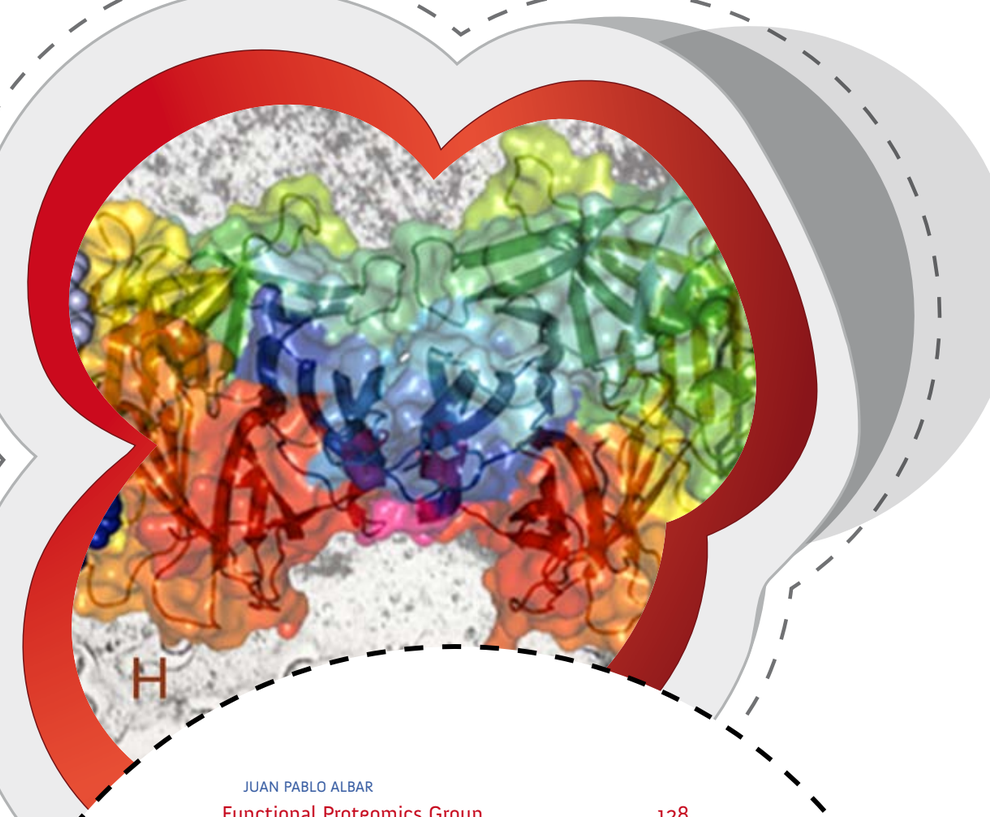
Macromolecular Structures

The Department of Structure of Macromolecules uses a variety of experimental and computational approaches to study structural and functional properties of biological macromolecules and macromolecular assemblies at different levels of complexity. Several groups are working on the analysis of cell-cell interactions (with emphasis on receptors of the immune system), and on the virus-cell interplay at distinct stages, including entry, replication, assembly and maturation. These groups work with various approaches, with special emphasis on X-ray crystallography, advanced electron and X-ray microscopy, and tomography. Cell organelles such as the centrosome, centrosomal complexes, chaperones and their cofactors, diverse nanomachines and virus-related complexes are under detailed analysis using correlative approaches, including X-ray crystallography and advanced cryo-electron 3D-microscopy methods. These studies involve extensive development of image acquisition and processing tools to achieve improved resolution. The department has incorporated nanoscopic approaches (optical and magnetic tweezers) to study properties of molecular motors involved in DNA repair and replication as well as other important macromolecular assemblies. The interplay between experimental and computational approaches is an important activity in the department. We also perform research on functional and clinical proteomics, functional characterisation of genes, proteins and protein expression in different experimental conditions and at increasing complexity levels.



Macromolecular Structures

Functional Proteomics Group

LEAD INVESTIGATOR
Juan Pablo Albar

POSTDOCTORAL SCIENTISTS
Alberto Paradela
Severine Gharbi
Miguel Marcilla

PREDOCTORAL SCIENTISTS
Salvador Martinez de Bartolomé
Rosana Navajas
Alberto Medina
Antonio Ramos
Miguel Angel López

TECHNICIANS
Virginia Pavón
Maria Dolores Segura

Proteomics is a technology-driven research area in which mass spectrometry and protein separation techniques are evolving in conjunction with powerful computational and bioinformatics tools

Advances in the standardisation of proteomic workflows and data formats are nonetheless developed at a suboptimal level. In recent years, the CNB functional proteomics group has targeted differential protein expression in a variety of tissues, cell types and organisms after various experimental treatments/conditions, and subsequent protein identification using mass spectrometry.

We have faced these proteomics challenges as follows:

1) The CAM project "Towards Functional Proteomics: an approach integrating Proteomics, Bioinformatics and Structural Biology" (CAM P2006-GEN-0166 2007-10) allowed us to partially characterise subcellular interactions between centrosomal proteins, as well as to assess macromolecular complex components by interactomics techniques based on affinity tags, stable isotopic labelling, mass spectrometry and peptide array approaches. These are considered key issues of the so-called functional proteomics.

2) Computational proteomics encompasses analysis of data from large-scale experiments and meta-annotation of proteins and protein complexes, including prediction of protein function, localisation and molecular features. The functional proteomics group has been involved in 1) application of probability-based methods for large-scale peptide and protein identification from tandem mass spectrometry data, 2) implementing methods for data mining visualization (PIKE tool, available at http://proteo.cnb.csic.es/pike), and 3) proteomics standard data formatting, reporting, storage and exchange (MIAPE generator web tool, available at http://www.proteored.org/MIAPE). The European project "International Data Exchange and Data Representation Standards for Proteomics" (ProteomeXchange) EU FP7-HEALTH-2010 and ProteoRed support most computational proteomics activities of our group.

3) Quality control, standardisation, reproducibility and robustness of proteomics workflows are issues that have been evaluated through the participation in multi-laboratory studies within the ProteoRed project (GE 2005X747_3) led by our group.

4) Clinical proteomics, including the development of protein profiling methods and biomarker discovery tools for diagnostic and prognostic purposes, is a new and rapidly evolving field. Our group leads the ISCIII project "Novel Diagnostic and Prognostic Proteins Search in Cardiovascular Diseases by Proteomic Analysis" (PI071049), in which we have applied differential quantitative proteomics using label-free and chemical labelling, mass spectrometry and bioinformatics tools to find potential plasma biomarkers associated to acute myocardial infarction. Several candidates are being validated by targeted proteomics and other techniques.

SELECTED PUBLICATIONS

Medina-Aunon JA, Paradela A, Macht M, Thiele H, Corthals G, Albar JP (2010) Protein Information and Knowledge Extractor: Discovering biological information from proteomics data. **Proteomics** 10:3262-71.

Hoogland C, O'Gorman M, Bogard P, Gibson F, Berth M, Cockell SJ, Ekefjärd A, Forsstrom-Olsson O, Kapferer A, Nilsson M, Martínez-Bartolomé S, Albar JP, Echevarría-Zomeño S, Martínez-Gomariz M, Joets J, Binz PA, Taylor CF, Dowsey A, Jones AR (2010) Minimum Information About a Proteomics Experiment (MIAPE). Guidelines for reporting the use of gel image informatics in proteomics. **Nat Biotechnol** 28:655-6.

Martínez-Bartolomé S, Blanco F, Albar JP (2010) Relevance of proteomics standards for the ProteoRed Spanish organization. **J Proteomics** 73:1061-6.

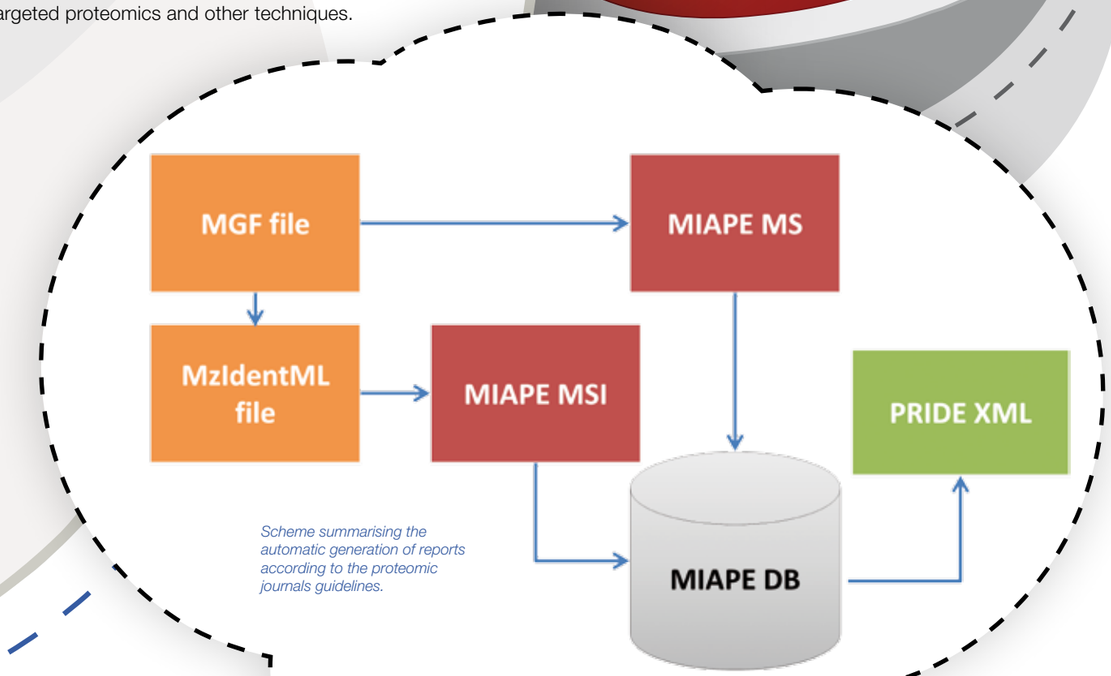
Paradela A, Marcilla M, Navajas R, Ferreira L, Ramos-Fernandez A, Fernández M, Mariscotti JF, García-del Portillo F, Albar JP (2010) Evaluation of isotope-coded protein labeling (ICPL) in the quantitative analysis of complex proteomes. **Talanta** 80:1496-502.

Martínez-Bartolomé S, Medina-Aunon JA, Jones AR, Albar JP (2010) Semi-automatic tool to describe, store and compare proteomics experiments based on MIAPE compliant reports. **Proteomics** 10:1256-60.

PATENTS

Albar JP, Mena MC, Hernando A, Lombardía M. Método para la extraccion y cuantificación de gluten nativo e hidrolizado. Número de solicitud: **P200930445.** 13 julio 2009, 15:13 (CEST). Oficina receptora: OEPM Madrid.

A Ramos Fernandez y JP Albar Ramírez. Método de identificación de péptidos y proteínas a partir de datos de espectrometría de masas. Número de patente asignado: **200930402.** 1 julio 2009, Oficina receptora: OEPM Madrid.



Scheme summarising the automatic generation of reports according to the proteomic journals guidelines.

Macromolecular Structures

LEAD INVESTIGATOR

Jose María Carazo

POSTDOCTORAL SCIENTIST

Carlos Oscar Sánchez
Roberto Marabini
Johan Busselez
Joaquín Otón
Natalia Jiménez Lozano

PREDOCTORAL SCIENTISTS

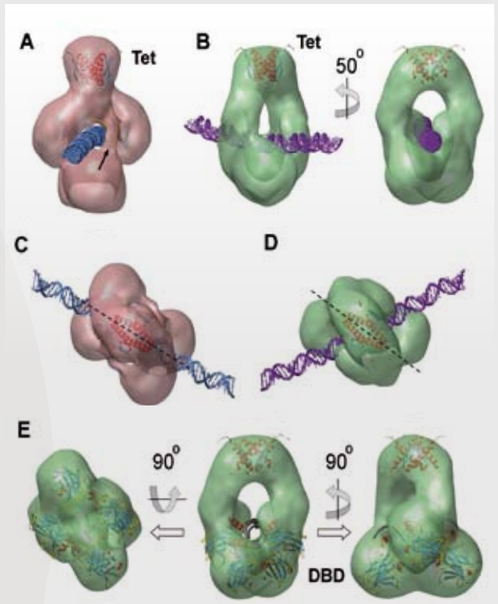
Ana Lucía Álvarez
José Ramón Macías
José Miguel de la Rosa Trevín

TECHNICIANS

Adrian Quintana
Juan José Vega
Jesus Cuenca Alba
Verónica Domingo
Blanca Benítez Silva

Biocomputing Unit

Our group focuses in the area of three-dimensional (3D) electron and X-ray microscopy, specifically developing new image processing approaches and applying them to challenging experimental biological systems. Consequently, we are a quite interdisciplinary group, where engineers, mathematicians, physicists, biologists and chemists find a common place to work.



The way electron and x-ray microscopes delivers 3D information is by a "tomographic" process by which thousands of different 2D images are combined into a 3D volume. The mathematics foundation of this combination are well routed in the area of inverse problems and are essentially the same when applied to human beings in a medical CT scanner than when applied to a macromolecular complex more than a hundred million times smaller. However, the world at the scale of the nanometer presents unique challenges of its own. Chief among them is the extraordinary flexibility of the molecular machines in charge of performing key tasks inside the cell. Indeed, their detailed composition and shape usually changes while performing their function. Obviously, we are not interested in obtaining a "blurred" 3D structure, average of the many shapes a given "nano machine" may have, but a detailed 3D structure of each of these stages. In this way we enter the world of 3D "classification" and "reconstruction", starting from thousands of very noisy images.

In order to address this problem we have developed new mathematical approaches that are able to distinguish subtle change variations at the nanometer scale, classifying the microscopic images and rendering new biological insights. Then, we port these development into our widely distributed image processing suite, named XMIPP, making them accessible as part of our committeemen in the context of the European project for strategic research infrastructures in structural biology, named Instruct.

Obviously, we are eager to test these developments into challenges biological systems, either systems we work internally at the laboratory, or working with external collaborators. In both cases the aim is to deliver new biological information through this new capability to handle conformational changes in 3D, proving "snap shot" of the actual "work" at the nano scale.

Recently we have entered the field of "Soft X-ray Tomography", expanding the approaches just described to the "meso scale", visualizing complete cells using synchrotron radiation. Indeed, Spain now has the third microscope of its class at the Spanish synchrotron ALBA, opening new perspective in biology we are actively getting involved with.

SELECTED PUBLICATIONS

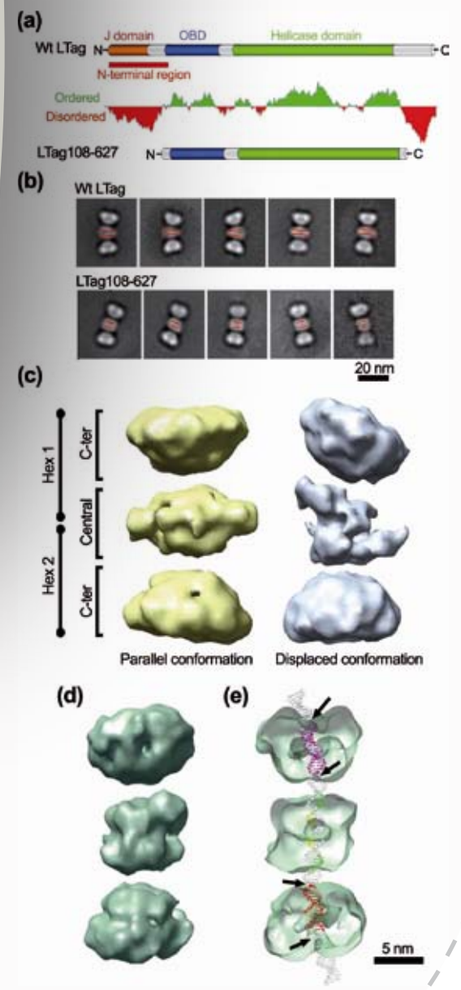
Sorzano CO, Bilbao-Castro JR, Shkolnisky Y, Alcorlo M, Melero R, Caffarena-Fernández G, Li M, Xu G, Marabini R, Carazo JM (2010) A clustering approach to multireference alignment of single-particle projections in electron microscopy. **J Struct Biol** 171:197-206.

Cuesta I, Núñez-Ramírez R, Scheres SH, Gai D, Chen XS, Fanning E, Carazo JM (2010) Conformational rearrangements of SV40 large T antigen during early replication events. **J Mol Biol** 397:1276-86.

Nickell S, Beck F, Scheres SH, Korinek A, Förster F, Lasker K, Mihalache O, Sun N, Nagy I, Sali A, Plitzko JM, Carazo JM, Mann M, Baumeister W (2009) Insights into the molecular architecture of the 26S proteasome. **Proc Natl Acad Sci USA** 106:11943-7.

Jiménez-Lozano N, Segura J, Macías JR, Vega J, Carazo JM (2009) aGEM: an integrative system for analyzing spatial-temporal gene-expression information. **Bioinformatics** 25:2566-72.

Scheres SH, Melero R, Valle M, Carazo JM (2009) Averaging of electron subtomograms and random conical tilt reconstructions through likelihood optimization. **Structure** 17:1563-72.



Macromolecular Structures

Structure of Macromolecular Complexes

LEAD INVESTIGATOR
José López Carrascosa

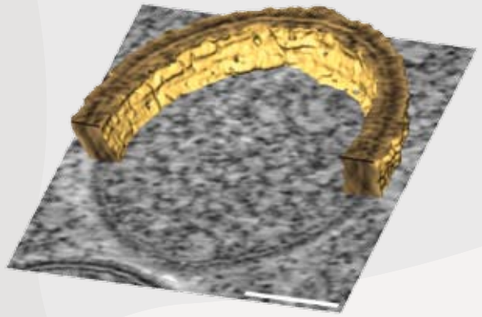
POSTDOCTORAL SCIENTISTS
Javier Chichón
Ana Cuervo
Borja Ibarra
Sonia Moreno
Maria José Rodríguez

PREDOCTORAL SCIENTISTS
Michele Chiappi
Verónica González
Maria Ibarra
Alina Ionel
Roberto Miranda

TECHNICIANS
Encarna Cebrián
Mar Pulido

The group is working on the study of virus assembly and maturation using an integrative approach of microscopic, biochemical and biophysical methods.

Our aim is to determine the molecular bases of the interaction of macromolecules and macromolecular complexes to yield functional biological machineries. We have centred our interest on the way viral proteins assemble into intermediate structures that are further matured to yield fully infective viral particles. Our model systems are bacteriophages (T7 and Φ29), and eukaryotic viruses such as vaccinia. Using bacteriophages, we have reached subnanometer resolution by electron cryo-microscopy and three-dimensional reconstruction. The combination of these structures with computational modelling has led to the definition of the molecular basis of capsid expansion and stabilization characteristic of certain virus families.



Tomographic reconstruction of a maturation intermediate of vaccinia within the infected cell. A plane of the tomographic reconstruction (gray levels) is coupled to a section of t he tomographic reconstruction of the viral envelope.

The intracellular maturation of eukaryotic viruses is currently studied using vaccinia and other complex viruses. We are following the rearrangements involved in the sequential generation of intermediate morphogenetic particles within the cytoplasm by combining electron cryo-microscopy and low temperature processing methods to retrieve three-dimensional tomographic reconstructions of virus-infected cells in a native preservation state. The tomographic volumes are used as templates to map by computer modelling other structural data, from x-rays or single particle cryo-microscopy, to yield high resolution topographic maps. As electron microscopes have limited penetration power, extension of the use of tomographic methods to the cell environment demands other types of microscopy. We are developing x-ray microscopy methods to obtain three-dimensional cryo-tomographic reconstructions of whole cells at intermediate resolutions between electron and light microscopy. These developments are coordinated with the set-up of the microscope beam line at the Spanish synchrotron ALBA.

We are also currently exploring the potential of combining different biophysical methods for the analysis of biological material at the nanoscopic level. We have worked intensively on the use of atomic force microscopy to study the nanomechanical properties of viral capsids, and the structure of other protein-nucleic acids complexes. Together with the group of Jose M. Valpuesta (CNB), we are developing optical and magnetic tweezers to study molecular motors at the single molecule level, as well as to explore their potential for cell organelle manipulation.

SELECTED PUBLICATIONS

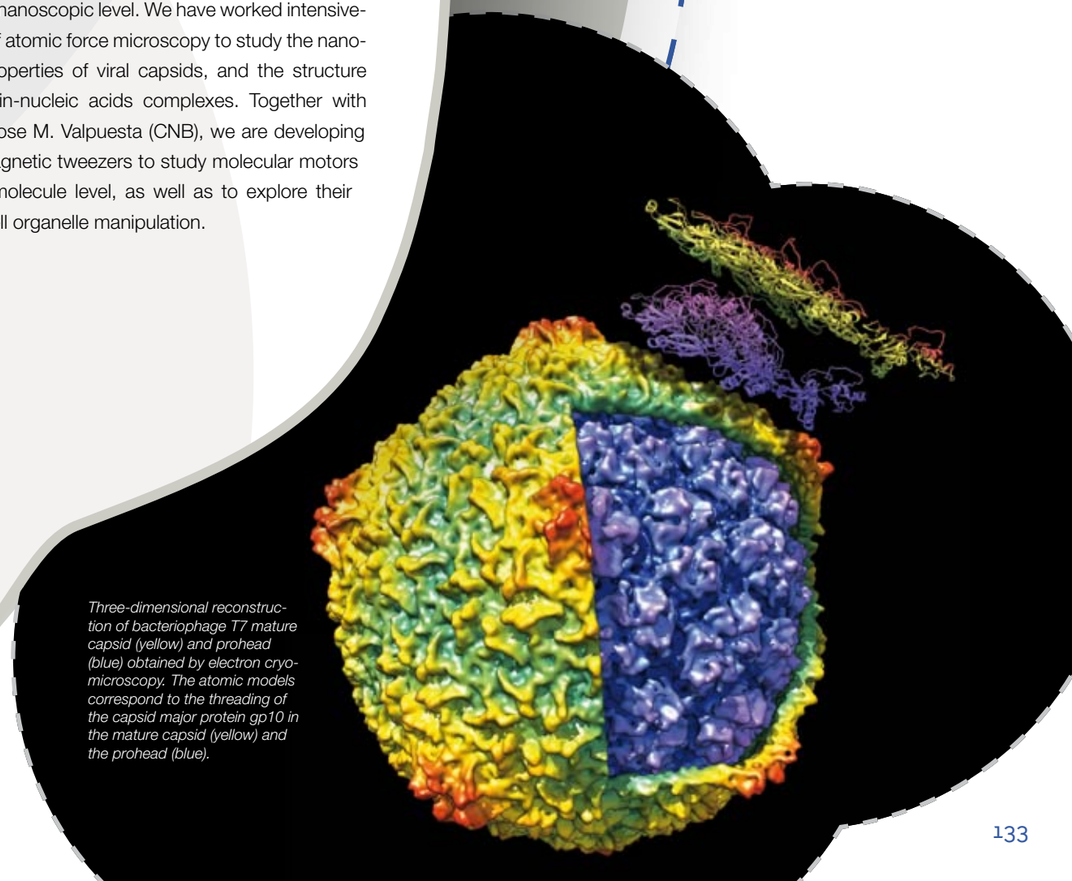
Carrascosa JL, Chichón FJ, Pereiro E, Rodríguez MJ, Fernández JJ, Esteban M, IEM S, Guttmann P, Schneider G (2009) Cryo-X-ray tomography of vaccinia virus membranes and inner compartments. **J Struct Biol** 168:234-239.

Chichón FJ, Rodríguez MJ, Risco C, Fraile-Ramos A, Fernández JJ, Esteban M, Carrascosa JL (2009) Membrane remodeling during vaccinia virus morphogenesis. **Biol Cell** 101:401-404.

Hormeño S, Ibarra B, Chichón FJ, Habermann K, Lange BMH, Valpuesta JM, Carrascosa JL, Arias-González JR (2009) Single Centrosome Manipulation Reveals Its Electric Charge and Associated Dynamic Structure. **Biophys J** 97:1022-1030.

Arechaga I, Martínez-Costa OH, Ferreras C, Carrascosa JL, Aragón JJ (2010) Electron microscopy analysis of mammalian phosphofructokinase reveals an unusual 3-dimensional structure with significant implications for enzyme function. **FASEB J** 24:4960-4968.

Luque D, González JM, Garriga D, Ghabrial SA, Havens WM, Trus B, Verdaguer N, Carrascosa JL, Castón JR (2010) The T=1 capsid protein of *Penicillium chrysogenum* virus is formed by a repeated helix-rich core indicative of gene duplication. **J Virol** 84:7256-7266.



Three-dimensional reconstruction of bacteriophage T7 mature capsid (yellow) and prohead (blue) obtained by electron cryo-microscopy. The atomic models correspond to the threading of the capsid major protein gp10 in the mature capsid (yellow) and the prohead (blue).

Macromolecular Structures

LEAD INVESTIGATOR
José M. Casasnovas

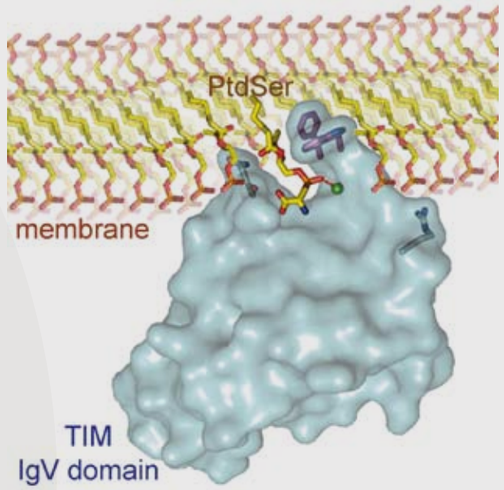
POSTDOCTORAL SCIENTIST
César Santiago

PREDOCTORAL SCIENTISTS
Ángela Ballesteros
Meriem Echbarthi
Gaurav Mudgal

TECHNICIAN
Nuria Cubells

Cell-Cell and Virus-Cell Interactions

A large variety of glycosylated proteins that participate in cell-cell and virus-cell interactions populate the cell and viral membranes. The surface proteins allow the cell to communicate with its environment and participate in cell migration and in virus entry processes.



Our group analyses receptor-ligand recognition using biochemical techniques and X-ray crystallography, studying receptors of the immune system, some of which have been subverted by viruses to enter host cells. Our research also focuses on receptor-mediated virus entry into cells, with viruses of medical interest that use the receptor molecules for attachment and cell membrane penetration. We study receptor recognition and the dynamics of virus entry with rhinovirus, measles virus and coronavirus. Our research provides new knowledge on cell adhesion and virus entry processes, and it can lead to the development of anti-viral therapies and treatments for immune disorders. Below we highlight some of our recent results, related to both immune processes and viral infections.

TIM proteins: a family of PtdSer receptors that regulate immunity

The T cell/transmembrane, immunoglobulin and mucin domain (TIM) gene family plays a critical role in regulating immune responses, including transplant tolerance, autoimmunity, the regulation of allergy and asthma, and the response to viral infections. We demonstrated that the TIM proteins are pattern recognition receptors, specialized in recognition of the phosphatidylserine (PtdSer) cell death signal. The figure on the left shows a surface representation of the structure of the mTIM-4 IgV domain bound to PtdSer in a lipid membrane bilayer, which was determined in our laboratory. Cells expressing TIM-1, TIM-3 and TIM-4 can mediate elimination of apoptotic cells, which display PtdSer on their outer membrane leaflet. This biological process is essential for maintaining tissue homeostasis and prevention of autoimmunity and inflammatory reactions.

Measles virus binding to the CD46 cell receptor

Measles virus (MV) has a glycoprotein, known as haemagglutinin (H), anchored to its lipid envelope. The H protein is specialized in the recognition of cell surface receptor molecules such as CD46; this process is essential for virus entry into host cells. We have been able to visualize this molecular event by determining the crystal structure of the measles virus H protein bound to the CD46 receptor protein. The image on the right shows a surface representation of the dimeric H protein structure (rainbow colours) bound to two CD46 molecules (C alpha trace in blue). An electron microscopy image of measles virus particles entering into host cells is shown below the structural representation. Binding to the CD46 receptor molecule is essential for MV entry into human cells and it begins the virus infective cycle.

SELECTED PUBLICATIONS

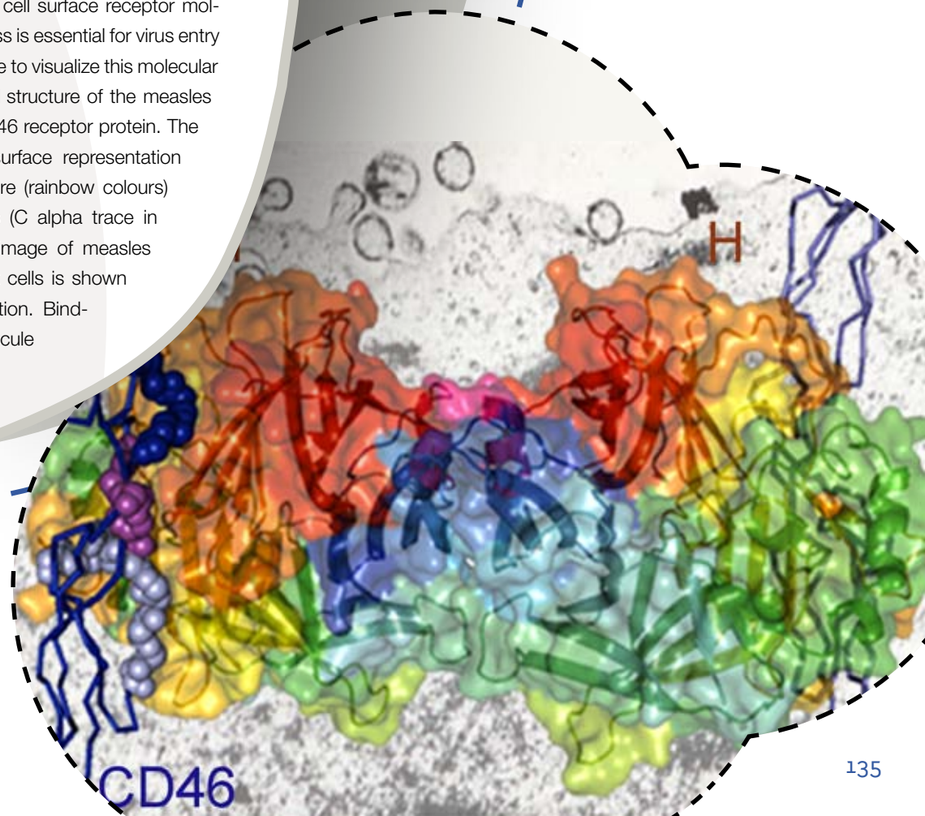
Santiago C, Celma ML, Stehle T, Casasnovas JM (2010) Structure of the measles virus hemagglutinin bound to the CD46 receptor. **Nat Struct Mol Biol** 17124-9.

DeKruyff RH, Bu X, Ballesteros A, Santiago C, Chim YL, Lee HH, Karisola P, Pichavant M, Kaplan GG, Umetsu DT, Freeman GJ, Casasnovas JM (2010) T Cell/Transmembrane, Ig, and Mucin-3 Allelic Variants Differentially Recognize Phosphatidylserine and Mediate Phagocytosis of Apoptotic Cells. **J Immunol** 184:1918-30.

Freeman GJ, Casasnovas JM, Umetsu DT, DeKruyff RH (2010) TIM genes: a family of cell surface phosphatidylserine receptors that regulate innate and adaptive immunity. **Immunol Rev** 235:172-89.

Persson BD, Schmitz NB, Santiago C, Zocher G, Larvie M, Scheu U, Casasnovas JM, Stehle T (2010) Structure of the Extracellular Portion of CD46 Provides Insights into Its Interactions with Complement Proteins and Pathogens. **PLoS Pathog** 6:e1001122.

Stehle T, Casasnovas JM (2009). Specificity switching in virus-receptor complexes. **Curr Opin Struct Biol** 19:181-188.



Macromolecular Structures

Structural Biology of Viral Macromolecular Assemblies

Our studies attempt to elucidate structure-function-assembly relationships of viral macromolecular assemblies by three-dimensional cryo-EM in combination with X-ray structures.

LEAD INVESTIGATOR

José Ruiz Castón

POSTDOCTORAL SCIENTISTS

Daniel Luque Buzo
Nerea Irigoyen Vergara

PREDOCTORAL SCIENTISTS

Elena Pascual Vega
Josué Gómez Blanco
Mariana Castrillo Briceño

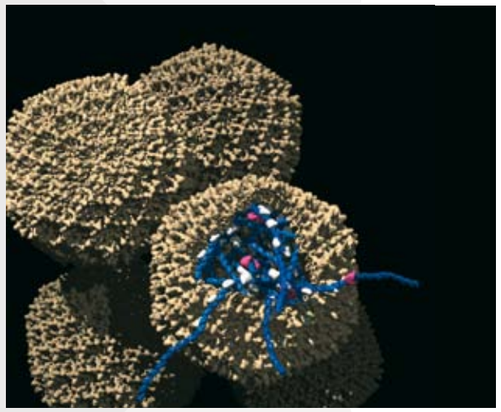
VISITOR

Nuria Verdaguer Massana

We are investigating several viral systems with different levels of complexity: double-stranded (ds)RNA viruses such as infectious bursal disease virus (IBDV) and *Penicillium chrysogenum* virus, and virus-like particles of a single-stranded RNA virus, rabbit hemorrhagic disease virus. Understanding the structural basis of the polymorphism of the coat proteins has been our major goal. The general rules that govern coat protein conformational flexibility can be extended to other macromolecular assemblies that control fundamental processes in biology. We also focus on the structural basis of dsRNA virus replication. All dsRNA viruses, from the mammalian reoviruses to the bacteriophage phi6, share a specialised capsid involved in transcription and replication of the dsRNA genome.

The IBDV has a T=13l capsid built of a single protein, VP2 (441 residues). VP2 is initially synthesised as a 512-residue precursor, pVP2. The conformational switch responsible for the VP2 structural polymorphism resides in the segment 443-453, which forms an amphipathic helix. The other major protein, VP3, participates in capsid assembly as a scaffolding protein. This process is mediated by electrostatic interactions of the VP3 acidic C-terminal region with the basic side of the amphipathic helix. In addition, VP2 has Asp431-mediated endopeptidase activity that does not disrupt the Ala441-Phe442 bond. Our structural analysis also showed an unusual feature for IBDV: the cargo space of the capsid is much larger than is necessary to enclose a single genome copy. Indeed, IBDV is an icosahedral polyploidy dsRNA virus that can package more than one complete genome copy. Multiploid IBDV particles propagate with higher efficiency than haploid virions. VP3 is also an RNA-binding protein and is closely associated with the viral genome, forming ribonucleoprotein (RNP) complexes. VP3 renders these RNP less accessible to nucleases. RNP complexes are functionally competent for RNA synthesis in a capsid-independent manner.

We analysed the capsid structure of a fungal dsRNA virus, *Penicillium chrysogenum* virus. This capsid shows that the asymmetric unit is formed by a repeated helical core, indicative of gene duplication. This basic repeated motif could provide insight into an ancestral fold and show structural evolutionary relationships of the dsRNA virus lineage. Our results indicate that a dimer as the asymmetric unit is a conserved arrangement favourable for managing replication and genome organization of dsRNA viruses.



Infectious bursal disease virus (IBDV), an avian pathogen, is a non-enveloped icosahedral virus that has a bisegmented dsRNA genome (blue) enclosed within a single-layered capsid with T=13l geometry (orange). dsRNA is bound to a nucleocapsid protein (VP3, white) and the RNA polymerase (VP1, pink). IBDV is the first icosahedral virus to be described that can package more than one complete genome copy.

SELECTED PUBLICATIONS

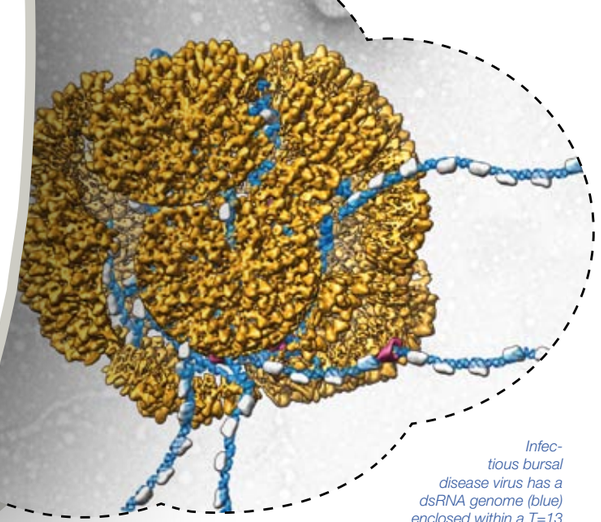
Luque D, González JM, Garriga D, Ghabrial SA, Havens WM, Trus BL, Verdaguer N, Carrascosa JL, Castón JR (2010) The T=1 capsid protein of Penicillium chrysogenum virus is formed by a repeated helix-rich core indicative of gene duplication. **J Virol** 84:7256-7266.

Saugar I, Irigoyen N, Luque D, Carrascosa JL, Rodríguez JF, Castón JR (2010) Electrostatic interactions between capsid and scaffolding proteins mediate the structural polymorphism of a double-stranded RNA viruavs. **J Biol Chem** 285:3643-3650.

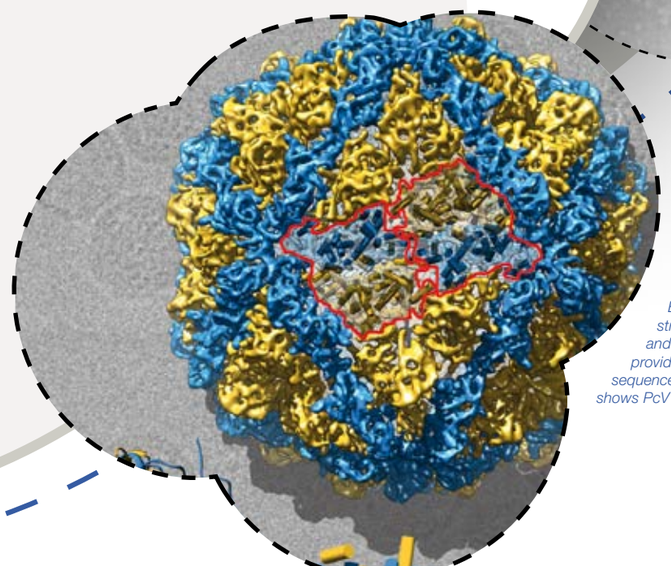
Luque D, Saugar I, Rejas MT, Carrascosa JL, Rodríguez JF, Castón JR (2009) Infectious Bursal Disease Virus: ribonucleoprotein complexes of a double-stranded RNA virus. **J Mol Biol** 386:891-901.

Luque D, Rivas G, Alfonso C, Carrascosa JL, Rodríguez JF, Castón JR (2009) Infectious Bursal Disease Virus is an icosahedral polyploid dsRNA virus. **Proc Natl Acad Sci USA** 106:2148-2152.

Irigoyen N, Garriga D, Verdaguer N, Rodríguez JF, Castón JR (2009) Autoproteolytic activity derived from the infectious bursal disease virus capsid protein. **J Biol Chem** 284:8064-8072.



Infectious bursal disease virus has a dsRNA genome (blue) enclosed within a T=13 capsid (yellow). The genome, organised as ribonucleoprotein complexes bound to a nucleocapsid protein (white) and the RNA polymerase (pink), can be released by mild disruption of viral particles (background) and is functionally competent for RNA synthesis.



Structure of Penicillium chrysogenum virus (PcV), a fungal double-stranded RNA virus. Three-dimensional cryo-EM reconstruction of the PcV virion at 8 Å resolution, which has a T=1 capsid formed by 60 copies of a single polypeptide. Boundaries for two-capsid proteins are outlined in red and their secondary structural elements are highlighted; each subunit has two similar domains (yellow and blue), suggesting ancestral gene duplication. This basic repeated motif could provide insight into an ancestral fold and αJoined motifs, with the same fold but lacking sequence similarity, have been described in other capsid proteins. The background shows PcV particles frozen in vitreous ice

Macromolecular Structures

Computational Methods for 3D Electron Microscopy

LEAD INVESTIGATOR
José Jesús Fernández

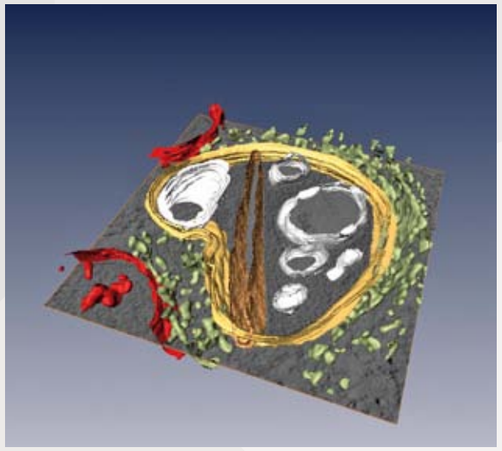
POSTDOCTORAL SCIENTIST
María del Rosario Fernández Fernández

PREDOCTORAL SCIENTISTS
José Ignacio Agulleiro Baldó
Antonio Martínez Sánchez

Knowledge of the structure of biological specimens is essential to understanding their functions at all scales.

Electron microscopy (EM) combined with image processing allows the investigation of the three-dimensional (3D) structure of biological specimens over a wide range of sizes, from cell structures to single macromolecules, providing information at different levels of resolution. Depending on the specimen under study and the structural information sought, different 3D EM approaches are used. Single particle EM makes it possible to visualize macromolecular assemblies at subnanometer or up to near-atomic resolution. Electron tomography turns out to be a unique tool for deciphering the molecular architecture of the cell. In all cases, the computational methods of image processing play a major role. Computational advances have contributed significantly to the current relevance of 3D EM within structural biology.

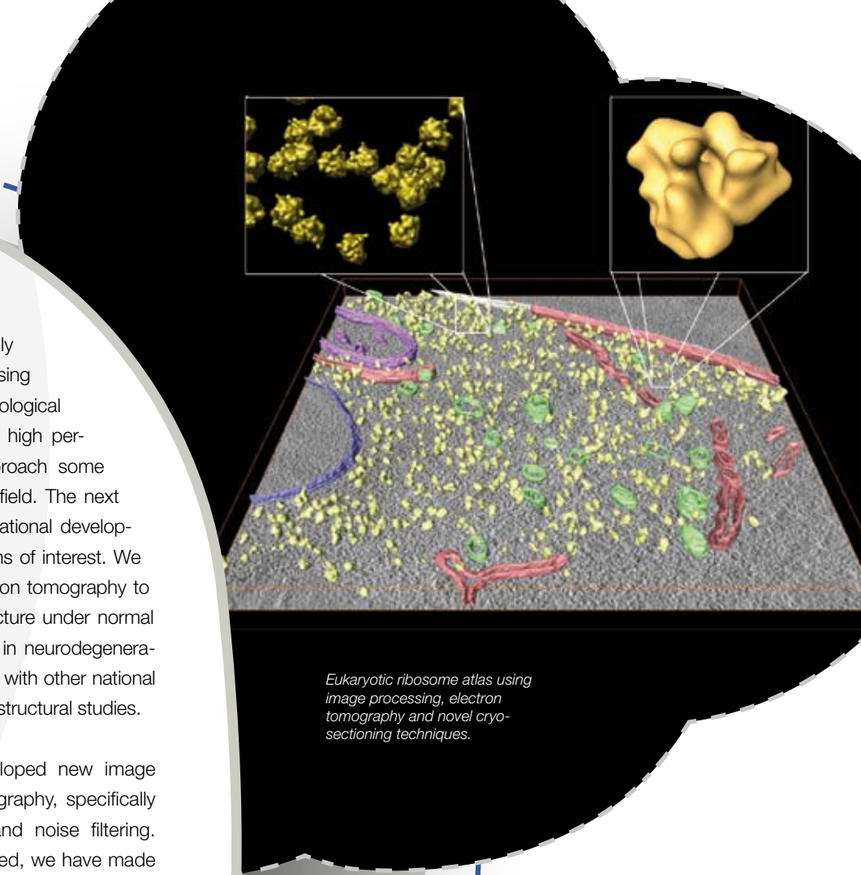
Rubella virus factory visualized in three dimensions using electron tomography.



Our research interests are focused mainly on the development of image processing methods for structural analysis of biological specimens by 3D EM. We also devise high performance computing strategies to approach some of the computational challenges in this field. The next step is the application of these computational developments to the study of biological problems of interest. We are interested in the application of electron tomography to explore alterations in subcellular architecture under normal and pathological conditions, particularly in neurodegenerative diseases. In addition, we collaborate with other national and international groups in experimental structural studies.

In the last few years, we have developed new image processing methods for electron tomography, specifically for three-dimensional reconstruction and noise filtering. As far as structural studies are concerned, we have made significant advances towards the construction of an atlas of the eukaryotic ribosome using image processing, electron tomography and novel cryosectioning techniques, in collaboration with the Carrascosa and Peters groups. We have also collaborated with the Risco group in the structural characterization of rubella virus factories by electron tomography.

Dr. M. R. Fernández Fernández recently joined the group. She worked previously on different aspects of the molecular and cellular biology of Huntington's disease neurodegeneration. We are now conducting a project focused on the use of electron tomography to understand structural alterations at the subcellular level in Huntington's disease.



Eukaryotic ribosome atlas using image processing, electron tomography and novel cryosectioning techniques.

SELECTED PUBLICATIONS

Fernández JJ (2009) TOMOBFLOW: Feature-preserving noise filtering for electron tomography. **BMC Bioinformatics** 10:178.

Pierson J, Fernández JJ, Bos E, Amini S, Gnaegi H, Vos M, Bel B, Adolfsen F, Carrascosa JL, Peters PJ (2010) Improving the technique of vitreous cryo-sectioning for cryo-electron tomography: Electrostatic charging for section attachment and implementation of an anti-contamination glove box. **J Struct Biol** 169:219-225.

Fontana J, López-Iglesias C, Tzeng WP, Frey TK, Fernández JJ, Risco C (2010) Three-dimensional structure of Rubella virus factories. **Virology** 405:579-591.

Fernández-Fernández MR, Ferrer I, Lucas JJ (2011) Impaired ATF6α processing, decreased Rheb and neuronal cell cycle re-entry in Huntington's disease. **Neurobiol Dis** 41:23-32.

Agulleiro JI, Fernández JJ (2011) Fast tomographic reconstruction on multicore computers. **Bioinformatics** 27:582-3.

Macromolecular Structures

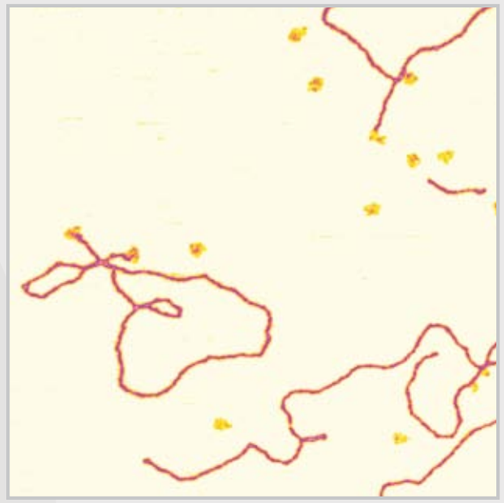
Molecular Biophysics of DNA Repair Nanomachines

LEAD INVESTIGATOR
Fernando Moreno Herrero

POSTDOCTORAL SCIENTIST
Carolina Carrasco Pulido

PREDOCTORAL SCIENTISTS
Benjamin Gollnick
María Eugenia Fuentes Pérez

The molecular biophysics group joined the CNB in 2009 to develop single-molecule techniques and to study the mechanisms of molecular motors and DNA repair protein machines.



AFM image showing AddAB proteins (yellow) and DNA molecules (red). Often AddAB appears bound at the end of a DNA molecule. Image size, 1 micrometer.

DNA breaks are a potentially catastrophic form of DNA damage that can lead to cell death, premature ageing or cancer. Consequently, nature has developed robust mechanisms to repair DNA breaks that involve dynamic, large-scale DNA manipulations by molecular motors and protein machines. Understanding how these machines work in model organisms will underpin a deeper knowledge of the basic enzymology of DNA repair, and provide important insights into the molecular basis of carcinogenesis.

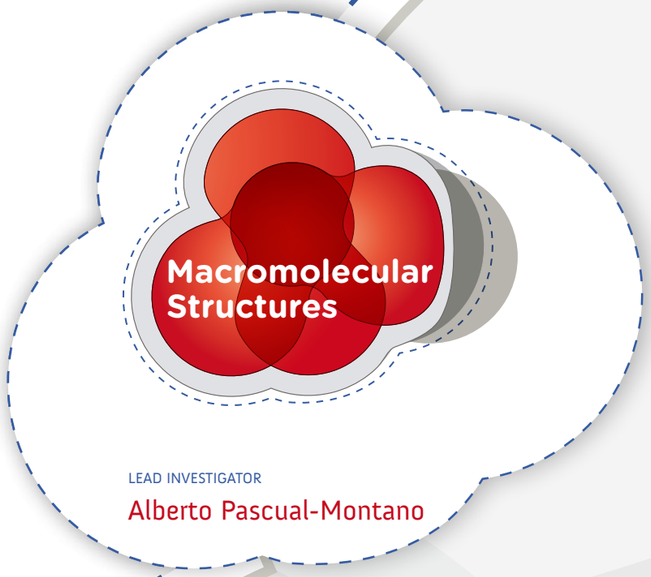
Over the last two years, a large effort has been made to set up the techniques and to establish research lines that are now starting to give results. During this short time, we have built a magnetic tweezers (MT) instrument that can manipulate single DNA molecules and measure force and torque applied by molecular motors. The group will also have a custom-adapted atomic force microscope (AFM) and an optical tweezers (OT), whose construction is expected to be finished in 2011. We follow two main research lines, supported by a solid collaboration with the group of Dr. M.S. Dillingham at Univ. Bristol, i) to study the dynamics and mechanisms of the helicase-nuclease AddAB from Bacillus subtilis and ii) to study the mechanisms of cohesion and condensation of DNA by structural maintenance of chromosome proteins. Using AFM and MT, we have been able to monitor in real time the translocation of AddAB –a molecular motor involved in DNA-end processing in homologous recombination– and to visualise individual proteins in the act of moving along a single DNA molecule. Surprisingly, we found that the coupling between translocation and unwinding in a helicase is a complex parameter and that unwinding is promoted by intermediates that differ from the classical "Y" shape. Moreover, the helicase activity of AddAB was activated by encounter with a specific sequence in the DNA track and that generated a loop structure that we captured using our single-molecule techniques.

SELECTED PUBLICATIONS

Yeeles JTP, Dillingham MS, Moreno-Herrero F (2010) Atomic Force Microscopy Shows that Chi Sequences and SSB Proteins Prevent DNA Reannealing Behind the Translocating AddAB Helicase-Nuclease. **Biophys J** 98:65a.

Yeeles J, Longman E, Moreno-Herrero F, Dillingham MS (2010) Activation of a Helicase Motor Upon Encounter With a Specific Sequence in the DNA Track. **Biophys J** 98:66a.

AFM image showing AddAB caught in the act of translocating along DNA. Unwound strands immediately re-anneal behind the translocating motor. Dark structures are single-stranded DNA with bound AddAB proteins. Image size, 1 micrometer.



Functional Bioinformatics

LEAD INVESTIGATOR
Alberto Pascual-Montano

POSTDOCTORAL SCIENTIST
Rubén Nogales-Cadenas

PREDOCTORAL SCIENTISTS
Mariana Lara Neves
Edgardo Mejía-Roaz

To help in understanding the biology that underlies experimental settings, our group is dedicated to the development of new methods and analytic techniques to solve specific biological questions.

We concentrate on functional bioinformatics, whose focus is on the functional characterization of genes and proteins in different experimental conditions. We also develop new methods for the analysis and interpretation of biological data, and centre on three major areas: analysis of gene expression information, functional analysis of annotations and scientific literature analysis.

For gene expression analysis, we have developed several techniques to bicluster data (find functional patterns of small set of genes that show coherent expression in a small subset of experimental conditions). This development is based a novel matrix factorization approach that can produce not only a clustering the genes and experimental conditions simultaneously, but also aid in its interpretation.

In this line, we have also developed several new techniques for the functional characterization of lists of genes or proteins. The novelty of our proposal lies in the combination of several information sources and determination of which combination of functional annotations is significantly enriched in the gene or protein list. This has opened a new research area known as modular functional enrichment.

These studies are complemented by one of the richest sources of biological information currently available: the scientific literature. We have developed several methods and tools with which to analyse the set of Medline abstracts related to lists of genes and proteins, and to summarize its content into semantic features that can later be integrated into functional analysis. In this way, a global understanding of biological events is possible.

A large set of high-quality bioinformatics software has also been developed, published and made available to the scientific community.The figure summarises the developments of our group in functional bioinformatics. More details can be found at http://bioinfo.cnb.csic.es

SELECTED PUBLICATIONS

Vazquez M, Nogales-Cadenas R, Arroyo J, Botías P, Raul García, Carazo JM, Tirado F, Pascual-Montano A, Carmona-Saez P (2010) MARQ: an online tool to mine GEO for experiments with similar or opposite gene expression signatures. **Nucleic Acids Res** 38 ppl:W228-32.

Medina I, Carbonell J, Pulido L, Madeira S, Gotees S, Conesa A, Tarrasa J, Pascual-Montano A, Nogales-Cadenas R, Santoyo-Lopez J, Garcia-Garcia F, Marba M, Montaner D, Dopazo J (2010) Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. **Nucleic Acids Res** 38 Suppl:W210-213.

Neves ML, Carazo JM, Pascual-Montano A (2010) Moara: a Java library for extracting and normalizing gene and protein mentions. **BMC Bioinformatics** 11:157.

Nogales-Cadenas R, Carmona-Saez P, Vazquez M, Vicente C, Yang X, Tirado F, Carazo JM, Pascual-Montano A (2009) GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. **Nucleic Nucleic Acids Res** 37:W317-22.

Vazquez M, Nogales-Cadenas R, Chagoyen M, Carmona-Saez P, Pascual-Montano A (2009) SENT: Semantic Features in Text. **Nucleic Acids Res** 37:W153-159.

Cell Structure Lab

Macromolecular Structures

LEAD INVESTIGATOR
Cristina Risco Ortiz

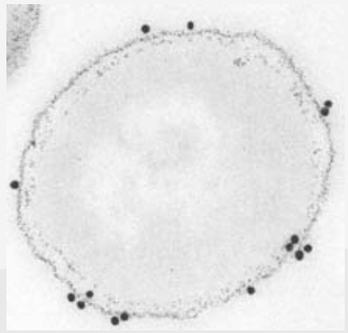
POSTDOCTORAL SCIENTIST
Elia Diestra Villanueva

PREDOCTORAL SCIENTISTS
Noelia López Montero
Laura Sanz Sánchez
Isabel Fernández de Castro Marín
Juan Fontana Jordán de Urríes

TECHNICIAN
Eva Sanmartín Conesa

UNDERGRADUATE STUDENT
Alejandra Sánchez Pineda

VISITORS
Nebraska Zambrano
Alejandro Sánchez

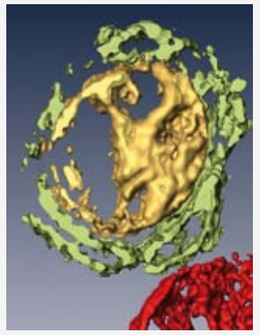


Viral factories are complex structures built by many different viruses as a physical scaffold for viral replication and morphogenesis.

Signalling events involved in the assembly of viral factories are mainly unknown. We are interested in characterizing the factories of several RNA viruses that are important pathogens for humans, simultaneously using viruses as very valuable tools to study cell architecture. One of our main goals is to develop new methods for specific detection of macromolecules in 3D maps of whole cells. This will permit progress in one of today's most ambitious challenges in structural biology: elaboration of molecular atlases of cells to understand the structure-function relationships that underlie cell functions. Within this context, our activities in 2009 and 2010 are summarized as follows:

- Correlative microscopy methods helped us to understand the dynamics of arbovirus infection in mosquito cells. We detected new anti-viral mechanisms that maintain infection under control in these cells (López-Montero and Risco, Cell Microbiol 2010).

- 3D maps of factories built by rubella virus have been obtained by electron tomography and reveal how cell organelles are modified and interact with each other in the factory scaffold (Fontana et al., Virology 2010). This work was done in collaboration with Dr. José J. Fernández in the Department of Structure of Macromolecules (CNB).



RUBV factory in 3D as studied by electron tomography (right) and detection of Hfq protein in E. coli with the MT clonable tag (left). Hfq molecules (small particles) are detected in the inner membrane, outer membrane (immunogold-labelled, large particles) and the periplasm between them.

- After validating the first clonable tag for electron microscopy that works in live cells, the "GFP" of ultrastructural analysis, we applied this new methodology to the study of proteins in bacteria (Diestra et al., J. Struct. Biol. 2009; Diestra et al., PLoS ONE 2009) and eukaryotic cells (Risco and Sanmartín, patent, 2010). This new technology based on the metal-binding protein metallothionein (MT) could be an important step that will provide us with a completely new vision of structure-function relationships in complex biological systems.

We are currently studying 1) the nature of contacts between cell organelles in viral factories, 2) how membrane-associated arrays of viral polymerases work and release the replicated viral genomes, 3) how replication and assembly are spatially connected inside viral factories, 4) the major structural changes triggered in immature viral particles to become infectious virions inside the factories, and 5) applications of new synthetic tags for correlative microscopy and molecular mapping in 3D EM.

SELECTED PUBLICATIONS

Diestra E, Fontana J, Guichard P, Marco S, Risco C (2009) Visualization of proteins in intact cells with a clonable tag for electron microscopy. **J Struct Biol** 165:157-68.

Diestra E, Cayrol B, Arluison V, Risco C (2009) Cellular electron microscopy imaging reveals membrane localization of the bacterial Hfq protein. **PLoS One** 4:e8301.

Chichón J, Rodríguez MJ, Risco C, Fernández JJ, Fraile-Ramos A, Esteban M, Carrascosa JL (2009) Membrane remodelling during vaccinia virus morphogenesis. **Biol Cell** 101:401-414.

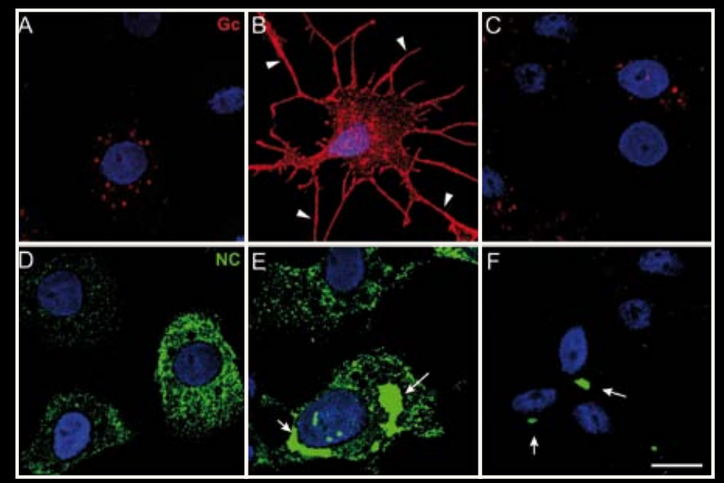
Fontana J, López-Iglesias C, Tzeng W-P, Frey TK, Fernández JJ, Risco C (2010) Three-dimensional structure of Rubella virus factory. **Virology** 405:579-591.

López-Montero N, Risco C (2011) Self-protection and survival of arbovirus-infected mosquito cells. **Cell Microbiol** 13:300-15.

PATENT

Cristina Risco and Eva Sanmartín. Marcador clonable para su uso en microscopía. Patent application: **P201031880** (Dec (2010). CSIC

Changes in the distribution of Gc (red) and NC (green) viral proteins in Bunyavirus-infected C6/36 mosquito cells. NC protein is progressively trapped in N-bodies (white arrows in F).



Macromolecular Structures

LEAD INVESTIGATOR
Carmen San Martín Pastrana

POSTDOCTORAL SCIENTIST
Ana J. Pérez Berná

PREDOCTORAL SCIENTISTS
Rosa Menéndez Conejero
Gabriela N. Condezo Castro
Alvaro Ortega Esteban

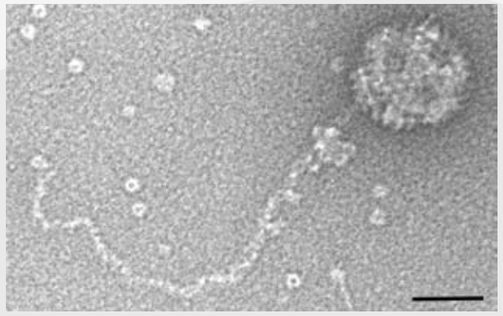
UNDERGRADUATE STUDENTS
Jorge Pablo Arribas Miguélez
Almudena Chaves Pérez

TECHNICIANS
María López Sanz
María Angeles Fernández Estévez

VISITOR
Inna Romanova

Structural and Physical Determinants of Viral Assembly

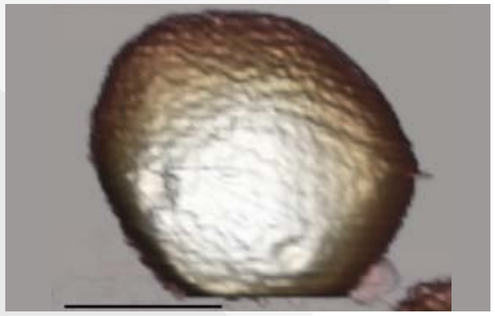
We are interested in the structural and physical principles that govern assembly and stabilization of complex viruses.



Two ways of analysing adenovirus capsid stability. Left, electron microscopy; right, atomic force microscopy (in collaboration with PJ de Pablo, Univ Autónoma de Madrid). Bars represent 50 nm.

As a model system we use adenovirus, a challenging specimen of interest both in basic virology and nanobiomedicine. We approach the problem from an interdisciplinary point of view, combining biophysics, computational, structural and molecular biology techniques. Sharing expertise and resources is a pillar in our work; accordingly, we maintain a variety of intra- and extramural collaborations that enrich the development and interpretation of our work.

Adenoviruses are pathogens of clinical relevance in the increasingly large immunocompromised population. They are also widely used as vectors for gene therapy, vaccination and oncolysis. The viral particle is composed of more than 10 distinct proteins plus the dsDNA viral genome, for a total molecular weight of 150 MDa. During the 2009-2010 period, we determined the structural differences between the mature and immature adenovirus particle. A thermosensitive mutant that does not package the viral protease produces viral particles containing unprocessed precursors of several capsid and core proteins. These virions, which represent the immature assembly intermediate, are not infective due to a defect in entry and uncoating. Subnanometer resolution difference maps calculated between mature and immature virus particles, and interpreted with the help of available crystal structures, revealed the conformational changes required to switch the stability requirements from assembly to uncoating during the infectious cycle. In the immature virus, precursors of minor coat proteins pIIIa, pVIII and pVI increase the network of interactions that hold together the icosahedral protein shell, while core protein precursors pVII and



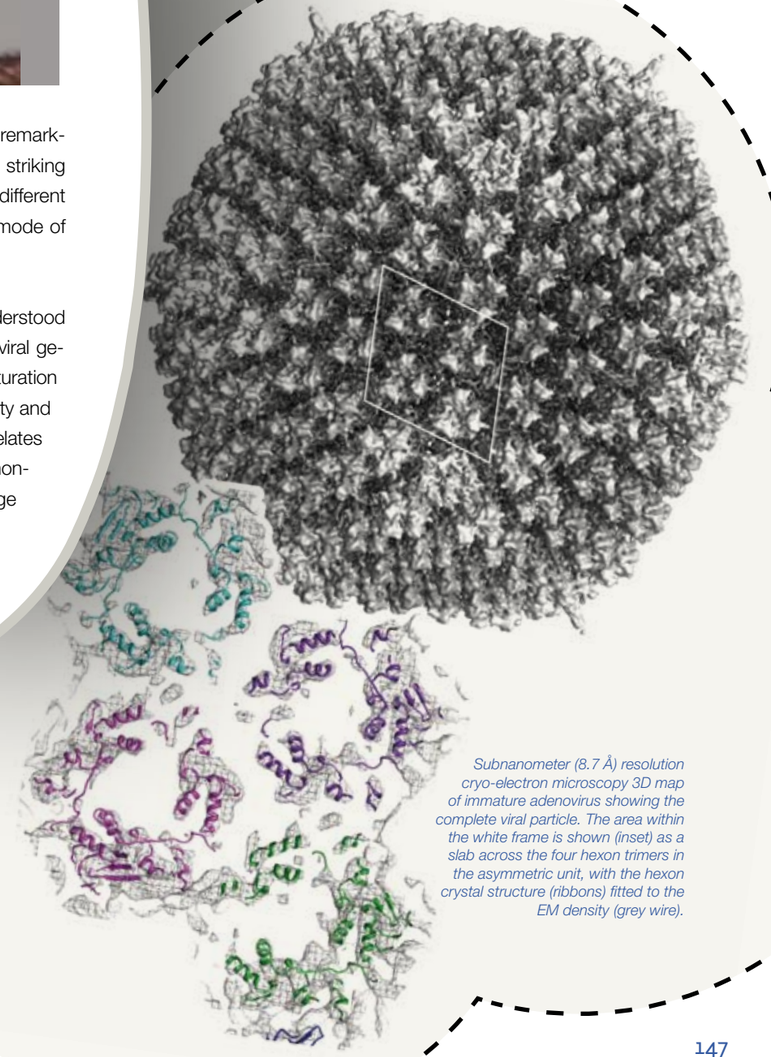
pre-μ act to tightly compact the viral genome into a remarkably stable, spherical structure. Interestingly, this striking change in core organization correlates with a different disassembly pattern, providing a glimpse into the mode of DNA packing within the virion.

Our current research lines focus on the less understood aspects of adenovirus assembly, such as how the viral genome is packaged into the capsid, how virion maturation occurs, the key elements that modulate virion stability and mechanical properties, how adenovirus evolution relates to that of its hosts and finally, the organization of non-icosahedral virion components. Accurate knowledge of adenovirus structure and biology is fundamental to both the discovery of anti-adenovirus drugs and the design of new, efficient adenoviral therapeutic tools.

SELECTED PUBLICATIONS

Pérez-Berná AJ, Marabini R, Scheres SHW, Menéndez-Conejero R, Dmitriev IP, Curiel DT, Mangel WF, Flint SJ, San Martín C (2009) Structure and uncoating of immature adenovirus. **J Mol Biol** 392:547-557.

Sorzano COS, Recarte E, Alcorlo M, Bilbao-Castro JR, San Martín C, Marabini R, Carazo JM (2009) Automatic particle selection from electron micrographs using machine learning techniques. **J Struct Biol** 167:252-260.



Subnanometer (8.7 Å) resolution cryo-electron microscopy 3D map of immature adenovirus showing the complete viral particle. The area within the white frame is shown (inset) as a slab across the four hexon trimers in the asymmetric unit, with the hexon crystal structure (ribbons) fitted to the EM density (grey wire).

Macromolecular Structures

LEAD INVESTIGATOR

Mark J. van Raaij

DOCTORAL STUDENT

Carmela García Doval

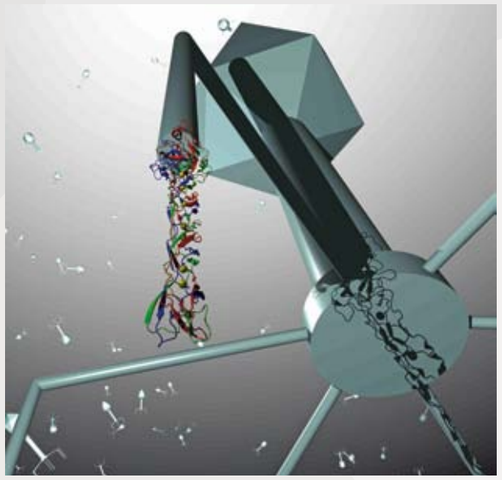
Schematic image of bacteriophage T4 with the structure of the receptor-binding tip of the long tail fibre highlighted. The phage is viewed from the bottom, i.e. from the side of the bacterium about to be infected (image courtesy of Florencio Pazos of the Sequence Analysis and Structure Prediction Unit, CNB).

Structural Biology of Viral Fibres

Adenovirus and bacteriophages such as T4, T5, T7 and lambda attach to their host cell via specialised fibre proteins.

These fibres all have the same basic architecture: they are trimeric and contain an N-terminal virus attachment domain, a long thin shaft domain and a more globular C-terminal cell attachment domain. These fibrous proteins are very stable to denaturation by temperature or detergents. In 2010, we determined the structures of the porcine adenovirus type 4 fibre head and galectin domains and of the receptor-binding tip of the bacteriophage T4 long tail fibre protein gp37.

The adenovirus NADC-1 isolate, a strain of porcine adenovirus type 4, has a fibre containing an N-terminal virus attachment region, shaft and head domains and a C-terminal galectin domain, connected to the head by an RGD-containing sequence. The crystal structure of the head domain is similar to previously solved adenovirus fibre head domains, but specific residues for binding coxsackievirus and adenovirus receptor (CAR), CD46 or sialic acid are not conserved. The structure of the galectin domain revealed an interaction interface between its two carbohydrate recognition domains, locating the two sugar binding sites face-to-face. Other tandem-repeat galectins may have the same arrangement. We showed that the galectin domain binds carbohydrates containing lactose and *N-acetyl*-lactosamine units. Modification of the galectin domain of this fibre should allow targeting to specific carbohydrate receptors.



Bacteriophages are the most numerous replicating entities in the biosphere, but little high-resolution structural detail is available on their receptor-binding fibres. We solved the structure of the receptor-binding tip of the bacteriophage T4 long tail fibre. It showed an unusual elongated six-stranded anti-parallel β-strand needle domain containing seven iron ions coordinated by histidine residues arranged co-linearly along the core of the biological unit. At the end of the tip, the three chains intertwine to form a broader head domain, which contains the putative receptor interaction site. The structure, a previously unreported beta-structured fibrous fold, provides insights into the remarkable stability of the fibre and suggests a framework through which mutations expand or modulate receptor-binding specificity. Modification of the bacteriophage fibre receptor binding specificities might lead to improved detection and elimination of specific bacteria.

We also collaborated with other research groups and have determined the structures of cyclic antibiotic peptides synthesised by the group of Dr Mark Overhand at Leiden University (The Netherlands) and of bacterial dehydroquinases complexed with inhibitors synthesised by the group of Dr Concepción González Bello of the University of Santiago de Compostela (Spain).

Structure of the galectin region of the porcine adenovirus type 4 fibre illustrating the cooperative nature of tri-N-acetyl-lactosamine (yellow) binding by the N- and C-terminal carbohydrate-recognition domains (shown in blue and red, respectively).

SELECTED PUBLICATIONS

Leiman PG, Arisaka F, van Raaij MJ, Kostyuchenko VA, Aksyuk AA, Kanamaru S, Rossmann MG (2010) Morphogenesis of the T4 tail and tail fibers. **Virol J** 7:355.

Bartual SG, Otero JM, Garcia-Doval C, Llamas-Saiz AL, Kahn R, Fox GC, van Raaij MJ. Structure of the bacteriophage T4 long tail fiber receptor-binding tip. **Proc Natl Acad Sci USA** 107:20287-92.

Kapoerchan VV, Knijnenburg AD, Niamat M, Spalburg E, de Neeling AJ, Nibbering PH, Mars-Groenendijk RH, Noort D, Otero JM, Llamas-Saiz AL, van Raaij MJ, van der Marel GA, Overkleeft HS, Overhand M (2010) An adamantyl amino acid containing gramicidin S analogue with broad spectrum antibacterial activity and reduced hemolytic activity. **Chemistry** 16:12174-81.

Peón A, Otero JM, Tizón L, Prazeres VF, Llamas-Saiz AL, Fox GC, van Raaij MJ, Lamb H, Hawkins AR, Gago F, Castedo L, González-Bello C (2010) Understanding the key factors that control the inhibition of type II dehydroquinase by (2R)-2-benzyl-3-dehydroquinic acids. **ChemMedChem** 5:1726-33.

Guardado-Calvo P, Muñoz EM, Llamas-Saiz AL, Fox GC, Kahn R, Curiel DT, Glasgow JN, van Raaij MJ (2010) Crystallographic structure of porcine adenovirus type 4 fiber head and galectin domains. **J Virol** 84:10558-68.

Macromolecular Structures

LEAD INVESTIGATORS
José Mª Valpuesta
& Jaime Martín-Benito

POSTDOCTORAL SCIENTISTS
Jorge Cuéllar
Elías Herrero
Begoña Sot

PREDOCTORAL SCIENTISTS
Sara Alvira
Srdja Drakulic
Silvia Hormeño
Álvaro Peña
María Ángeles Pérez
Marina Serna
Hugo Yébenes

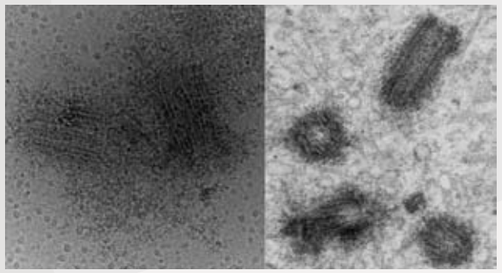
TECHNICIANS
Rocío Arranz
Ana Beloso

Structure and Function of Molecular Chaperones

Our group is interested in the structural and functional characterisation of macromolecular complexes, using electron microscopy and image processing techniques as our main tools.

In particular, we are very much interested in the study of molecular chaperones, a group of proteins involved in assisting not only in the folding of other proteins but also in their degradation. These two processes are carried in most cases by the coordinated functions of different chaperones that form transient complexes, thus forming an assembly line that make more efficient the protein folding and degradation processes. We are currently working with chaperones such as CCT, Hsp70, PFD, and PhLP.

We also study the structure of the centrosome and of some centrosomal complexes and proteins. The centrosome is



Left, Cryoelectron microscopy image of the centrosome. Right, Image of several centrioles.

the major microtubule organising centre (MTOC) in most animal cells. Typically, centrosomes are made of a pair of centrioles embedded in the amorphous, pericentriolar material (PCM). We analyse the overall structure of the centrosome using two approaches, the first one of which is electron tomography, a technique that has recently undergone major improvements, and which might allow the structural characterization of entire centrosomes in near-native conditions. The second approach is X-ray tomography, a technique that certainly allows the reconstruction of whole centrosomes, albeit at lower resolution than electron tomography; we plan to use the facilities being set up in the dedicated beam line at the Spanish ALBA Synchrotron. We are also working on the structural characterisation of centrosomal proteins using conventional electron microscopy.

Finally, we are also interested, in collaboration with the group of Dr. José L. Carrascosa, in the characterisation of the forces involved in the function of certain proteins, and in the manipulation of cell organelles using single-molecule techniques such as optical and magnetic tweezers.

SELECTED PUBLICATIONS

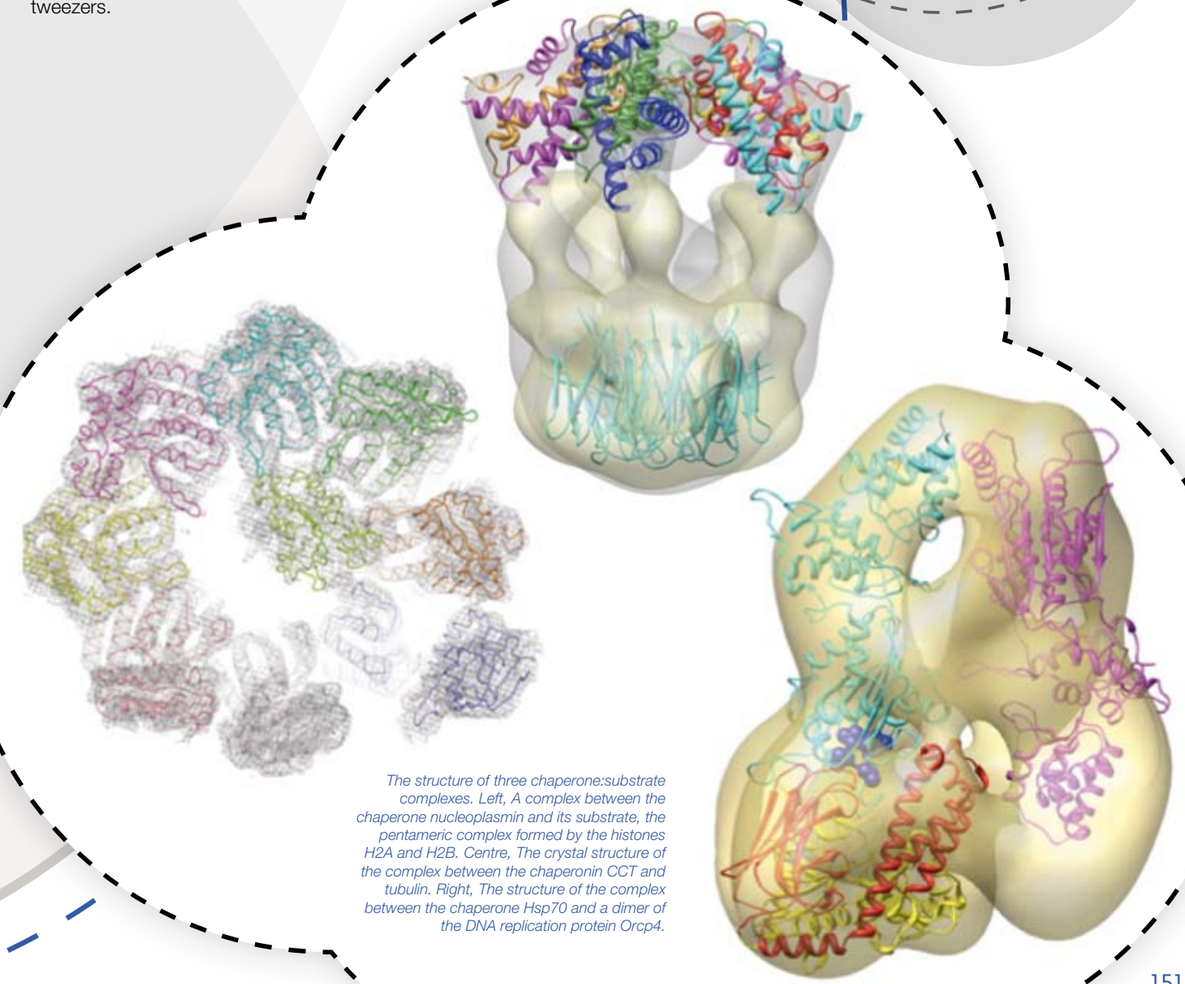
Alvarado-Kristensson M, Rodríguez MJ, Silió V, Valpuesta JM, Carrera AC (2009) SADB kinase-mediated γ-tubulin phosporylation regulates centrosome duplication. **Nat Cell Biol** 11:1081-92.

Hormeño S, Ibarra B, Chichón FJ, Habermann K, Lange BMH, Valpuesta JM, Carrascosa JL, Arias-Gonzalez JR (2009) Single centrosome manipulation reveals its electric charge and associated dynamic structure. **Biophys J** 97:1022-30.

Albert A, Yunta AC, Arranz R, Salido E, Valpuesta JM, Martín-Benito J (2010) Direct observation of a folding intermediate of a substrate protein bound to the chaperonin GroEL. **J Biol Chem** 285:6371-76.

Moreno-del Álamo M, Sánchez-Gorostiaga A, Serrano AM, Prieto A, Cuéllar J, Martín-Benito J, Valpuesta JM, Giraldo R (2010) The lSM Motif is the Main Target for Hsp70 Chaperones in the AAA+ Domain of Yeast Orc4p and Modulates Essential Interactions with Orc2p. **J Mol Biol** 403:24-39.

Ramos I, Martin-Benito J, Finn R, Bretaña L, Aloria K, Arizmendi JM, Ausió J, Muga A, Valpuesta JM, Prado A (2010) Nucleoplasmin binds histone H2A-H2B dimers through its distal face. **J Biol Chem** 285:33771-78.



The structure of three chaperone:substrate complexes. Left, A complex between the chaperone nucleoplasmin and its substrate, the pentameric complex formed by the histones H2A and H2B. Centre, The crystal structure of the complex between the chaperonin CCT and tubulin. Right, The structure of the complex between the chaperone Hsp70 and a dimer of the DNA replication protein Orcp4.