Description of the Programme’s areas of research and strategic goals

The Structural Biology Programme (SBP) has two strategic goals. On the one hand, we use structural and molecular biology to investigate the molecular mechanisms of proteins and macromolecular complexes that contribute to cancer progression. The Programme studies protein kinases as well as protein complexes involved in the complicated cellular response to DNA damage and genomic instability. A better understanding of how these macromolecules work and the determination of their atomic structures provides the knowledge needed to understand their roles in cancer and potentially guide new therapeutic opportunities. On the other hand, the Programme uses bioinformatics tools, computational cancer genomics, and computational oncology to better understand the complexity of cancer, predict therapy responses and develop new therapeutic strategies. SBP is currently composed of 1 Senior Group, 5 Junior Groups and 4 Units.

Summary of milestones & major achievements during 2020

2020 has been a difficult year for the Programme due to the Covid-19 pandemic, and this struggle was especially challenging for 2 new junior groups that had just started to assemble their team at CNIO. Despite the difficulties, everyone in the Programme has made a substantial effort to keep our research moving; we made some important contributions to the mechanistic understanding of the mismatch repair machinery in response to DNA replication errors, to the role of RUVBL1 and RUVBL2 ATPases in the regulation of essential cellular processes, and to the understanding of the activation of cytosolic hybrid histidine kinases and the Focal Adhesion Kinase (FAK). In addition, work by the groups in the Programme has contributed to the analysis of tissue-specific alternative splicing, to the characterisation of tumour-immune heterogeneity in advanced ovarian cancer, and to creating new tools for drug repositioning and prioritisation. 

“Our Programme uses structural biology and computational and genomic tools to improve our understanding of the complexity of cancer and of the proteins involved.”
Our current work dedicates special attention to study RUVBL1 and RUVBL2, two highly conserved AAA+ ATPases that are essential for several cellular processes relevant in cancer, including Fanconi anaemia, chromatin remodelling, nonsense-mediated mRNA decay (NMD), and the assembly and activation of large macromolecular complexes such as the those formed by mTOR and ATR kinases. Interestingly, RUVBL1-RUVBL2 inhibitors show anti-oncogenic potential, and cancer cells with high mTOR activity are dependent on the functions of RUVBL1-RUVBL2 for survival. How RUVBL1 and RUVBL2 perform their functions is only partially understood. Our work provides novel structural and mechanistic understanding of how these ATPases work, which will be useful for exploring new ways to target these proteins. For this, we combine biochemistry, molecular and cell biology with cryo-electron microscopy methods that allow us to visualise individual macromolecular complexes and to resolve their structure at high resolution.

“We have improved the structural understanding of how RUVBL1 and RUVBL2 are regulated, information needed to explore new ways to target these ATPases as a therapeutic opportunity against cancer.”
RuvBL-like 1 (RUVBL1) and RuvBL-like 2 (RUVBL2) are 2 highly conserved AAA+ ATPases, which have been found to be essential in a wide range of unrelated cellular processes, including transcriptional regulation, chromatin remodelling, DNA repair, Fanconi anaemia, nonsense-mediated mRNA decay (NMD), and the assembly and activation of complexes formed by the kinases of the phosphatidylinositol-3-kinase–related kinase (PIKK) family, such as mTOR, ATR, and SMG1. RUVBL1 and RUVBL2 are involved in cancer through their contribution to these cellular processes. They are essential for tumour cell growth, and they are overexpressed in many cancer types such as hepatocellular carcinoma, colon, breast or lung cancer. Interestingly, recent years have seen the development of several inhibitors of RUVBL1 and RUVBL2 ATPase activity for their use against cancer cells.

RUVBL1 and RUVBL2 work as a heterohexameric complex, but how this complex is capable of performing such a diversity of functions remains poorly understood. During 2020 we used cryo-electron microscopy (cryo-EM) studies combined with other techniques to characterise 3 cellular processes in which these ATPases are essential.

On the one hand, we characterised in detail how a domain of the RNA polymerase II-associated protein 3 (RPAP3) protein can specifically recognise the ATPase domain of RUVBL2. RPAP3 interacts not only with RUVBL1-RUVBL2, but also with HSP90 and other proteins that are needed to assemble and activate a growing list of complexes, including RNA polymerase II and complexes of the PIKK family of kinases such as mTOR, ATR, and SMG1. In addition, in collaboration with Jens Luders at the IRB (Barcelona), we found that RUVBL1-RUVBL2 participates in the assembly of the human γ-tubulin ring complex, a macromolecular assemblage that regulates microtubule formation. This finding was used to describe, for the first time, a procedure for the reconstitution of the human γ-tubulin ring complex that will be useful for functional and structural studies.

Furthermore, we started to characterise what is the function of RUVBL1 and RUVBL2 in nonsense-mediated mRNA decay (NMD), in collaboration with the Electron Microscopy Unit. NMD is a surveillance pathway that regulates gene expression by targeting some RNAs for degradation, especially aberrant mRNAs containing premature termination codons. RUVBL1–RUVBL2 ATPase activity is essential for the initiation of NMD but why and how this happens is still unknown. In 2020 we showed that DHX34, an RNA helicase that participates in the initiation of an NMD response, interacts with RUVBL1–RUVBL2, inducing conformational changes that regulate their ATPase activity. These findings suggest that factors required for NMD are coupled to RUVBL1 and RUVBL2 to regulate their ATPase activity. Further work will be needed to understand in full how RUVBL1 and RUVBL2 regulate NMD.

Further work will be needed to understand further how the NMD is regulated by RUVBL1 and RUVBL2 and how this complex is capable of performing such a diversity of functions. The study of related factors that are involved in the regulation of NMD will be an important area for future research.
ANNUAL REPORT 2020

Basic Research

Potent and specific protein kinase inhibitors, leading to be translated into the design and development of more insults. Crucially, such atomic and molecular information in cancer due to oncogenic mutations and other oncogenic complexes to transmit signals inside the cell. We place inputs, and how they assemble into macromolecular protein regulated by posttranslational modifications and allosteric kinase function: how protein kinases are activated and development of better therapeutics. Our research focuses of oncogene activation and signalling is key for the design and research. Understanding the structural and molecular bases is a crucial but still pending challenge in current cancer.

Rational and precise targeting of oncogene-driven signalling is a crucial but still pending challenge in current cancer research. Understanding the structural and molecular bases of oncogene activation and signalling is key for the design and development of better therapeutics. Our research focuses on the structural and molecular understanding of protein kinase function: how protein kinases are activated and regulated by posttranslational modifications and allosteric inputs, and how they assemble into macromolecular protein complexes to transmit signals inside the cell. We place special emphasis on how these mechanisms are corrupted in cancer due to oncogenic mutations and other oncogenic insults. Crucially, such atomic and molecular information can be translated into the design and development of more potent and specific protein kinase inhibitors, leading eventually to more effective drugs for the treatment of cancer patients.

We apply an integrated and multidisciplinary approach combining molecular biology for the generation of suitable constructs, protein biochemistry and biophysics for protein purification, quality assessment and functional evaluation, mass spectrometry (MS) for the quantification of posttranslational modifications, X-ray crystallography, and in vivo validation using Drosophila models. Furthermore, we use structure-guided drug discovery and MD simulation approaches to exploit structural and functional vulnerabilities for the design, development, and optimisation of optimal protein kinase inhibitors.

In 2020, we made significant progress on the research lines initially established in the laboratory:

1. We elucidated important structural and molecular details about the precise mechanism of the catalytic activation and auto-regulation of the c-Src oncogene. By applying a systematic phospho-proteomic approach, we identified new c-Src autophosphorylation sites and revealed that “canonical” activating and repressive tyrosine residues actually play other important roles and functions not previously envisioned.

2. Another research line was directed at dissecting the function of CDC2c-RET, a RET oncogenic fusion and driver in NSCLC. We successfully purified this challenging protein in different isoforms and length-variants and, by applying an integrated approach, demonstrated that the full-length construct behaves as an active dimer in solution. Auto-phosphorylation assays demonstrated fast kinetics compared to RET wild-type constructs. Further phospho-site mapping by MS and dissection of the activation mechanism highlighted important roles for catalytic activity and substrate specificity through unexpected elements.

3. A third research line focused on the exploitation of structural and functional vulnerabilities in RET for the rational design and development of highly specific inhibitors. Our current paradigm is based on recently developed second generation RET inhibitors LOXO-292 and BLU-667 that showed excellent results in both preclinical models and early clinical trials, resulting in their timely FDA approval for the treatment of RET-rearranged or -mutated cancers. We are applying an integrated approach combining structural data, molecular docking, structure-guided molecular dynamics simulations, screening with both virtual and chemical libraries, together with biophysical and biochemical tools for functional validation. Following this multidisciplinary approach, we identified an allosteric interface in RET with a good druggability score that can potentially be targeted with allosteric inhibitors. Furthermore, we uncovered a new sub-pocket within the ATP-binding site that is exploited by highly specific second-generation type I RET inhibitors (FIGURE). This information will be crucial for the design and development of highly specific, clinically successful third generation RET inhibitors able to overcome refractory RET mutations.
Safeguarding genetic information is essential to all forms of life. Two key cellular processes keep it free from errors: DNA replication and repair. Importantly, when they do not work correctly, genetic information may be damaged or lost, ultimately leading to disease. Deregulation and malfunction of the protein machinery that safeguards our genome are a hallmark of cancer, but it remains unclear how this happens at the molecular level. The devil is in the detail, and we aim to understand what and when something goes wrong with these molecular machines, so that we can act on it to correct it as well as to prevent it from happening.

These macromolecules are like real-life machines, with intricate mechanisms that allow them to perform their activities. To understand how they work, we use cryo-electron microscopy and biochemistry in an integrative approach. Beyond fundamental research, this structural information provides the necessary detail for drug development.

**Mismatch repair**

The DNA mismatch repair machinery (MMR) corrects the errors introduced by DNA polymerases during DNA replication and is critical for genome stability. The MutS protein loads onto newly synthesised DNA and searches for mismatches. The recognition of an error in the DNA leads to an ATP-dependent conformational change that transfers MutS into a sliding clamp state. Only this MutS state can activate the MutL ATPase, which, in turn, promotes the removal of the DNA for repair. These protein complexes are incredibly dynamic and flexible. Because of this, critical steps of this process have remained elusive to structural analysis. Using cryo-Electron Microscopy (cryo-EM), we captured multiple functional steps and studied the conformational changes that these proteins undergo to recognise the mismatch and license the downstream events that lead to repair. These studies were carried out in collaboration with Titia Siems (Netherlands Cancer Research Institute) and Meindert Lamers (Leiden University).

**DNA replication and repair - focus on mitochondria**

Eukaryotic cells have 2 genomes: nuclear and mitochondrial. However, how the mitochondrial genome’s integrity is maintained through the equilibrium between DNA replication, repair and degradation, and organelle dynamics, remains unclear. We are interested in understanding these pathways because of their implications for ageing and disease and, in particular, their relationship to cancer. By combining in vitro reconstitution of DNA replication complexes with cryo-EM imaging, we aim to capture the replication machinery at different functional stages, allowing us to understand in detail its mechanisms and how it is regulated.

**FIGURE**

Mismatch repair studies. (A) Scheme representing the initial stages of the DNA mismatch repair pathway: mutS binding and DNA scanning, mismatch binding, clamp formation, mutL recruitment, and sliding clamp formation. These steps control the licensing of DNA repair. (B) Cryo-EM micrograph of MutS protein on DNA (long strings) (left) and 2D class averages of the protein after image processing (right). These images are used for high-resolution structural analysis.
RESEARCH HIGHLIGHTS

Cancer fitness landscape: from within a gene (cis) to between genes (trans)

The classic 2-hit model postulates that both alleles of a tumour suppressor gene must be inactivated by a combination of 2 different alterations for tumour progression. However, some cancer genes increase tumour fitness after only a single hit and, in some cases, a second hit may actually be detrimental. To comprehensively understand the cancer fitness landscape, we analysed >10,000 tumours and classified cancer genes as 2 hits, 1 hit, or having optimal activity levels, which is a dangerous approximation because the activity-fitness functions of individual cancer genes are often diverse depending on the context. Specifically, mutations in other cancer genes frequently switch individual drivers from requiring 2 hits to 1 hit being sufficient to promote tumour progression. These results will provide the correct genetic model for a cancer gene, depending on their contexts, and emphasise a frequent redundancy between a second hit occurring in the same gene or in a second gene in a pathway during tumour progression. These studies were conducted in collaboration with Fran Supek (IRB, Barcelona) and Ben Lehner (CRG, Barcelona).

Inherited variants of Mendelian disease-associated genes in cancer genomics

Hereditary diseases are caused by pernicious mutations in certain genes or chromosomes. Usually, the abnormalities appear in newborns or during infancy, but sometimes they also occur in adults, such as is the case with Huntington's disease. In cases of late onset, it is reported that not only does it cause a single disease, but it also changes the concomitant pathways or affects cancer development if the stress from toxicity is sustained. The occurrence of cancer is apparent when there is an accumulation of additional variations. Using large-scale cancer genomics data, we identified the contribution of Mendelian disease-associated genes to cancer risk across more than 30 cancer types. These results will enable cancer prevention through genetic testing aimed at reflecting individual disease susceptibility to various diseases. These studies were carried out in collaboration with Young-il Goh (Seoul National University, South Korea).

“By analysing large-scale cancer genomics data, we aim to further pursue novel questions about cancer-type and context-specific tumour progression to understand tumour heterogeneity.”

FIGURE A systematic understanding of context-dependent cancer fitness landscape by analyzing large-scale cancer genomics. (A) A cancer gene will change its optimal model of tumour progression depending on the context, from cis-regulation to trans-regulation. (B) Novel insight into the increment of cancer risk by Mendelian disease-associated genes.
ANNUAL REPORT 2020

Our main research areas include:

- Predicting therapy response using genomic signatures of chromosomal instability (CIN). By therapeutically targeting CIN, we aim to improve outcomes in these tumours.

- Developing single cell/nucleus sequencing approaches to detect ongoing CIN.

We are applying these technologies at the earliest stages of tumour development in patients with premalignant lesions to help prevent aggressive, difficult to treat cancers.

OVERTVIEW

In the Computational Oncology Group, we are tackling some of the deadliest cancers by targeting the causes of chromosomal instability. Pancreatic, oesophageal, lung and ovarian cancers have some of the lowest survival rates, but they also share a common trait that we can exploit – extreme chromosomal instability. By therapeutically targeting CIN, we aim to improve outcomes in these tumours.

Our main research areas include:

- Using model systems to develop therapeutic strategies to target CIN.
- Predicting therapy response using genomic signatures of CIN in patient biopsies.

2020 started well for the Computational Oncology Group with the publication of work the group was involved in as part of the pan-cancer analysis of whole-genomes project. However, the onset of the Covid-19 pandemic meant the first year of operations for the Computational Oncology Group did not go as planned. It did, however, make some of the small wins a lot more important!

We now have a brand-new laboratory setup designed specifically to interrogate tumour DNA copy number using a low-cost, low-pass whole-genome sequencing strategy. We have established our computational lab and have worked closely with the Bioinformatics Unit to upgrade CNIO’s scientific computing infrastructure to handle the imminent influx of sequencing data.

The Group is also slowly growing in size – María José García joined from Javier Benítez’s Group upon his retirement and has been the driving force behind getting the Computational Oncology Group’s laboratory operations up and running. She brings with her a wealth of experience preparing clinical samples for DNA sequencing and her own funded project on DNA mismatch repair in ovarian cancer. Blas Chaves has joined for his Master’s project at Universidad Complutense de Madrid (UCM) and has demonstrated his aptitude in the lab and in front of the computer. He will be characterising different types of chromosomal instability in model systems using copy number signatures.

We have developed collaborations with Marcos Malumbres’ laboratory where we will determine the types of CIN caused by knockout of CDKs, and with Felipe Cortés’ laboratory looking at CIN induced by etoposide treatment. We have started our project with Sam Janes at University College London (UCL) detecting CIN in premalignant lung cancers. Finally, we have made significant progress characterising pan-cancer patterns of CIN and how they relate to drug response – look out for a publication in 2021!

It looks an exciting year moving ahead with 3 new members and a chance to finally put our new laboratory infrastructure into action to tackle some of the deadliest cancers.

In our first year of operation we have established computational and laboratory infrastructure that will allow us to observe chromosomal instability at the earliest stages of tumour evolution.

“...and a chance to finally put our new laboratory infrastructure into action to tackle some of the deadliest cancers.”

Our main research areas include:

- Developing single cell/nucleus sequencing approaches to detect ongoing CIN.
- Using model systems to develop therapeutic strategies to target CIN.
- Predicting therapy response using genomic signatures of CIN in patient biopsies.

2020 started well for the Computational Oncology Group with the publication of work the group was involved in as part of the pan-cancer analysis of whole-genomes project. However, the onset of the Covid-19 pandemic meant the first year of operations for the Computational Oncology Group did not go as planned. It did, however, make some of the small wins a lot more important!

We now have a brand-new laboratory setup designed specifically to interrogate tumour DNA copy number using a low-cost, low-pass whole-genome sequencing strategy. We have established our computational lab and have worked closely with the Bioinformatics Unit to upgrade CNIO’s scientific computing infrastructure to handle the imminent influx of sequencing data.

The Group is also slowly growing in size – María José García joined from Javier Benítez’s Group upon his retirement and has been the driving force behind getting the Computational Oncology Group’s laboratory operations up and running. She brings with her a wealth of experience preparing clinical samples for DNA sequencing and her own funded project on DNA mismatch repair in ovarian cancer. Blas Chaves has joined for his Master’s project at Universidad Complutense de Madrid (UCM) and has demonstrated his aptitude in the lab and in front of the computer. He will be characterising different types of chromosomal instability in model systems using copy number signatures.

We have developed collaborations with Marcos Malumbres’ laboratory where we will determine the types of CIN caused by knockout of CDKs, and with Felipe Cortés’ laboratory looking at CIN induced by etoposide treatment. We have started our project with Sam Janes at University College London (UCL) detecting CIN in premalignant lung cancers. Finally, we have made significant progress characterising pan-cancer patterns of CIN and how they relate to drug response – look out for a publication in 2021!

It looks an exciting year moving ahead with 3 new members and a chance to finally put our new laboratory infrastructure into action to tackle some of the deadliest cancers.

In our first year of operation we have established computational and laboratory infrastructure that will allow us to observe chromosomal instability at the earliest stages of tumour evolution.

“...and a chance to finally put our new laboratory infrastructure into action to tackle some of the deadliest cancers.”

Our main research areas include:

- Developing single cell/nucleus sequencing approaches to detect ongoing CIN.
- Using model systems to develop therapeutic strategies to target CIN.
- Predicting therapy response using genomic signatures of CIN in patient biopsies.
OVERVIEW

This Unit focuses on the technical and scientific management of Nuclear Magnetic Resonance (NMR) Spectroscopy and molecular biophysics instrumentation available at the Structural Biology Programme. It provides CNIO researchers with equipment and experimental support for a variety of techniques used in biophysical studies of molecules involved in cancer. This includes the in vitro characterisation of the structure and dynamics of proteins by NMR, and characterisation of the affinity and kinetics of the interactions of proteins with other biopolymers and small molecules that could represent initial hits in the drug discovery process, or could serve as research compounds for biophysical and functional characterisation or profiling of proteins with other biopolymers and small molecules that could represent initial hits in the drug discovery process, or could serve as research compounds for biophysical and functional characterisation of the affinity and kinetics of the interactions, including spectrophotometers, a fluorimeter, isothermal titration and differential scanning calorimeters, a circular dichrograph, dynamic and multi-angle static light scattering devices, and two biosensor instruments: surface plasmon resonance (SPR), and biolayer interferometry (BLI). Research Groups mostly from, but not limited to, the Structural Biology Programme used these technologies throughout 2020 (i.e., the Haematological Malignancies Clinical Research Unit, the Monoclonal Antibodies Unit, and the Experimental Therapeutics Programme – ETP).

The Unit hosts a 700 MHz NMR spectrometer that is equipped with probes and a sample changer to run up to 120 samples automatically. This provides medium throughput for the screening of small molecule protein binders (together with ETP), as well as for metabolite quantification that, in 2020, was done in collaboration with the CNIO-Lilly Cell Signalling Therapies Section (ETP), and the Growth Factors, Nutrients and Cancer and Metabolism and Cell Signalling Groups (Molecular Oncology Programme). During 2020, we incorporated a QTOF mass spectrometer that will complement our battery of techniques for the quality control of purified proteins with the information contained in their intact mass spectra. For example, we examined several reference proteins (see FIGURE), verifying that the instrument can determine the mass with high precision and accuracy employing nanogram amounts. In addition, HPLC-MS measurements of biofluids were also initiated. Collectively, with our client groups, we will continue implementing sample preparation protocols and developing spectroscopic and analytical tools to characterise metabolites present in different biological samples.

“In 2020 we initiated work with a QTOF mass spectrometer that will allow the quality control of purified proteins from their intact mass spectrum, as well as the targeted characterisation or profiling of metabolites in liquid samples of cancer model systems.”

**PUBLICATIONS**

- Izquierdo-Garcia JL, Comella-Del-Barrio J, Dominguez J (2020). Discovery of Nutrients and Cancer and Metabolism and Cell Signalling Therapies Section (ETP), and the Growth Factors, Nutrients and Cancer and Metabolism and Cell Signalling Groups (Molecular Oncology Programme). During 2020, we incorporated a QTOF mass spectrometer that will complement our battery of techniques for the quality control of purified proteins with the information contained in their intact mass spectra. For example, we examined several reference proteins (see FIGURE), verifying that the instrument can determine the mass with high precision and accuracy employing nanogram amounts. In addition, HPLC-MS measurements of biofluids were also initiated. Collectively, with our client groups, we will continue implementing sample preparation protocols and developing spectroscopic and analytical tools to characterise metabolites present in different biological samples.

**PUBLICATIONS**

- Izquierdo-Garcia JL, Comella-Del-Barrio J, Dominguez J (2020). Discovery of human samples. In addition, in 2020, we adopted a mass spectrometer for the characterisation of intact proteins and for metabolite studies using HPLC-MS methods.

**FIGURE** (A) Mass spectrum of perfused lysosome (120 µg/mL), at 3 µl/min with indication of the identified cationic forms (+13 to +8), allowing determination of a mass of 14304 Da (expected 14313 Da). (B) HPLC-MS chromatograms of human urine injected in a reverse-phase column as followed by base peak ion counts (lower traces, left scale) and UV absorbance (upper traces, right scale). Two consecutive runs of the same sample measured in positive (red) and negative (black) Electro Spray Ionization mode are shown.

**FIGURE** (A) Mass spectrum of perfused lysosome (120 µg/mL), at 3 µl/min with indication of the identified cationic forms (+13 to +8), allowing determination of a mass of 14304 Da (expected 14313 Da). (B) HPLC-MS chromatograms of human urine injected in a reverse-phase column as followed by base peak ion counts (lower traces, left scale) and UV absorbance (upper traces, right scale). Two consecutive runs of the same sample measured in positive (red) and negative (black) Electro Spray Ionization mode are shown.
Bioinformatics is a key discipline for understanding the cancer genome and for the future of cancer therapeutics. Bioinformatics-based approaches have the ability to transform the vast amount of biological data into comprehensive models that provide a deep understanding of cancer disease and the complex genotype-phenotype relationships needed to identify molecular cancer-driving alterations and novel therapeutic targets.

The CNIO Bioinformatics Unit (BU) has several objectives: (i) to develop new computational methodologies and bioinformatics tools to enable the integration of biological and clinical data, (ii) to achieve genome analysis in cancer patients’ data to identify new biomarkers and drug response mechanisms, (iii) to provide bioinformatics support with data analysis and interpretation using computational and statistical methods, and (iv) to maintain the scientific computing facilities at the CNIO and provide training in bioinformatics tools and methods.

"DREIMT provides the largest drug-immune expression signature associations database available, allowing the users to generate hypotheses and explore druggable targets throughout the immune system."

In 2020, the CNIO Bioinformatics Unit published more than 20 peer-reviewed articles (see full list on our web site https://bioinformatics.cnio.es/) as a result of our ongoing research projects and scientific collaborations. We studied the effect of drugs on the biology and activity of the immune system and its interaction with cancer cells. To this end we developed DREIMT (Troshlé K et al., 2020), a new hypothesis-generation web tool that performs drug prioritisation analysis for immunomodulation. DREIMT provides immunomodulatory drugs targeting up to 70 immune cell subtypes through a curated database that integrates 4,960 drug profiles and 2.6K immune gene expression signatures. DREIMT is the largest database for drug-immune expression signature associations currently available. DREIMT also provides tools to suggest potential immunomodulatory drugs targeting user-supplied gene expression signatures. DREIMT is fully accessible to the scientific community at http://www.dreimt.org. Additionally, we applied our method PanDrugs (https://www.pandrugs.org/) in the context of Pan-Cancer analysis of whole genomes (The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020).

During 2020, the Unit also analysed alternative splicing variants in vertebrates and found that they played a significant role in the evolution of brain and heart tissues (Rodriguez JM et al., 2020). Furthermore, our results in the analysis of SINE Alu genomic elements confirmed that such elements contribute to the expansion of the human proteome despite little evidence of their biological relevance (Martinez-Gomez L et al., 2020).

With regard to academic and knowledge transfer activities, we co-organised the Master’s degree in Bioinformática Aplicada a Medicina Personalizada y Salud (ISCIII-ENS) (visit our web page for a full list of activities).

**OVERVIEW**

**BIOINFORMATICS UNIT**

Fátima Al-Shahrour  
Unit Head  
Staff Scientist  
Michael Tress  
Graduate Students  
Santiago García (since March) (IFM), (CARM), Maria José Jiménez, Laura Martinez, Fernando Pinto, Kevin Troulé

*Plan de Empleo Joven de la Comunidad de Madrid (UPJEM) Employment Plan, Community of Madrid*

**RESEARCH HIGHLIGHTS**

DREIMT provides the largest drug-immune expression signature associations database available, allowing the users to generate hypotheses and explore druggable targets throughout the immune system.

**SELECTED PUBLICATIONS**

- Please see BU’s web site for a list of all publications.
Technical advances in the last decade have positioned cryogenic electron microscopy (cryo-EM) as one of the most powerful and effective technologies available to investigate the structures of macromolecules at near-atomic resolution. Among several cryo-EM structural determination methods, single-particle analysis is the most popular for structural biologists, as it has relatively well-established methods for sample preparation, data collection, image processing, and structural determination. At the CNIO we have in place a 120 kV, Tecnai G2 Spirit microscope equipped with a TVIPS CMOS detector that is used to obtain images of negatively stained samples and to screen vitrified samples. For medium resolution structural studies, the Unit is equipped with a JEM-2200FS cryo-EM and a K3 direct electron detector camera. Our scientific activity throughout 2020 involved collaborations with all the Research Groups from the Structural Biology Programme, several Groups from other Programmes, as well as with scientists outside the CNIO. For instance, in collaboration with the Cell Division and Cancer Group, we monitored centriole structure and organisation as a consequence of lack of Cep135, a protein involved in centrosomal and spindle dynamics; in collaboration with the Microenvironment and Metastasis Group, we studied the morphological changes of extracellular vesicles, mainly exosomes isolated from prostate cancer cells; in collaboration with the Macromolecular Complexes in DNA Damage Response Group, we continued our collaboration on the structural characterisation of several protein complexes, e.g., RUVBL1/2 complexes and DNA repair complexes; and, finally, in collaboration with Genome Integrity and Structural Biology Group, we have been setting-up a pipeline to use a cryo-EM as a tool for drug discovery.
RESEARCH HIGHLIGHTS

Our Unit works closely with the Experimental Therapeutics Programme on several projects. To fulfil the need of recombinant proteins, we produced, throughout the year, full-length and kinase domain human MASTL, full-length mouse TRF1 and human TRF1 dimerization domain, for biochemical, in vitro, thermo-stability and structural analyses. Furthermore, to support drug discovery projects, we performed several thermal shift assays (thermofluor) in the presence of compounds developed and tested at the Medicinal Chemistry Section and the Biology Section, respectively.

We also continued our close collaboration with the CNIO Monoclonal Antibodies Unit on the production of proteins to generate highly specific antibodies against several cancer-associated proteins such as HASPN, HANK, CB85C, CB85G and CB85J, and other protein tools such as Cas9. Additionally, we ran a number of internal collaborations with other CNIO Groups and Units, providing them with recombinant proteins for biochemical and/or cell-based functional assays; this was the case, for example, with the Telomerases and Telomerase Group, the Experimental Oncology Group, the Genomic Instability Group, the Cell Division and Cancer Group, the Melanoma Group, the H2O-CNIO Lung Cancer Clinical Research Unit, the Macromolecular Complexes in DNA Damage Response Group, the H2O-CNIO Haematological Malignancies Clinical Research Unit, and the Transformation and Metastasis Group.

The Unit maintained collaborations with various external groups: the Experimental Biology Department, CIB-CSIC, Spain; the Pharmacology and Therapeutics Department, Roswell Park Cancer Institute, USA; the Department of Biomedicine, University of Bergen, Norway; the Department of Crystallography and Structural Biology, Instituto Quimico-Fisico Rocasolano, CSIC, Spain; the Department of Immunology, Genetics and Pathology, Uppsala University, Sweden; the Cancer Immunotherapy Unit, 12 de Octubre University Hospital, Spain; and the Division of Pulmonary and Critical Care Medicine, Fibrosis Research Center, and Center for Immunology and Inflammatory Diseases, Harvard Medical School, USA.

Throughout 2020, the Unit also proceeded with its own scientific projects. We continued working on targeting the function of the Mdm2-MdmX E3 complex activity in the context of an NIH-funded collaborative project with the Department of Pharmacology and Therapeutics at Roswell Park Cancer Institute. In addition, we are recombinantly producing a T cell-recruiting bispecific antibody (named ATTACK) for structural and functional purposes, in collaboration with the company LeadArtis, the Department of Microbiology (Immunology) of the Complutense University of Madrid, and the Cancer Immunotherapy Unit of the 12 de Octubre University Hospital; a project funded by the Nuevas Comunicaciones programme of the Spanish Ministry of Science, Innovation and Universities. The Unit is also taking part in 2 collaborative projects with the Biomedical Application of Radioisotopes Unit of CERMAT, the Bioactive Nanomaterials Structure Group of the Complutense University of Madrid, and the CNIO’s Molecular Imaging Unit to develop new antibody-based positron emission tomography (immunoPET) imaging tools for tumour visualisation and pretargeted clickable antibody fragments for therapeutic applications; both projects are supported by BVIA Foundation grants. Finally, in 2020 we were awarded a BVIA Foundation grant, jointly with the Cancer Immunotherapy Unit of the 12 de Octubre Hospital’s Research Institute (+12), to design a new immunotherapy method to fight Covid-19.

Overview

The Crystallography and Protein Engineering Unit (XTPEU Unit) is a core facility created to provide on-demand services at different levels to fulfil the needs of our users. By offering services ranging from protein cloning to solving the 3D structures of proteins, we help our users to further comprehend how target proteins work. With this purpose in mind, we produce proteins for different types of biochemical/biophysical/in vitro/in vivo assays and for monoclonal antibody production, also offering macromolecular structural determination at high-resolution (atomic) by X-ray crystallography, and at low-resolution in solution by small-angle X-ray scattering (SAXS).

Protein co-crystallisation, in the presence of inhibitors or small molecules, is routinely done in our laboratory in combination with studies on protein thermal stability (thermofluor assay), to accelerate the guided drug discovery process.

“By fragment screening on crystals, we visualise direct interactions between small molecules and proteins, speeding up the identification of new targetable sites in drug discovery projects.”

Publications

