The Programme’s research areas and strategic goals

The research performed within the Structural Biology Programme (SBP) focuses on two major strategic areas. 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. For this, we apply a combination of approaches and methods, but single-particle cryo-electron microscopy (cryoEM) remains one of our core structural technologies. The Programme studies protein kinases, protein complexes involved in the cellular response to DNA damage and genomic instability, proteins participating in DNA replication as well as molecular chaperones. 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. To achieve these goals, the Groups are supported by specialised Units with expertise in several technologies and methods needed for protein production and characterisation, including cryoEM, NMR, protein crystallography, and biophysical methods such as bio-layer interferometry, SEC-MALS, Surface Plasmon Resonance and others. These platforms are used not only by the Groups within the Programme but also by several others at CNIO. 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. As part of these efforts, the Bioinformatics Unit has significant synergies with several groups at CNIO and develops computational tools and methods of interest in cancer research. SBP currently consists of 1 Senior Group, 4 Junior Groups and 5 Units.

Summary of milestones & major achievements in 2021

During 2021, the Structural Biology Programme made substantial contributions in several areas of cancer research. Cryo-EM was used to advance our understanding of the molecular and structural mechanisms of several cellular pathways relevant in cancer. For example, the Programme determined the atomic structures of molecular chaperones involved in mTOR activation and spliceosome assembly, studied mechanisms that transport amino acids across the cell membrane that can be targeted against cancer, and advanced in the molecular understanding of some of the mechanisms required to repair lesions in the DNA that accumulate and give rise to cancer. On the computational and bioinformatics front, the Programme contributed to the study of genetic interactions in cancer genes as well as the analysis of markers in cancer progression, among several other efforts. We also developed several computational pipelines and tools to help cancer studies such as single-cell RNA-seq analyses.

Several of SBP’s Units and Groups made significant contributions to the work of other CNIO groups as part of synergistic collaborations, helping to understand the toxicity of peptides linked to amyotrophic lateral sclerosis, contributing to the study of how melanoma-derived small extracellular vesicles induce metastasis and how centrosome duplication defects cause microcephaly, and helping to characterise trimerbodies as potential therapeutic proteins in cancer. Finally, it is imperative to mention that our expertise in protein structure and purification was put at the service of research against the pandemic. The Structural Biology Programme helped in studies about the prognosis of COVID-19 patients admitted to intensive care units and in the study of mucosal immunotherapy as potential protection against SARS-CoV-2 infection.

“Our Programme uses structural biology, computational and genomic tools to improve our understanding of the complexity of cancer and of the proteins involved.”
Our Group uses cryo-electron microscopy (cryoEM) to determine the 3D structure of large macromolecular complexes of relevance in cancer at high resolution. Structural information, in combination with molecular and cell biology and biochemistry, is then used to propose how these molecules work and increase our understanding of the molecular basis of cancer. Most of our efforts are currently focused on 2 major areas of research: the study of chaperones essential for the activation of several macromolecular complexes relevant in cancer such as those formed by the mTOR kinase, and the study of complexes implicated in the repair of DNA double-strand breaks. In addition, and in collaboration with other groups, we are studying the structure and mechanisms of several amino acid transporters. "We have characterised the structure and the molecular mechanisms of 2 protein complexes (TELO2-TTI1-TTI2 and LAT2/CD98hc) and 1 long non-coding RNA, considered important players in cancer."
RESEARCH HIGHLIGHTS

Structure of the TELO2-TTI1-TTI2 complex and its role in mTOR activation

As part of a collaboration with Laurence H Pearl and Chrisostomos Prodromou at the Genome Damage and Stability Centre, University of Sussex, UK, we helped to determine the structure of the TELO2-TTI1-TTI2 complex using cryoEM. This complex is essential for the maturation and activation of mTOR, a serine/threonine protein kinase that regulates several essential processes such as cell growth, cell proliferation, cell motility, autophagy, and protein synthesis. The mTOR signalling pathway is often activated in tumours, and the pathway is being studied intensively in the search for anti-cancer therapies. The structure of the TELO2-TTI1-TTI2 complex that we helped to resolve, together with biochemical experiments, revealed some of the mechanisms involved in the activation of mTOR by chaperones.

Long non-coding RNAs in DNA double-strand breaks in hepatocellular carcinoma

Long non-coding RNAs (lncRNA) are now considered essential players in cancer but the mechanisms are poorly understood. As part of a consortium involving several institutions in Europe and the USA, and directed by Puri Fortes at the Centre for Applied Medical Research (CIMA), University of Navarra, in Pamplona (Spain), we contributed to studying the mechanisms of NIHCOLE, a novel lncRNA induced in hepatocellular carcinoma (HCC), whose expression is associated with poor prognosis and survival. In a close partnership between our group at CNIO and the group of Fernando Moreno-Herrero at the CNB-CSIC in Madrid, and with funding from the local Government of Madrid, we used single-molecule imaging methods (AFM and electron microscopy) to characterise the structure of this lncRNA. These images show that NIHCOLE functions as a scaffold promoting the assembly of large multimeric complexes of proteins involved in the repair of DNA double-strand breaks.

Structure of heteromeric transporters of neutral amino acids

Amino acids play a central role in cellular metabolism. The transfer of amino acids across the plasma membrane is performed by proteins that bind and transport these molecules from the extracellular medium into the cell, and vice versa. Heteromeric Amino Acid Transporters (HATs) are a family of amino acid transporters that harmonise amino acid concentrations at each side of the plasma membrane, and they play a significant role in cancer and several inherited diseases. Several loss-of-function mutations in human LAT2/CD98hc are associated with age-related hearing loss and cataracts, and its overexpression in pancreatic cancer cells sustains glutamine-dependent mTOR activation to promote glycolysis and chemoresistance. LAT1/CD98hc is also linked to cancer and autism.

Each member of the HAT family displays a preference for transporting a certain set of amino acids. This specialisation explains the function of each HAT family member in certain physiological processes and diseases. The molecular mechanisms explaining why each family member shows exquisite preference for transporting some amino acids but not others had been mostly unknown. We determined the structure of one such member of the HAT family, LAT2/CD98hc, using cryoEM. This structural information, together with molecular dynamics and mutational and functional studies, enabled us to specify a few residues present in the substrate-binding pocket that contribute to determining substrate preference.

*PUBLICATIONS*


*AWARDS AND RECOGNITION*

Advisory Board for R&D and Innovation of the Autonomous Government of Navarre, Spain.
Scientific Advisory Board, Biomedical Research Centre of the Government of Navarre (NAVARRABIO), Spain.
External Scientific Advisory Board, Molecular Biology Institute of Barcelona (IBMB), Spain.
Rational and precise targeting of oncogene-driven signalling is a crucial and yet today outstanding 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 constructs 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 identification and quantification of post-translational modifications; X-ray crystallography for the 3D-visualization of proteins; and Drosophila as an in vivo model for data validation. Furthermore, we use structure-guided drug discovery and MD simulation approaches to exploit structural and functional vulnerabilities for the design, development, and optimization of protein kinase inhibitors as therapeutic agents in cancer.

In 2021, we made significant progress in all our laboratory research projects, which are materializing successfully and, as a result, several papers were submitted for publication and are under review.

1. c-Src codifies a non-receptor tyrosine kinase that is activated by a plethora of signalling receptors that are fundamental in the aetiology of cancer. Despite being the object of intense study over the last 40 years, the precise mechanism by which auto-phosphorylation regulates c-Src intrinsic activity and conformational state independent of external inputs, and how this process is corrupted in cancer, remains elusive. In our work we show that c-terminal Tyr 530 is a de facto c-Src auto-phosphorylation site with slow time-resolution kinetics and a strong intermolecular component. By contrast, activation-loop Tyr 419 undergoes very fast kinetics and a cis-to-trans phosphorylation switch that controls c-terminal phosphorylation, substrate specificity and substrate-like properties. In line with these findings, a Drosophila mutant at the equivalent residue in the activation loop shows tissue-specific functionality and milder but transforming phenotypes compared with wild-type or constitutive active variants. Furthermore, we provide evidence that the intrinsically disordered N-terminal region of c-Src does not promote direct dimerization in the "apo" or the ATP-complexed states, and that c-Src Tyr 530 auto-phosphorylation is associated with a lowered catalytic status. A crystal structure of the c-Src-Ponatinib complex in a DFG-out state reveals unusual active-like features and provides a clear snapshot of c-terminal Tyr 530 intramolecular phosphorylation between enzyme and substrate acting kinases. Altogether these data indicate that c-Src must adopt an alternative conformation to the inactive-closed state independent of c-terminal Src kinase phosphorylation, and that a sequential and coordinated cis-to-trans phosphorylation switch between the activation and c-terminal segments simultaneously controls c-Src catalytic and non-catalytic functions (Cuesta and Contreras et al., submitted for publication).

2. Gene fusion products are known drivers in human cancers, and are currently drug targets for personalised therapy. A second research line in the lab was established and directed at dissecting the functional and structural determinants for 2 RET oncogenic fusion products, CDC42-RET and KIF4B-RET. By applying an integrated approach, we demonstrated that full-length constructs behave like active dimers in solution. Auto-phosphorylation and enzymatic assays demonstrated fast kinetics compared to wild-type RET, and further phospho-proteomic characterisation by MS highlighted important roles for catalytic activity and substrate specificity through unexpected allosteric inputs by distant elements to the catalytic site (Hurtado et al., in preparation).

3. A third research line focuses 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 the 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 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, and screening with both virtual and chemical libraries together with biophysical and biochemical tools for functional validation. Following this approach, we identified an allosteric interface in RET with good druggability score that can be potentially targeted with allosteric inhibitors. Furthermore, we uncovered a cryptic-pocket within the ATP-binding site that is exploited by highly specific second-generation type I RET inhibitors. This information will be crucial to designing and developing highly specific third generation RET inhibitors that are clinically successful and able to overcome refractory RET mutations (Shehata et al., in press).

4. We initiated a new research line focused on the structural and functional characterization of human FAK. We want to explore how auto-phosphorylation drives the functional and conformational landscape of FAK, in a full-length setting, and how phosphorylation interferes with the assembly and interaction with substrates and signalling partners such as RET (both wild-type and oncogenic variants) and c-Src. Using a phospho-proteomic approach we already identified unexpected phospho-sites and revealed a previously unknown switch for FAK catalytic activation by N-terminal elements, which could be therapeutically exploited to design and develop next generation FAK inhibitors.

**RESEARCH HIGHLIGHTS**

<table>
<thead>
<tr>
<th>KINASES, PROTEIN PHOSPHORYLATION AND CANCER JUNIOR GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ivan Plaza Manacho, Junior Group Leader</td>
</tr>
<tr>
<td>Post-Doctoral Fellows</td>
</tr>
<tr>
<td>Julia M. Contreras, Ruben Julio</td>
</tr>
<tr>
<td>Martinez-Tomé (until December)</td>
</tr>
<tr>
<td>Graduate Students</td>
</tr>
<tr>
<td>Nicole Cuenda, Ana Martín, Housaifa Ahmed Shehata</td>
</tr>
<tr>
<td>Students in Practice</td>
</tr>
<tr>
<td>Avana Frous (January-June)</td>
</tr>
<tr>
<td>(Universiti de Lorraine, Nancy, France)</td>
</tr>
<tr>
<td>Marina Rodriguez (since October)</td>
</tr>
<tr>
<td>(Université Autonome de Madrid, Madrid, Spain)</td>
</tr>
<tr>
<td>Visiting Scientist</td>
</tr>
<tr>
<td>Yanara Astudillo (Universidad Tecnológica Ecuatoriana - Fundación Carolina, Quito, Ecuador)</td>
</tr>
</tbody>
</table>

| OVERVIEW |

| RESEARCH HIGHLIGHTS |

| In 2021, we made significant progress in all our laboratory research projects, which are materializing successfully and, as a result, several papers were submitted for publication and are under review. |
| 1. c-Src codifies a non-receptor tyrosine kinase that is activated by a plethora of signalling receptors that are fundamental in the aetiology of cancer. Despite being the object of intense study over the last 40 years, the precise mechanism by which auto-phosphorylation regulates c-Src intrinsic activity and conformational state independent of external inputs, and how this process is corrupted in cancer, remains elusive. In our work we show that c-terminal Tyr 530 is a de facto c-Src auto-phosphorylation site with slow time-resolution kinetics and a strong intermolecular component. By contrast, activation-loop Tyr 419 undergoes very fast kinetics and a cis-to-trans phosphorylation switch that controls c-terminal phosphorylation, substrate specificity and substrate-like properties. In line with these findings, a Drosophila mutant at the equivalent residue in the activation loop shows tissue-specific functionality and milder but transforming phenotypes compared with wild-type or constitutive active variants. Furthermore, we provide evidence that the intrinsically disordered N-terminal region of c-Src does not promote direct dimerization in the “apo” or the ATP-complexed states, and that c-Src Tyr 530 auto-phosphorylation is associated with a lowered catalytic status. A crystal structure of the c-Src-Ponatinib complex in a DFG-out state reveals unusual active-like features and provides a clear snapshot of c-terminal Tyr 530 intramolecular phosphorylation between enzyme and substrate acting kinases. Altogether these data indicate that c-Src must adopt an alternative conformation to the inactive-closed state independent of c-terminal Src kinase phosphorylation, and that a sequential and coordinated cis-to-trans phosphorylation switch between the activation and c-terminal segments simultaneously controls c-Src catalytic and non-catalytic functions (Cuesta and Contreras et al., submitted for publication). |
| 2. Gene fusion products are known drivers in human cancers, and are currently drug targets for personalised therapy. A second research line in the lab was established and directed at dissecting the functional and structural determinants for 2 RET oncogenic fusion products, CDC42-RET and KIF4B-RET. By applying an integrated approach, we demonstrated that full-length constructs behave like active dimers in solution. Auto-phosphorylation and enzymatic assays demonstrated fast kinetics compared to wild-type RET, and further phospho-proteomic characterisation by MS highlighted important roles for catalytic activity and substrate specificity through unexpected allosteric inputs by distant elements to the catalytic site (Hurtado et al., in preparation). |
| 3. A third research line focuses 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 the 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 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, and screening with both virtual and chemical libraries together with biophysical and biochemical tools for functional validation. Following this approach, we identified an allosteric interface in RET with good druggability score that can be potentially targeted with allosteric inhibitors. Furthermore, we uncovered a cryptic-pocket within the ATP-binding site that is exploited by highly specific second-generation type I RET inhibitors. This information will be crucial to designing and developing highly specific third generation RET inhibitors that are clinically successful and able to overcome refractory RET mutations (Shehata et al., in press). |
| 4. We initiated a new research line focused on the structural and functional characterization of human FAK. We want to explore how auto-phosphorylation drives the functional and conformational landscape of FAK, in a full-length setting, and how phosphorylation interferes with the assembly and interaction with substrates and signalling partners such as RET (both wild-type and oncogenic variants) and c-Src. Using a phospho-proteomic approach we already identified unexpected phospho-sites and revealed a previously unknown switch for FAK catalytic activation by N-terminal elements, which could be therapeutically exploited to design and develop next generation FAK inhibitors. |
Basic Research

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 these 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 to correct it and 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 (cryo-EM) and biochemistry in an integrative approach. Beyond fundamental research, this structural information provides the necessary detail for drug development.

“Using cryo-EM, we have captured the DNA mismatch repair machinery in multiple functional steps and studied conformational changes that these proteins undergo to recognise the mismatch and license the events that lead to repair.”

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. Recognition of an error in 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-EM, we captured 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 T. Sixma (Netherlands Cancer Research Institute) and M. Lamers (Leiden University).

DNA replication & 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, particularly their relationship to cancer. By combining in vitro reconstitution of DNA replication complexes with cryo-EM imaging, we aim to capture the replication machinery in different functional stages, allowing us to understand in detail its mechanisms and how it is regulated.

Department of Genome Integrity and Structural Biology Junior Group

Rafael Fernández-Leiro
Junior Group Leader
Post-Doctoral Fellow
Maria Dolores Moreno

Graduate Students
Ester Casajús, Samuel Míguez

Technician
Aynacelli Grande (TS)

† Azulita Superior (Advanced Degree)

Students in Practice
Emma Arawn (since September)
(University Politécnica de Madrid, Spain), Anna Martina Lippert (until June) (Universidad Internacional Menéndez Pelayo, Madrid, Spain).

Visiting Scientist
Sven Haagensen Stove (University of Bergen, Bergen, Norway)

Maria Sajo (Universidad Complutense de Madrid, Spain)

RESEARCH HIGHLIGHTS

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. Recognition of an error in 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-EM, we captured 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 T. Sixma (Netherlands Cancer Research Institute) and M. Lamers (Leiden University).

DNA replication & 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, particularly their relationship to cancer. By combining in vitro reconstitution of DNA replication complexes with cryo-EM imaging, we aim to capture the replication machinery in different functional stages, allowing us to understand in detail its mechanisms and how it is regulated.

Publications


FIGURE 1 Mismatch repair studies. The background of the image shows a cryo-EM micrograph of MutS protein (white circle) on DNA (grey strings). Circular insert shows a 2D class average of the protein after image processing. The bottom rectangular insert highlights the multiple structures solved in the successful steps of the DNA repair process: MutS loading and DNA scanning, mismatch binding, clamp formation, and MutL recruitment and sliding clamp formation. These steps control the licensing of the DNA mismatch repair pathway.

• Publications
**RESEARCH HIGHLIGHTS**

**High-order genetic interactions between two genes**

To better understand the dominance and dosage sensitivity of cancer genes, we systematically quantified the interactions between mutations and copy number changes. We found that many cancer genes do not behave like consistent models, but have activity-fitness functions that change across cancer types. To gain a better understanding of this switch, we identified one cause of these changes to be mutations in trans, as higher-order interactions. Most trans interactions were found to be in the same cancer signaling pathways and to share their functions. Our manuscript (in revision) will report the first analysis of high-order interactions in cancer genomics, based on studies conducted in collaboration with F. Supek (IRB Barcelona, Spain) and B. Lehner (CRG, Barcelona, Spain).

Furthermore, we expand this concept to 2 different aspects: (1) time-dependent high-order interaction with germline variants, and (2) condition-specific high-order interactions with cancer-causing factors.

**Looking beyond genomics to see cancer using TP53 and KRAS model**

Over several years, more than 1,000 somatic drivers have been discovered by analyzing huge amounts of genomics data. However, we need a next-level analysis to obtain a complete view of their working model in cancer. To overcome this missing link, we proposed to map position-specific protein interaction networks by integrating genomics and large-scale Y2H screening. Specifically, we focused on the 2 most important cancer genes, TP53 and the RAS family. In 2021, we created clones for > 10 TP53 hotspot variants and > 50 RAS family variants and conducted large-scale Y2H screening with a complete human library. Our screening results will provide the systematic protein-interaction networks to show how protein interactions can be differentially changed depending on mutations. This will point the way to identifying new precision treatments based on differential protein interaction networks across patients. These studies were carried out in collaboration with Yang’s Lab (CRAG, Barcelona, Spain).

**OVERVIEW**

Cancer is one of the most complex human diseases, involving genetic, environmental, and even unknown factors. Over the past several decades, our knowledge of cancer has rapidly accumulated thanks to different omics technologies, including genomics and proteomics. However, we still lack a complete understanding of the cancer fitness landscape across conditions. For example, how do cancer genes change their working models of tumour progression depending on cancer types or contexts? What kind of trans-interactions exist between 2 genes or many genes beyond single-gene level alterations? How can protein complexes or interactions be perturbed depending on different mutation positions? Based on large-scale genomics and proteomics analyses, we aim to pursue these questions.

“High-order interactions identify mutations that change the dominance and dosage sensitivity of cancer genes. These high-order interactions in the same pathway can be alternative evolutionary paths.”

---

**PUBLICATIONS**

ANNUAL REPORT 2021

Basic Research

Our main research areas include:

- To improve outcomes in these tumours.

- Instability (CIN). By therapeutically targeting CIN, we aim to exploit - extreme chromosomal instability, which we can exploit - extreme chromosomal instability (CIN). By therapeutically targeting CIN, we aim to improve outcomes in these tumours.

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, which we can exploit - extreme chromosomal instability (CIN). By therapeutically targeting CIN, we aim to improve outcomes in these tumours.

Our early Christmas present — the cellenONE single cell sorter — now allows us to perform single cell DNA sequencing of human tissue samples, so watch this space for exciting new data in 2022!

On the computational side, David, a PhD student from Luis Malumbres’ Lab, and Blas is continuing as a PhD student building a large collection of induced CIN models in collaboration with the CNIO laboratories of Marcos Malumbres, Ana Losada, Felipe Cortés-Ledesma, Óscar Fernández-Capetillo, Juan Ménendez, and Miguel Ángel Quintela-Fandino. As part of this project, Blas and Bárbara Hernando developed the first computational tool to come out of the lab, CiSpire, which identifies related cell lines based on their genome-wide copy number profile (preprint here: https://doi.org/10.1101/2021.09.28.462193).

Following on from her excellent master’s thesis identifying mismatch repair deficient ovarian cancers (ongoing project led by Maria José), Maria Escobar started her PhD where she will be using ongoing CIN to predict drug response in organoids and predict risk of progression in premalignant lung lesions. Alice Cádiz joined the lab with a training fellowship and is supporting all lab-based projects and keeping our databases in check.

On the computational side, David, a PhD student from Luis G. Paz-Ares’ group, joined the team and will work on mutational signatures in lung cancer. Bárbara joined as a post-doc and will be using ongoing CIN to predict drug response in organoids and predict risk of progression in premalignant lung lesions.

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

“...We have completed a proof-of-concept study showing that different types of CIN can be studied at high resolution, using single cell DNA sequencing, and induced via genome-editing.”

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

2021 saw the Computational Oncology Group shed its pandemic shackles and welcome 5 extremely talented new lab members and complete crucial proof-of-concept experiments.

Both Blas Chaves and Agustín Sánchez completed their master thesis projects. Agustín employed machine learning techniques to explore the relationship between DNA methylation and copy number signatures, and Blas quantified patterns of DNA copy number change at single cell resolution induced by knock-out of CDK18. Agustín has since taken a position in the Marcos Malumbres’ Lab, and Blas is continuing as a PhD student building a large collection of induced CIN models in collaboration with the CNIO laboratories of Marcos Malumbres, Ana Losada, Felipe Cortés-Ledesma, Óscar Fernández-Capetillo, Juan Ménendez, and Miguel Ángel Quintela-Fandino. As part of this project, Blas and Bárbara Hernando developed the first computational tool to come out of the lab, CiSpire, which identifies related cell lines based on their genome-wide copy number profile (preprint here: https://doi.org/10.1101/2021.09.28.462193).
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 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 research compounds for biophysical and functional studies. Furthermore, we use NMR to characterise the metabolic profiles of biofluids, cell growth media, and cell and tissue extracts from both animal models of cancer and human samples. In addition, in 2021, we successfully installed a multiple well microplate reader that includes numerous detectors (absorbance, fluorescence, intensity, polarisation and time resolved modes, luminescence, alpha-screen, etc.) with excellent sensitivity for both in-solution and adherent cells measurements.

The Unit provides a broad range of instrumentation for the biophysical characterisation of biomolecules and their interactions, including spectrophotometers, a fluorimeter, isothermal titration and differential scanning calorimeters, a circular dichrograph, dynamic and multi-angle static light scattering devices, two biosensor instruments — surface plasmon resonance (SPR) and bilayer interferometry (BIL) — and a multiple well microplate reader with numerous detectors. Research groups mostly from, but not limited to (i.e., Haematological Malignancies Clinical Research Unit, Monoclonal Antibodies Unit, Molecular Oncology Group and the Experimental Therapeutics Programme – ETP) the Structural Biology Programme used these technologies throughout the year.

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 the Experimental Therapeutics Programme), as well as for metabolite quantification that in 2021 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).

In collaboration with the latter group, we also implemented protocols to detect intracellular metabolites derived from the chemotherapeutic drugs 5-F-uracil and 5-F-uridine, using 19F-NMR spectroscopy. For example, FIGURE 1 shows representative spectra that enable characterisation of the metabolic conversion of the drug into different nucleotides and activated sugars, and how it is affected by overexpressing an enzyme involved in purine metabolism and a mutant thereof, as well as the effect of chemical inhibitors acting upstream in the metabolic pathway. 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.

**OVERVIEW**

“In 2021 we installed a new and versatile multiple well microplate reader that will allow, among other HT applications, quantification, inhibitor discovery, affinity determination and enzyme kinetics on liquid samples, as well as viability and other functional assays on cell samples.”
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, (iv) to maintain the scientific computing facilities at the CNIO, and (iv) to provide training in bioinformatics tools and methods.

**OVERVIEW**

Bioinformatics is a key discipline for understanding the cancer genome and 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, (iv) to maintain the scientific computing facilities at the CNIO, and (iv) to provide training in bioinformatics tools and methods.

**SELECTED PUBLICATIONS**


2. Pozo F, Martínez-Gómez L, Walsh TA, Rodríguez JM, Domenico T, Abascal F, Addisol M, Tress ML (2021). Beyondcell: a comprehensive bioinformatics pipeline that performs basic and advanced single-cell RNA-seq analysis and TRIFID, a method for classifying the functional importance of splice isoforms (Pozo et al. 2021). Additionally, the BU served as an active node of the European network ELIXIR (https://www.elixir-europe.org/), leading the ELIXIR Cancer Data Focus Group to provide the framework and expertise for the systematic analysis and interpretation of cancer genomes. With regard to academic and knowledge-transfer activities, we co-organised the Master’s degree programme in Biocomputing Applied to Personalised Medicine and Health at the National Institute of Health Carlos III (Master in Bioinformática Aplicada a Medicina Personalizada y Salud, ENS-ISCHI).

**RESEARCH HIGHLIGHTS**

In 2021, the Bioinformatics Unit published more than 22 peer-reviewed articles as a result of our ongoing research projects and scientific collaborations (see the full list of our activities on our website: https://bioinformatics.cnio.es/). We studied cancer cell biology and drug response at single-cell resolution. To this end we developed Beyondcell (FIGURE 1), a computational methodology to identify tumour cell subpopulations under drug exposure, thereby revealing sensitive, innate, and acquired drug resistant cancer cells. Through this method we are able to propose possible treatment strategies to overcome such resistance and to identify drug-response markers. Using Beyondcell, we characterised single-cell variability in drug response in 198 cancer cell lines, finding recurrent patterns of drug heterogeneity and their relationship with the cells’ functional status. Our software also allowed us to explore inter- and intra-tumour heterogeneity, linking it to clinical drug response data and to successfully predict responders and non-responders to immunotherapy among melanoma patients (Fastero-Torre et al. 2021). Beyondcell is fully accessible at https://gitlab.com/bu.cnio/beyondcell.

During 2021, our group also assessed the clinical importance of tandem exon duplication-derived substitutions (Martínez-Gómez et al. 2021) and implemented Bolito (García-Jimeno et al. 2022), a comprehensive bioinformatics pipeline that performs basic and advanced single-cell RNA-seq analysis and TRIFID, a method for classifying the functional importance of splice isoforms (Pozo et al. 2021). Additionally, the BU served as an active node of the European network ELIXIR (https://www.elixir-europe.org/), leading the ELIXIR Cancer Data Focus Group to provide the framework and expertise for the systematic analysis and interpretation of cancer genomes. With regard to academic and knowledge-transfer activities, we co-organised the Master’s degree programme in Biocomputing Applied to Personalised Medicine and Health at the National Institute of Health Carlos III (Master in Bioinformática Aplicada a Medicina Personalizada y Salud, ENS-ISCHI).

**FIGURE 1**

(A) Beyondcell workflow. (B) Beyondcell identified the presence of therapeutic clusters (TCs) in BRAF-inhibitor resistant human melanoma cells.

Trametinib had a positive Beyondcell score in all TCs, with a higher sensitivity in TC5, showing it could be proposed to target BRAFi-unresponsive cells.
The main objective of the Electron Microscopy (EM) Unit is to provide scientific-technical support to researchers to answer their biological questions using different transmission EM techniques. We regularly use negative staining and cryo-EM and help with image processing by performing 2D analysis and 3D reconstruction. We also offer support for choosing adequate EM techniques and performing sample preparation on different types of EM grids. Furthermore, we provide the necessary training for the use of our microscopes and auxiliary equipment. More advanced studies are typically delivered through research collaboration.

**OVERVIEW**

“We dedicate our main effort to ensuring efficient access to and use of existing infrastructure in the Unit. We also provide personalised scientific support and training for researchers.”

**PUBLICATIONS**


- Rodríguez CO, Escudero-Bravo D, Díaz L, Barrioscoso P, Garcia-Martín C, Silabert JG, Boskovic J, Gualler V, Emeli-Murugam E, Llorente D and Palacin M (2021). Structural basis for substrate specificity of the heteromeric amino acid transporter HAT family of neutral amino acid transporters. These findings provide the structural bases for mutations in LAT2/CD98hc (HAT member) that alter substrate specificity and that are associated with several pathologies. Finally, in collaboration with the Genome Integrity and Structural Biology Group, we further improved our set-up to use a cryoEM as a tool for drug discovery.

In our studies, we take advantage of the continuous technical advances in cryogenic electron microscopy (cryoEM). Specifically, we use single-particle cryoEM to elucidate the structures of macromolecules at near atomic resolution. At the CNIO we have a 120 kV Tecnai G2 Spirit microscope equipped with a TVIPS CMOS detector that is used to obtain images of negatively stained samples, to screen vitrified samples, and for small-scale data collection. For medium resolution structural studies, we use an EM-2200FS cryo-electron microscope equipped with a 200 kV field emission gun and a K3 direct electron detector.

Our scientific activity throughout 2021 involved collaborations with the research groups of the Structural Biology Programme, as well as with groups from other Programmes and with scientists outside the CNIO. For example, together with CNIO’s Macromolecular Complexes in DNA Damage Response Group, we contributed to the analyses of secreted extracellular vesicles (EVs) that influence the tumour microenvironment and promote distal metastasis. In particular, we imaged melanoma-secreted EVs that have been associated with lymph node, pre-metastatic niche formation in murine models. With the Macromolecular Complexes in DNA Damage Response Group, we pursued our work to structurally characterise several protein complexes e.g., different RUVBL1/2 and DNA repair complexes. Our collaboration also continued with M. Palacín’s group (IRB Barcelona), with whom we contributed to revealing the molecular mechanisms controlling substrate specificity within the heteromeric amino acid transporter (HAT) family of neutral amino acid transporters. These findings provide the structural bases for mutations in LAT2/CD98hc. (HAT member) that alter substrate specificity and that are associated with several pathologies. Finally, in collaboration with the Genome Integrity and Structural Biology Group, we further improved our set-up to use a cryoEM as a tool for drug discovery.

In particular, we imaged melanoma-secreted EVs that have been associated with lymph node, pre-metastatic niche formation in murine models. With the Macromolecular Complexes in DNA Damage Response Group, we pursued our work to structurally characterise several protein complexes e.g., different RUVBL1/2 and DNA repair complexes. Our collaboration also continued with M. Palacín’s group (IRB Barcelona), with whom we contributed to revealing the molecular mechanisms controlling substrate specificity within the heteromeric amino acid transporter (HAT) family of neutral amino acid transporters. These findings provide the structural bases for mutations in LAT2/CD98hc. (HAT member) that alter substrate specificity and that are associated with several pathologies. Finally, in collaboration with the Genome Integrity and Structural Biology Group, we further improved our set-up to use a cryoEM as a tool for drug discovery.
The goal of fragment-based screening is to expose protein crystals to libraries of fragments and to solve the crystal structures of the complexes. Our first target was the dimerization domain of TRF1 (Telomeres and Telomerase Group). We identified 3 well-defined fragments bound to the protein that could be further exploited to develop new inhibitors.

Our Unit works closely with the Experimental Therapeutics Programme on several projects: human TRF1 dimerization domain, the kinase domains of human MASTL and human DDR1 for biochemical and structural analyses. Furthermore, to support drug discovery projects, we performed several thermal shift assays (thermosfluor) in the presence of compounds developed in the Medicinal Chemistry Section.

We provide the proteins needed by the CNIO Monoclonal Antibodies Unit to generate antibodies, including the CD85 family, CSF3R, CLLU1, RANK, TACI and PILRA, among others. The Unit is also engaged in several internal collaborations with other CNIO groups, providing them with recombinant proteins for biochemical and/or cell-based functional assays.

Throughout 2021, the Unit also sustained its own scientific projects. We continued working on targeting the function of the Mdm2-MdmX E3 complex activity (NIH funded) as well as on the production of a T-cell recruiting bispecific antibody, ATTACK (funded by Retos Colaboración). The Unit is also taking part in 2 projects supported by BBVA Foundation grants. The first is a collaborative project with the Biomedical Application of Radioisotopes Unit at CIEMAT and the CNIO’s Molecular Imaging Unit to develop new antibody-based positron emission tomography (immunoPET) imaging tools for tumour visualisation. The second project, carried out in collaboration with the Immuno-oncology and Immunotherapy Unit at the Hospital de la Carretera, has resulted in the generation of 2 new synthetic nanobodies capable of targeting the spike protein of the SARS-CoV-2 virus. These nanobodies will soon be tested in mouse models at the Poxvirus and Vaccine Laboratory by M. Esteban’s group at the CNB.