – EGC 2005" Springer Verlag Berlin, 444-453

Kim, W.K. & Ison, J.C. (2005). Survey of the geometric association of domain-domain interfaces. *Proteins-Structure, Function and Bioinformatics*, 61, 1075-1088

Oinn, T. et al. (2005). Lessons in creating a workflow environment for the life sciences. *Concurr. Comput.-Pract. Exp.*, in press

Pillai, S. et al. (2005). SOAP-based services provided by the European Bioinformatics Institute. *Nucleic Acids Res.* 33, W25-W28

## Peter Stoehr

MSc Applied Genetics 1978, Birmingham University, UK.
Statistician in Agriculture Faculty, Newcastle University.
Analyst/Programmer with AFRC, Cambridge and Harpenden, UK.
At EMBL since 1988; at EMBL-EBI since 1994.

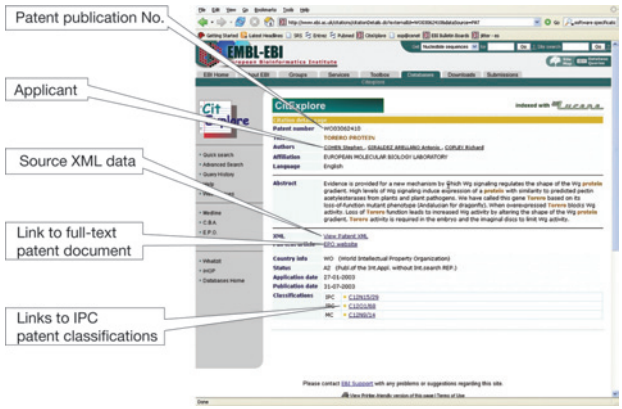

# Literature resource development

## Previous and current research

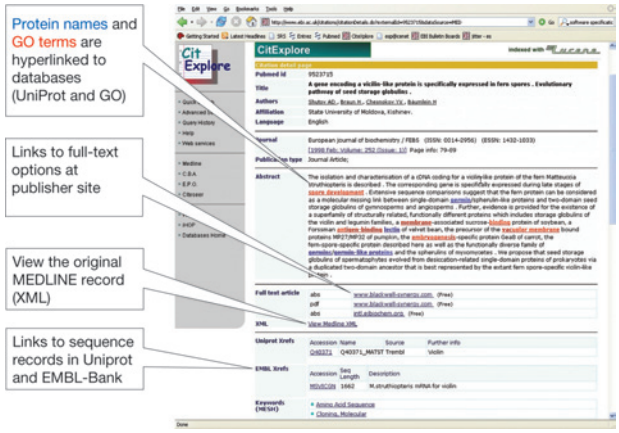
We continue to maintain literature data resources consisting of MEDLINE and patent applications for use within EMBL-EBI and in public services. We are also investigating other abstract and indexing resources (A&I) to complement MEDLINE, either in geographical or in subject coverage. In particular, we have incorporated abstracts from the Chinese Biological Abstracts (CBA) resource. These data are made available in-house for text-mining research and for database projects to improve the management of their journal citations. They are also built into EMBL-EBI's public services. A goal of the public services is to integrate the literature with EMBL-EBI's biological databases; where feasible we achieve this by implementing novel biomedical text-mining research methods.

## Future projects and goals

We expect during the summer of 2006 to launch a new service, CiteXplore, which is a query interface to our citation database of literature and patent abstracts. The search engine is based on the open-source Lucene system, with an Oracle database back-end. Links to full text, MeSH and Gene Ontology (GO) ontologies, and text-mining tools from the Rebholz group and others are integrated into the system, and we will continue to improve these features based on user feedback. We include cross-references to many of the EMBL-EBI's biological databases. The application is available at www.ebi.ac.uk/citations.




We will continue to supplement the our literature data with abstracts from journals that are cited in our biological databases but not already covered by MEDLINE, in many cases by direct interaction with journal publishers. We also have begun to capture full-text patents through a collaboration with the European Patent Office (EPO), and to investigate contributing to PubMedCentral (www.pubmedcentral.org), the free archive of full-text biomedical literature.




*Top: A MEDLINE record in CiteXplore. Left: a patent record in CiteXplore.*

## Selected references


Kirsch, H. et al. (2006). Distributed modules for text annotation and IE applied to the biomedical domain. *Int. J. Med. Inform.* 75, 496-500.

Zdobnov, E.M. et al. (2002). The EBI SRS server – new features. *Bioinformatics* 18, 1149-1150

# EMBL Monterotondo, Italy

The EMBL Monterotondo campus is situated in a green park on the banks of the Tiber river, 20km 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, which was established to capture new opportunities in mouse functional genomics and to exploit applications of mouse genetic manipulation to biomedical problems, is centred in Monterotondo. 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 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 Head of EMBL Monterotondo Outstation since 2001.


# Regenerative mechanisms in heart and skeletal muscle

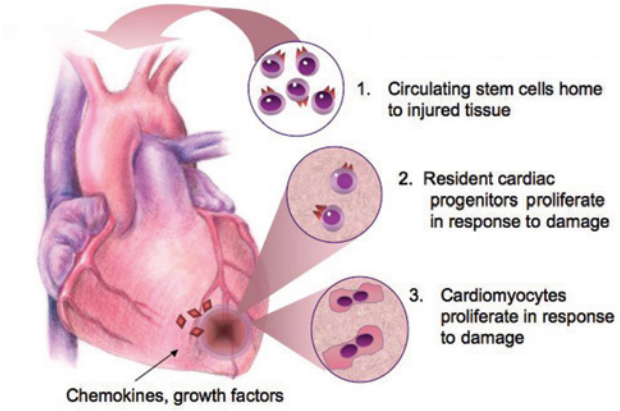
## Previous and current research

Our laboratory focuses on the biology and enhancement of muscle and cardiovascular regeneration, using the mouse as a model. Regeneration is a homeostatic mechanism evolved to maintain or restore the original architecture of a damaged tissue by recapitulating part of its original embryonic development. Despite inevitable physiological changes, mammalian organs are relatively resilient to age because they can rebuild themselves according to their original body plans. Tissues that undergo high cellular turnover such as blood and skin regenerate continually throughout adult life whereas replenishment of post-mitotic tissues such as skeletal and cardiac muscle occurs through the induced proliferation of resident progenitor cells in response to injury or chronic disease. Our focus has been to intervene in signalling mechanisms at work in the regeneration process to increase the efficiency of mammalian tissue repair. We are interested in insulin-like growth factors that attenuate muscle atrophy and improve repair in ageing, muscular dystrophy and cardiomyopathies through stem cell-mediated regeneration. Our recent applications of a local Insulin-Like Growth Factor 1 isoform (mIGF-1) to various neuromuscular and cardiovascular pathologies data implicate this growth factor as a powerful enhancer of the regeneration response, mediating the recruitment of circulating progenitor cells to sites of tissue damage and augmenting local repair mechanisms. Selective muscle fibre loss and fibrosis in ageing and diseased skeletal muscle can be blocked by transgenic or viral delivery of IGF-1, which promote recruitment of stem cells to sites of injury. Recent work has implicated both the NFκB and calcineurin-mediated signalling pathways downstream of IGF-1 in skeletal muscle regeneration. Analysis of different IGF-1 isoforms has revealed a surprising diversity of functions in muscle hypertrophy and regeneration. We have extended these findings to studies on the heart. The identification of common pathways underlying regeneration in cardiac and skeletal muscle tissue in the mouse has lent insight into intracellular signalling mechanisms that represent attractive targets for clinical intervention in various neuromuscular and cardiac diseases.

## Future projects and goals

Using a combination of molecular and cell biology with gene mutational strategies in the mouse we will continue to investigate:

- Mechanisms of muscle ageing: analysis and intervention
- Interactions of IGF-I, inflammatory and anti-fibrotic pathways in heart and skeletal muscle
- Origin and function of stem cells in heart and skeletal muscle regeneration
- Stem cell-mediated gene therapy in heart and skeletal muscle
- Development of mouse models and phenotypes for human disease




*Different modes of cardiac regeneration.*

## Selected references


Dobrowolny, G., Giacinti, C., Pelosi, L., Nicoletti, C., Winn, N., Barberi, L., Molinaro, M., Rosenthal, N. & Musaro, A. (2005). Muscle expression of a local Igf-1 isoform protects motor neurons in an ALS mouse model. *J. Cell Biol.* 168, 193-199

Musaro, A., Giacinti, C., Borsellino, G., Dobrowolny, G., Pelosi, L., Cairns, L., Ottolenghi, S., Cossu, G., Bernardi, G., Battistini, L., Molinaro, M. & Rosenthal, N. (2004). Stem cell-mediated muscle regeneration is enhanced by local isoform of insulin-like growth factor 1. *Proc. Natl. Acad. Sci. USA* 101, 1206-1210

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

Schulze, P.C., Fang, J., Kassik, K.A., Gannon, J., Cupesi, M., MacGillivray, C., Lee, R.T. & Rosenthal, N. (2005). Transgenic overexpression of locally acting insulin-like growth factor-1 inhibits ubiquitin-mediated muscle atrophy in chronic left-ventricular dysfunction. *Circ. Res.* 97, 418-426


## Cornelius Gross

PhD 1995, Yale University, USA.
Postdoctoral research at Columbia University, USA.
Group Leaser 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


Carola, V., Frazzetto, G. & Gross, C. (2006). Identifying interactions between genes and early environment in the mouse. *Genes Brain Behav.* 5, 189-199

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

Klemenhagen, K.C., Gordon, J.A., David, D.J., Hen, R. & Gross, C.T. (2006). Increased fear response to contextual cues in mice lacking the 5-HT1A receptor. *Neuropsychopharmacology* 31, 101-111

Walitza, S., Renner, T.J., Dempfle, A., Konrad, K., Wewetzer, C.h, Halbach, A., Herpertz-Dahlmann, B., Remschmidt, H., Smidt, J., Linder, M., Flierl, L., Knolker, U., Friedel, S., Schafer, H., Gross, C., Hebebrand, J., Warnke, A. & Lesch, K.P. (2005). Transmission disequilibrium of polymorphic variants in the tryptophan hydroxylase-2 gene in attention-deficit/hyperactivity disorder. *Mol. Psychiatry* 10, 1126-1132

## Liliana Minichiello

PhD, University of Naples, Italy, and The National Cancer Institute, NIH, Bethesda, MD, USA.
Postdoctoral work at EMBL Heidelberg 1993-1999.
Interdisciplinary Team Leader at EMBL Monterotondo since 2000.
Visiting Professor, Stem Cell Center, University of Lund, Sweden since 2005.


# 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.

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 synaptic plasticity, 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 PLCγ 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).

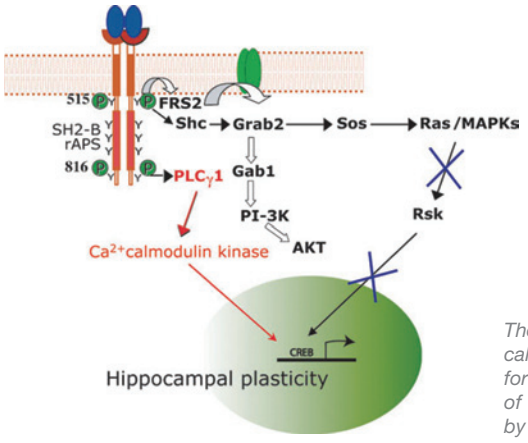
## Future projects and goals

For further comprehensive understanding of how signalling molecules interlink with each other in the formation of a transductosome 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 is inserted into the mouse gene of interest by homologous recombination in ES cells). This allows 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 determining functional and/or morphological feature of GABAergic interneurons. Selective knockouts of key molecules expressed in GABAergic interneurons are fundamental to demonstrate their functional importance. As the neurotrophins and their cognate receptors are expressed also in these cell type, in particular BDNF/TrkB, our aim is to specifically ablate trkB from GABAergic interneuros or from a particular subset of these neurons by the use of the cre-lox system. We have started generating transgenic mice expressing the cre recombinase under specific promoters in bacterial artificial chromosomes (BACs). The use of BACs increases the chance of correct temporal and spatial gene expression.

We are using a similar approach to understand the *in vivo* relevance of neurotrophins and their cognate receptors in neurodegenerative disorders like Alzheimer's and Huntington's diseases.

Our goal is to define the molecular mechanisms by which neurotrophin receptors perform their various functions in the developing and adult vertebrate nervous system.




*The PLCγ 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.*

## Selected references


von Bohlen und Halbach, O., Minichiello, L. & Unsicker, K. (2005). Haploin-sufficiency for trkB and trkC receptors induces cell loss and accumulation of α-synuclein in the substantia nigra. *FASEB J.* 19, 1740-1742

Hovelmeyer, N., Hao, Z., Kranidioti, K., Kassiotis, G., Buch, T., Frommer, F., von Hoch, L., Kramer, D., Minichiello, L., Kollias, G., Lassmann, H. & Waisman, A. (2005). Apoptosis of oligodendrocytes via fas and TNF-R1 is a key event in the induction of experimental autoimmune encephalomyelitis. *J Immunol.* 175, 5875-5884

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/PLCγ-mediated control of neuronal migration. *EMBO J.* 23, 3803-3814

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


## 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. Visiting Professor, Stem Cell Center, University of Lund, Sweden from 2005.


# Transcription factor function in development, physiology and disease

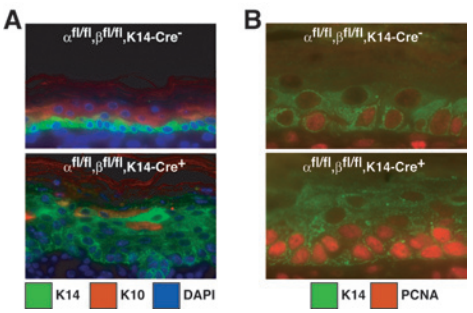
Transcription factors play important roles in regulation of cellular proliferation, differentiation, and in the function of fully differentiated cells. The main focus of the laboratory is how transcription factors link the processes of cellular differentiation and self-renewal/proliferation, both at the stem cell level and during terminal cellular differentiation. We use mouse genetics to study *in vivo* the C/EBP family of basic region-leucine zipper transcription factors. These proteins play essential roles in the development of the hematopoietic system, adipose tissues, epithelia and granulosa cells. We use conditional mutagenesis to delete one or more C/EBPs from specific cell types, and point mutagenesis to specifically alter C/EBP protein-protein interactions or post-translational modifications.

## Previous and current research

We previously defined E2F repression and interaction with the SWI/SNF complex as C/EBPα functions essential for adipose and myeloid differentiation (Pedersen et al., 2001; Porse et al., 2001). We have now found that E2F repression is required also for myeloid tumour suppression, as mice homozygous for mutations that disable C/EBPα-E2F interaction have increased myeloid progenitor proliferation and develop an acute myeloid leukemia (AML)-like disease (Porse et al., 2005). In contrast, we did not observe any effect on progenitor proliferation upon deletion of the Cdk2/Cdk4 interaction domain of C/EBPα (Porse et al., 2006). Mutations in the gene encoding C/EBPα are found in AML patients, and the most common type results in specific loss of the 42kDa C/EBPα isoform (p42), while preserving expression of the 30kDa isoform (p30). Only p42 has E2F repression activity, and when we generated p42 knockout mice we found that also these developed AML. We are now generating knockin mice with other AML-derived mutations, and investigating the role of C/EBPα mutations in the formation of self-renewing leukemic stem cells. We are currently investigating the role of C/EBPs in keratinocytes, and have observed a similar role for C/EBPs in the transition from proliferation to differentiation; however, in this case C/EBPα and C/EBPβ function redundantly, and removal of both is necessary to cause keratinocyte hyperproliferation and impair their differentiation (see figure). Other main projects of the lab involve studying the role of post-translational modifications of C/EBPs in metabolism and macrophage activation, and the transcriptional regulation of C/EBPβ.

## Future projects and goals

The future focus of the laboratory will be to elucidate the molecular mechanisms by which C/EBPs control differentiation in non-hematopoietic tissues (neurons, skin, liver), and to determine the signalling pathways that regulate C/EBPs through post-translational modification. A major effort will deal with the effects of leukemogenic mutations on hematopoietic stem cell function, in order to determine how malignant, self-renewing tumour stem cells arise.




*Defective keratinocyte differentiation and epidermal hyperplasia upon conditional ablation of C/EBPα and C/EBPβ expression in the skin. A) Skin sections from control mice (carrying floxed alleles of both C/EBPα and β, but not the Keratin 14-Cre transgene; upper panel) and C/EBPα/β double knockout mice (lower panel. As above, but with the K14-Cre transgene). The sections were stained with antibodies against the basal cell marker (Keratin 14, green), a marker for differentiating keratinocytes (Keratin 10, red) and the DNA stain DAPI (blue). The lower boundary of the epidermis is defined by the Keratin 14 positive basal cells; the outer surface of the skin is at the top of the panels. Note the expansion of the Keratin 14 expression domain and diminished Keratin 10 expression in double knockout mice. B) Skin sections as in A), stained with antibodies against Keratin 14 (green) and Proliferating Cell Nuclear Antigen (PCNA, a marker of proliferating cells; red). Note the expanded domain of cell proliferation in double knockout mice.*

## Selected references


Nerlov, C. (2004). C/EBPα 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/EBPα TBP/TFIIB and SWI/SNF recruiting domains is required for adipocyte differentiation. *Genes Dev.* 15, 3208-3216

Porse, B.T., Bryder, D., Theilgaard-Monch, K., Hasemann, M.S., Anderson, K., Damgaard, I., Jacobsen, S.E. & Nerlov, C. (2005). Loss of C/EBPα cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage. *J. Exp. Med.* 202, 85-96

Porse, B.T., Pedersen, T.A., Xu, X., Lindberg, B., Wewer, U.M., Friis-Hansen, L. & Nerlov, C. (2001). E2F repression by C/EBPα is required for adipogenesis and granulopoiesis *in vivo*. *Cell* 107, 247-258

## Walter Witke

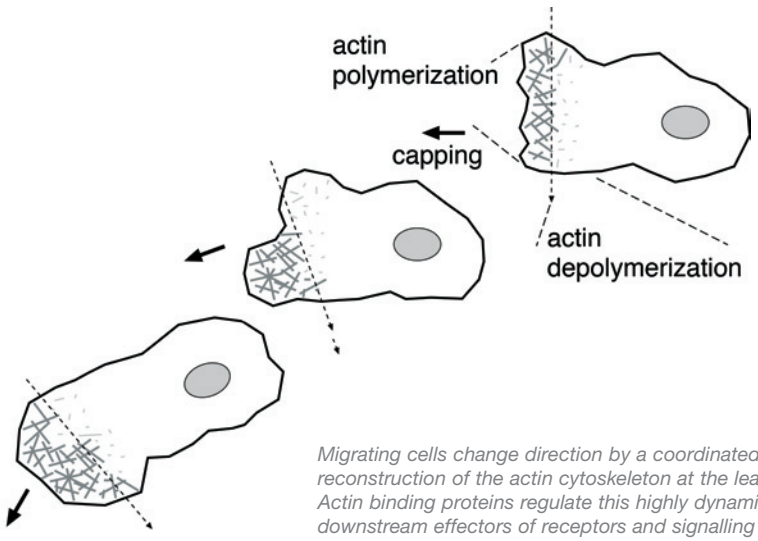
PhD 1991, University of Munich, Germany.
Postdoctoral research at the Brigham and Women's Hospital, Harvard Medical School, USA.
Group Leader at EMBL since 1996.


# Role of the actin cytoskeleton in cell migration and mouse physiology

## Previous and current research

The actin cytoskeleton is the key structure for regulating cell motility, membrane trafficking, cell polarity and cytokinesis. How the actin filament system participates in all these processes is a fundamental question in cell biology. In mammals the situation is even more complex, because different cell types have different requirements in terms of migration and motility. Furthermore, cell migration and cytokinesis during embryonic development have a different quality than, for example, cell migration in adult animals. While the first aims to build structured tissues from cells, the latter type of motility rather relates to aspects of the animal's physiology.

To serve these different requirements, cells need mechanisms to regulate the dynamics of actin filaments, and to link actin polymerisation to a specific cellular processes. A large number of actin binding proteins have been characterised by their activity to modulate actin polymerisation and filament alignment. Our group is focusing on two protein families which are thought to be essential for remodelling of the actin cytoskeleton *in vivo*. The first group of proteins – the profilins – sequester monomeric actin, catalyse the nucleotide exchange on actin monomers, and finally deliver the monomers to the growing ends of actin filaments. In the mouse, three profilin genes were identified. The second group of proteins – the actin depolymerising factors – display the opposite activity. These proteins recognise old actin filaments and depolymerise the filaments from the slow growing end. In mouse these activities are provided by the actin depolymerising factors Gelsolin, Cofilin and ADF. The main interest of our group is to understand the role of these proteins for regulating cell crawling during embryonic developments as well as later in the adult animal in the brain and the immune system.




*Migrating cells change direction by a coordinated breakdown and reconstruction of the actin cytoskeleton at the leading edge. Actin binding proteins regulate this highly dynamic process as downstream effectors of receptors and signalling cascades.*

## Selected references


Aidinis, V., Carninci, P., Armaka, M., Witke, W., Harokopos, V., Pavelka, N., Koczan, D., Argyropoulos, C., Thwin, M.M., Moller, S., Kazunori, W., Gopalakrishnakone, P., Ricciardi-Castagnoli, P., Thiesen, H.J., Hayashizaki, Y. & Kollias, G. (2005). Cytoskeletal rearrangements in synovial fibroblasts as a novel pathophysiological determinant of modeled rheumatoid arthritis. *PLoS Genet.* 1, e48

Gareus, R., Di Nardo, A., Rybin, V. & Witke, W. (2006). Mouse profilin 2 regulates endocytosis and competes with SH3 ligand binding to dynamin 1. *J. Biol. Chem.* 281, 2803-2811

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

# Index

# Notes

# Notes