Scientific Annual Review

A collection of scientific advances in the research lines of CIC bioGUNE

Issue 4

January 18th 2023

www.cicbiogune.es

CHEMISTRY

Click chemistry and biorthogonal chemistry. *p. 3*

GENETICS

Multi-omics and large-scale genetic studies. *p. 5*

VACCINES

Lyme vaccines and adjuvants. *p.6 and p.7*

The Metabolic Rewiring in Longevity

The rejuvenation strategies in aging, *p. 8*

CIC bioGUNE MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

Contents

The Center

CIC bioGUNE is a collaborative research center focused on Life Sciences, from Chemistry to Biomedicine, from basic to translational Science. Our cutting-edge scientific activity concentrates on discovering the molecular bases and mechanisms of disease to promote development of advanced therapies. Our activity explores four main biomedical research themes like Cancer, Metabolic Diseases, Rare Diseases, and Infectious Diseases organized in two research programmes "*Metabolism and Cell Signaling in Disease*" and "*Molecular Recognition and Host-Pathogen Interactions*".

With our collaborative philosophy we are deeply engaged in multidisciplinary research collaborations with local, national, and international colleagues and technology experts. The center is impinged in a heterogeneous network of Academic and Clinical Entities, Research and Technology Centers, and is member of the Basque Research and Technology Alliance (BRTA).

The scientific activity is supported by cutting-edge infrastructures and technology platforms, including advanced equipment for nuclear magnetic resonance (NMR), now recognized as ICTS, electron microscopy, a facility for monoclonal antibody production, as well as different core technology platforms where genomes, proteomes and metabolomes can be analyzed.

General View

Direction

CIC bioGUNE activities are strongly related to our specific mission: to build up an EU-referent knowledge pole in biosciences, which should be able to favour the development of the emerging sectors in the bioscience and health fields, and the incorporation of the proper technologies to be able to enhance the competitiveness of the corresponding industrial (biotech, pharma, etc) sectors. Specifically, CIC bioGUNE acts with a strong commitment of collaboration and coordination with the rest of social and scientific agents within BRTA and in the Basque Country to optimize the existing capacities, and jointly conform an integrative scientific and technological offer of excellence. This offer should be able to boost the evolution of the economy by strongly increasing its intrinsically high added value. Our research activities cover from the gene to animal models of cellular processes through the determination of biomolecular structure and assembly and the elucidation the key mechanisms and interaction patterns at the highest resolution. Our scientific objectives are transversal and target the complete characterization of the molecular basis of protein-based processes in human pathophysiology and immune defence, cell proliferation, and development. The final aim is to translate our findings to the clinic, with special interest in precision medicine.

Nobel Laureate & Emergency Calls: Glycoscience at the phone

Ana Ardá, June Ereño-Orbea, Luca Unione, and Jesús Jiménez-Barbero, *Chemical Glycobiology Laboratory*

2022 will remain as an important year for glycosciences, marked by the Nobel Prize in Chemistry to Prof. Carolyn Bertozzi. She took the fantastic tool of click chemistry, developed by Prof. Sharpless and Prof. Meldal (the other 2022 Nobel laureates), and adapted it to be used in human cells, creating the concept of bioorthogonality to chemically modify glycans in living cells without interfering with normal cellular functions, breaking down barriers in glycoscience. This scientific breakthrough is important for understanding the function of glycosylation on the cell surface, for the design of new probes for in vivo imaging of glycans or for the generation of new biotherapeutic molecules.

Society has received two main urgent calls along 2022: global pandemic and environmental impact. Dragging the Covid19 tail, this year has seen the monkeypox outbreak (WHO) and was exposed to the risk of the Camel flu (MERS-CoV). Using the combination of glycan AND infections in PUBMED, over 3500 research articles appeared in 2022. Members of the group have contributed to address the call by showing which glycans participate in such infectious events[1,2,3,4,5].

Key tools for glycobiology, such as genetic cell glycoengineering[6] and biorthogonal chemistry have continued evolving with outstanding applications[7,8,9]. At the same time, new evidences and clues on the involvement of glycan-lectin interactions in general[10,11], and glycan-galectins interactions in important processes related to health and disease, such as bacterial host immunity[12], viral infectivity[13], and cancer development and immunity[14,15] have been revealed, highlighting their potential use as therapeutic targets[16,17,18].

Over the years, scientists have struggled to produce large and complex glycans of well-defined structures, a fundamental step to advance in our knowledge about their structure-functional relationship. In this sense, after the solid-phase automated glycan assembly strategy, developed by Prof. Seeberger, in 2022 an automated glycan assembly strategy in solution has been proposed[19], reporting the chemical synthesis of a large library of biologically relevant carbohydrates including the "longest and largest ever-made homogeneous polysaccharide in the history"[20]. A fantastic alternative or complementary strategy for glycan synthesis, is to exploit the professional glycan makers in Nature, the glycoenzymes. This is however not straightforward, as many of them are difficult to produce. The recently reported universal glycoenzyme biosynthesis pipeline[21], dubbed SIMPLEx (solubilization of IMPs with high levels of expression), enables the topological conversion of secretory and membrane-bound (human) proteins into water-soluble variants. The method was used to the soluble expression of 100 functional GTs, within E. coli cytoplasm, in the 5–10 mg/L range.

The access to a plethora of unique glycans in large scale with a minimal environmental impact will push the boundaries of glycosciences in health, energy, and materials science. 2023 will be an exciting, sweet year.

The Aftermath of AlphaFold and the Nobel Prize in Chemistry 2022

Gonzalo Jiménez-Osés, *Computational Chemistry Laboratory*

Two of the most challenging and successful research lines explored in our group, namely functional protein design and site-selective protein modification have very recently experienced an unprecedented boost – and the corresponding hype – within the scientific community, due to the astounding impact of DeepMind's program AlphaFold and the award of the Nobel Prize in Chemistry 2022 for the development of click chemistry and biorthogonal chemistry. In the last years, our group has published a number of papers on site-selective modification of proteins and therapeutic antibodies, which is entirely based on the concepts of click and bioorthogonal chemistry, mainly fast, selective and biocompatible cycloaddition and conjugate addition reactions.

Let's focus on the first aspect. Since its unveiling at the Critical Assessment of Protein Structure Prediction contest in 2020 (CASP14), AlphaFold has become omnipresent in life-sciences research. The Alphabet (i.e. Google)-owned company DeepMind released the software's underlying code in 2021[22,23] and published an essay this year so that anyone could run the program with at least some basic knowledge on how it works[24]. Moreover, an AlphaFold database updated this year (www.alphafold.ebi.ac.uk) holds predicted structures of varying quality for almost every protein from all organisms represented in genome databases, a total of more than 200 million proteins from over 10 million species including the complete human proteome[25]. These predicted structures have been now made accessible also through the most widely used online servers including the Protein Data Bank (PDB) and UniProt. This means that for every known sequence in the UniProt data resource there will be either an experimentally determined structure in PDB, or an AlphaFold model in the AlphaFold database. This development represents a step-change for molecular biology: for the first time in history, for almost every protein of known sequence, a high-quality 3D model will be readily available.

Following in the footsteps of DeepMind, another Artificial Intelligence (AI) laboratory from tech company Meta (owner of Facebook, Instagram, and WhatsApp) released the structures of more than 600 million putative proteins in a database called the ESM Metagenomic Atlas (www.esmatlas.com). The structures are for proteins predicted to exist based on genetic data from large-scale metagenomic screens of soil, seawater, and other sources. The proteins themselves have yet to be isolated or identified using proteomic methods. The Meta AI algorithm used to make the new protein models, ESMFold[26], is not as accurate as AlphaFold, but it is faster. Such high speed is a result of how the tool predicts protein structures using a language model trained on sequence data – the order of amino acids in the linear chain that make up a protein. Normally, language models are trained on large volumes of text. To apply them to proteins, researchers instead fed the AI sequences of known proteins, which can be written down as a series of letters, each representing one of 20 possible amino acids. The network then learnt to fill in the sequences of proteins in which some of the amino acids were obscured.

A legion of research groups around the world are contributing to augment the capabilities of AlphaFold, such as enriching the predicted models with ligands and cofactors – a feature that is missing in the original code – and producing the AlphaFill databank with more than 12 million transplants performed on nearly a million AlphaFold models (www.alphafill.eu)[27]. Experts in crystallography and cryoEM have developed an iterative procedure in which AlphaFold models are automatically rebuilt on the basis of experimental density maps and the rebuilt models are used as templates in new AlphaFold predictions, with astonishing results[28].

In a post-AlphaFold world, new and exciting challenges wait ahead. Some of the most important and yet unsolved milestones in biology involve determining how proteins interact with other molecules such as other proteins (the 'interactome' problem) or drugs (the 'drug discovery' problem) and predicting the multiple shapes that some proteins can assume (the 'dynamic personalities of proteins' problem). Approaches to these problems are being proposed in light of the extreme power exhibited by AI. Scientists are now using AlphaFold to model thousands of binary protein interactions from the human

proteome[29], and to predict the structure of very large protein complexes combining AlphaFold and Monte Carlo tree search[30]; these studies begin to paint a picture of the complexities of human protein interactions, exploring how disease-causing mutations might change those relationships, suggesting potential disease mechanisms. Not by chance, Isomorphic Labs – also subsidiary of Alphabet – is a drug discovery company just created building on the AlphaFold breakthrough which uses artificial intelligence to – in their own words – transform drug discovery and treat and cure disease faster. Whether these promises grounded in the seemingly infinite power of AI will be fulfilled or not, and as the famous song says…"Only Time Will Tell".

The power of One: how a minority can change the face of cancer pair of Glass

Arkaitz Carracedo, *Cancer Cell Signaling and Metabolism Laboratory*

2022 has demonstrated that access to high resolution and high throughput strategies can help explain critical clinical questions in oncology. Studies all over the globe this year have provide clinically relevant answers by focusing on the minority of cells that inhabits the tumor. Single cell technologies enable individualized analysis of the transcriptome in complex tissues. In the recent past we could only assume that heterogeneity existed in tissues and tumors through targeted analysis of a few selected markers. However, we can now identify, quantify and catalogue intratissue heterogeneity in all its complexity. A few examples of the repercussion of this new perspective are highlighted below:

The high-relapse cells as the source of colorectal cancer recurrence. The mechanisms behind resistance to anticancer treatment is a predominant focus in cancer research. This year, the lab of Eduard Batlle identified a minority of tumor cells responsible of the recurrence of colorectal cancer. Importantly, they defined markers that would help detect these cells, and demonstrated that genetic ablation of this cell fraction sharply increased therapeutic success in preclinical models. This is a further refinement of the well-accepted hypothesis that proposes that specific cell phenotypes have unique capacity to disseminate or endure microenvironmental stresses, and these findings open a new opportunity to improve the management of aggressive disease[31].

Castration resistance before castration occurs. Androgen targeting therapies are among the first targeted therapies to be implemented. Despite their effectiveness in prostate cancer and the continuous development of new generation targeting molecules, resistance often emerges in patients. It was thought that the selection pressure of the therapy was the main driver of castration-resistant prostate cancer cells. However, single cell resolution and cell trajectory inference has recently demonstrated that there might be instances where tumors harbor pre-existing castration-resistant tumor cells. Their role and their relevance in the progression of the disease is an area of research that remains to be exploited[32].

The castrated immune cells. Hormone-dependent cancers offer a unique therapeutic opportunity, namely, to inhibit hormone receptors to curb tumor growth and progression. The effectiveness of anti-androgenic therapies in prostate cancer has been ascribed to the cancer-cell autonomous of the drug. However, this year Yang et al., reported that CD8 T cells are targets of androgen receptor blockade therapies and that their modulation contributes to the therapeutic efficacy of this sype of agents[33].

The awakening of cancer in our sleep. Tumor cell dissemination to circulation is a process that was thought to be continuously active since early stages of cancer development. Indeed, circulating tumor cells serve as a surrogate marker of cancer aggressiveness and therapy response. Surprisingly, the group of Nicola Aceto recently reported that the number of circulating cancer cells in patients oscillate in a circadian fashion. Specifically, during the sleep period the number of circulating tumor cells was dramatically increased. Beyond the repercussion of this finding for cancer biology, the circadian regulation of cancer cell spreading could offer new therapeutic opportunities if we understand the mechanisms that govern this process[34].

New partners in crime. Tumors consist of cancer cells, but are also composed non-transformed cells that compose the immune and non-immune stroma. Recent studies have demonstrated that there is another dimension to cancer composition, which includes the microbiota. A recent study exploited spatial single cell technologies to reveal that there is a cartography of tumor-stroma-microbiota interaction, so that the bacteria inhabit tumor areas that are less vascularized and heavily immunosuppressed. In addition, tumor-microbe interaction favors individualized migration and invasion, a feature of cancer cell aggressiveness. This new view of the tumor architecture opens intriguing questions about how tumors evolve as an ecosystem, and encourage the study of this aspect to promote future cancer research[35].

Tackling Cancer Heterogeneity

Maria dM Vivanco and Robert Kypta, *Cancer Heterogeneity Laboratory*

Tumour heterogeneity implies that cancer cells do not function in isolation. Tumours include epithelial cells, stromal cells, and immune cells, as well as small populations of normal stem cells and cancer stem cells (CSCs). In particular, the presence of CSCs proves a hurdle for therapy while represents a source of fitness for the tumour. CSCs are therapy-resistant and their enrichment after treatment leads to tumour recurrence and metastasis. We previously showed that normal and cancer stem/progenitor cells share some features, such as expression of Sox family transcription factors and increased Wnt signalling activity. While several drugs to improve cancer survival are in clinical trials, the challenge is to use them without compromising normal tissue homeostasis. This challenge can be met by developing more specific, patient-tailored inhibitors. The development of more specific Wnt inhibitors remains a goal for many groups. An epitope-directed selection strategy was recently used to identify Wnt receptor-selective antibodies[36]. This is reminiscent of the epitope-directed strategy we are using to exploit our knowledge about the epitope of our Wnt-11 antibodies to target other Wnt family members. Specific activators of Wnt signaling are also being sought for treatment of some diseases. Martin et al.[37] generated mutations in Wnt-7B that adjusted its signaling specificity so that it can be used to treat CNS disorders without side-effects.

An exciting new avenue for exploration in 2022 is the discovery that NALCN (sodium leak channel non-selective protein) is key regulator of metastasis, not only of cancer cells but also of non-malignant cells[38]. This observation unmasks a potential new target for antimetastatic therapies, and, as the author of the study Richard Gilbertson noted, "if validated through further research, this could have far-reaching implications for how we prevent cancer from spreading." The studies ongoing in our laboratory using intravital imaging of cells expressing fluorescent Wnt signalling reporters during metastasis of human cells in the chick embryo (manuscript accepted for publication) may prove useful to explore this area in more detail. Furthermore, our studies with normal breast tissue[39] in the context of breast cancer prevention may also be influenced and enlightened by these intriguing findings. Over the last few years, the relevant implications of breast cancer heterogeneity have become more evident, both among patients and within each

tumour. Several ongoing clinical trials now aim to identify ways to de-escalate treatment and stratify patients using biomarker guidance[40]. Precise prognosis is key for selection of the most effective adjuvant therapy in breast cancer. Molecular subtyping is increasingly used to complement immunohistochemical and pathological classification and to predict recurrence. We have compared molecular subtyping and pathological classification data in a cohort of 143 breast cancer patients and identified some critical discrepancies between molecular and pathological subtyping of the samples, including misclassification of HER2-positive tumours and the identification of a significant percentage of genomic high risk T1N0 tumours[41]. These results suggest that according to pathological subtyping, a considerable number of breast cancer patients would not receive the appropriate systemic therapy. Our findings support the need to determine the molecular subtype of invasive breast tumours to improve breast cancer management.

We continue to develop and optimise tools that target aberrant Wnt signalling in cancer and in other disease contexts, such as Crohn´s disease[42]. We are further exploring how other Wnt family members can be targeted and how new nanotechnologies can be combined with targeting antibodies to deliver therapeutic cargo. In addition, we are following on our development of Sox family inhibitors and studies of the crosstalk between tumour cells and the tumour microenvironment to obtain a deeper understanding of the process of development of resistance and reveal novel therapeutic targets against breast cancer.

Uncovering the mysteries of complex diseases with multi-omics and genome sequencing

Urko Martinez Marigorta, *Integrative Genomics Laboratory*

The Integrative Genomics lab studies the genetics of complex diseases. We use multi-omic profiles of patients to better understand how diseases develop. Our focus is on understanding the variability in how patients develop diseases and on developing predictors of prognosis and response to treatment. These steps are essential for discovering and characterizing biomarkers that can be used in the clinic to track prognosis and disease status. We will discuss here three important developments in genomic medicine.

The first one involves the emergence of multi-omic efforts as the strategy to "unlock the secrets of disease". Current human genomics excels at identifying variants associated with disease but translating this information into improved understanding of the underlying mechanisms is still challenging. Multi-omic efforts, which study multiple layers of biological data simultaneously, can provide a more comprehensive and nuanced understanding of complex diseases. By integrating data from multiple omics levels, researchers can gain insights into the underlying mechanisms of disease and identify potential targets for intervention. This strategy helps to identify important regulatory elements and pathogenic genes, a key step towards development of personalized medicine approaches. For example, Liu et al. performed a large genetic study for kidney function in 1.5 million individuals[43], and integrated the results with transcriptomic, epigenomic, and single-cell datasets to prioritize over 500 genes as likely causes of kidney dysfunction. This study highlights the potential of multi-omics data and analytical tools in translating GWAS findings into a better understanding of the mechanisms behind complex diseases.

Genome sequencing is proving useful for identifying genetic risks for common diseases and improving diagnostic yield for rare diseases. For instance, Guzauskas et al.[44] found that genomic screening for Lynch syndrome had an incremental cost-effectiveness ratio of $132,200 per quality-adjusted life-year. This suggests that genome sequencing for Lynch syndrome and other genetic conditions can be cost-effective. In turn, The UK 100,000 Genomes Project showed that whole-genome sequencing resulted in diagnoses in 35% of individuals with a monogenic cause and 11% with a complex cause. 25% of participants who received a genetic diagnosis had immediate clinical actionability. These results support the use of genome sequencing for the diagnosis of rare diseases.

Third, the new advances in polygenic risk estimation. Mars et al.[45] showed that polygenic risk scores (PRSs) and family history (FH) provide independent and complementary information on inherited disease susceptibility. The PRSs explained, on average, 10% of the effect of FH, but FH only 3% of the PRSs, with the PRSs being independent of both early- and late-onset family history. The PRS estimates stratified risk similarly in individuals with and without positive FH: a high PRS conferred a considerably elevated risk, whereas a low PRS compensated for the effect of FH. These findings provide opportunity for generalization in clinical practice through incorporation into existing risk assessment tools or to develop tailored screening tools based on family history and polygenic risk.

Large-scale genetic studies for precision medicine in gastroenterology

Isotta Bozzarelli, Mauro D'Amato and Cristina Esteban-Blanco, *Gastrointestinal Genetics Laboratory*

The Gastrointestinal Genetics Lab (MDAlab) focuses on the identification of genetic risk factors and mechanisms, aiming for translational application and improved therapeutic precision in gastroenterology. While we participate in multiple consortia harnessing the power of omics technologies for the discovery of biomarkers in inflammatory bowel disease, microscopic colitis, functional dyspepsia and other gastrointestinal (GI) conditions, we are at the forefront of genetic research in irritable bowel syndrome (IBS). This condition affects one in ten people with bowel symptoms including recurrent abdominal pain, bloating, diarrhea, and constipation; it strongly impacts quality of life, is a leading cause of work absenteeism, and consumes 0.5% of the healthcare annual budget. The heterogeneity of IBS, the lack of specificity of its various definitions, and a poor understanding of the underlying pathophysiological mechanisms all hamper the development of effective therapeutic options. Genetic predisposition is demonstrated for IBS, though poorly investigated until recent years, primarily due to our contributions to this line of research. As we recently reviewed[46], large-sample sizes from population-based biobanks have recently enabled the development of well-powered genetic studies in IBS, providing unprecedented opportunities for long-awaited research. We pioneered IBS genetic research by conducting genome-wide association studies (GWAS) to identify IBS genes, reporting their association with disease risk and/or altered bowel functions (including visceral hypersensitivity, gut motility etc) in multiple independent cohorts[47],[48].

Our most recent studies suggest that a subset of "organic" IBS might have been identified, accounted for by a group of carbohydrate "malabsorbers" who carry hypomorphic variants in the sucrase-isomaltase gene (SI), coding for an intestinal disaccharidase that digests sucrose and starch. Similar to the congenital form of sucrose intolerance (CSID, due to homozygous SI loss-of-function mutations), we have shown that SI variants coding for enzymes with reduced disaccharidase

activity (hypomorphic variants) confer increased risk of IBS, possibly via accumulation of undigested disaccharides in the large bowel[49,50]. Additionally in a pilot retrospective study of IBS-D patients, we recently found that SI carriers are also less likely to benefit from a low-FODMAP diet, as their symptoms may be due to malabsorption of sucrose and starch, which are not specifically targeted by this diet[51].

A different prioritized line of research in the MDAlab takes has been established following a novel "outside the box" approach to understanding bowel (dys)function via the genetic characterization of the so called endophenotypes. Indeed, we looked into the genetics of intestinal (dys)motility via large-scale GWAS meta-analyses of stool frequency and consistency as proxies for gut transit time[52]. Analyses currently ongoing highlighted several mechanisms involving ion channel activity, synaptic transmission, enteric nervous system and digestive tract development. Druggable pathways were also identified, including some already targeted by current IBS medications.

Our studies reveal genetic factors and mechanisms crucial to IBS pathophysiology, providing a rationale for developing personalized therapeutic and dietary approaches in molecularly-defined subgroups of patients. These new routes and novel actionable targets will positively impact the clinical management of IBS and the dysmotility syndromes.

Cancer immunotherapy meets glycobiology

Asís Palazón, *Cancer Immunology & Immunotherapy Laboratory*

Cancer immunotherapy is a type of treatment that uses the body's own immune system to fight cancer. There have been significant advances in recent years. Some of the most well-known cancer immunotherapies include immune checkpoint inhibitors[53], which block proteins that inhibit the immune system's ability to recognize and attack cancer cells, and CAR-T cell therapy[54], which involves genetically engineering a patient's own immune cells to recognize and kill cancer cells. Other promising areas of research in cancer immunotherapy include tumor vaccines, oncolytic viruses, and combination therapies that combine different immunotherapies or combine immunotherapy with other cancer treatments[55].

An emerging approach that could constitute a new generation of cancer treatments exploits glycobiology, a research field that has been awarded with The Nobel Prize in Chemistry in 2022 (Carolyn R. Bertozzi). Glycoscience is the study of carbohydrates and their role in biological processes. Carbohydrates, also known as glycans, are complex sugars that play important roles in various biological processes, including cell recognition, signaling, and immune system function. In cancer, glycans can be found on the surface of cancer cells and can play a role in the development, progression, and metastasis of cancer. Researchers in the field of glycoscience are interested in understanding how glycans contribute to cancer and in developing new therapies that target glycans in cancer cells. This includes the use of drugs that target glycans on cancer cells or the use of antibodies that specifically bind to glycans on cancer cells. Other approaches being explored in the field of glycoscience and cancer include the use of enzymes that modify glycans[56], the use of small molecules that mimic glycans, and the use of vaccine therapies that stimulate the immune system to target glycans on cancer cells.

An example of an actionable pathway in this space is the Siglec-sialic acid immune pathway[57]. When this pathway is disrupted, it can lead to various immune disorders such as autoimmune diseases, neurological conditions, allergies, and cancer. The potential for Siglec-targeted therapy to be an effective treatment for cancer is largely dependent on a better understanding of Siglecs and their functions, as well as identifying ways to harness this knowledge to trigger an immune response against tumors. This can be a challenge due to the diversity of Siglec function and the variety of sialic acid synthesis and binding characteristics. There are several potential strategies for activating an immune response through modulation of the Siglec-sialic acid pathway, including blocking the immune-suppressive effects of inhibitory Siglecs, activating immune-stimulating Siglecs, and altering the synthesis and expression of the sialic acid glycocalyx. Antibody therapies that target Siglecs have the potential to improve immune responses by neutralizing immune-suppressive signaling from cancer cells that overexpress sialic acids. One such therapy, called NC318, targets Siglec-15, which is expressed on certain types of cancer cells and tumor-associated macrophages. It has been shown to block Siglec-15 function and halt tumor growth and metastasis in mouse models, and initial clinical data has shown tolerability. Antibodies that target Siglec-7 and -9 are also being studied for their potential to reduce tumor burden. Preclinical data in mouse models has shown that blocking Siglec-7 and -9 signaling can significantly reduce tumor growth. Several agents are currently being tested in preclinical studies using these approaches, and some therapies, such as NC318 (anti-Siglec-15, Nextcure) and E-602 (panSiglec, Palleon pharma), are being evaluated in clinical trials.

The development of a vaccine against Lyme borreliosis circles back to OspA

Sarai Araujo-Aris y Juan Anguita, *Inflammation and Macrophages Plasticity Laboratory*

Many infectious diseases lack effective vaccine formulations. This is particularly relevant for antimicrobial resistant (AMR) pathogens, for which the availability of chemotherapies is very limited or inexistent. However, this is not the case for Lyme borreliosis, a tick-borne infectious disease of high incidence in the Northern hemisphere. As Gary Wormser describes in his recent review[58], a safe and efficacious vaccine (LYMErix) was developed, tested and commercialized in late 1998. The vaccine was based on a non-adjuvanted lipidated outer surface protein A preparation from Borrelia burgdorferi *sensu stricto*, the main causative agent of Lyme borreliosis in the United States. The molecular basis for protection was recently published by Haque and collaborators using a panel of human-derived monoclonal antibodies[59]. Nevertheless, several issues arose with the vaccine, including the no inclusion of children below 15 years of age, its exclusive American market and the claim of potential induction of autoimmunity due to a weak homology with the antigen LFA-1. The latter could contribute to joint inflammatory symptoms, which nevertheless, were never substantiated in the subsequent analysis of the data from the clinical trial, which found no increased incidence of arthritis (Lyme or otherwise, autoimmune) in the vaccinated individuals, even in those that had previously suffered from Lyme borreliosis. To complicate matters even more, hesitancy among the targeted population increased due to the need for timely boosts in order to maintain high circulating antibody titers. This is due to the mechanism of action of the vaccine, that targets an antigen that is exclusively expressed in the tick vector and therefore, does not generate an anamnestic immune response. The vaccine was finally retired from the market due to poor sales in 2002. Another recent account on the development, clinical trials, commercialization and problems with LYMErix is covered in a review from Dattwyler and Gomes-Solecky3[60].

Since then, multitude of attempts have been made to find a suitable and efficient vaccine candidate. A revision of the antigen identity and engineering efforts has been recently

published by Chen and collaborators[61]. This year a new family of Borrelia proteins, called protein family 12 (PF12) has been tested and found to be potentially useful as part of multivalent vaccines against Lyme borreliosis[62]. Since the number of cases of Lyme borreliosis has more than doubled since the vaccine stopped being available, the search for new candidates and formulations has continued and expanded to other targets, including the ability of the ticks to feed efficiently and therefore, blunt the transmission of the bacterium to the mammalian host. This is relevant for Borrelia since transmission to the host takes around 40-48 h to occur from the time the tick attaches and starts feeding, although other microorganisms transmitted by ticks may not be affected. In this sense, it is highly relevant the recent use of mRNA vaccine technology by the group of Erol Fikrig to target 19 antigens from tick saliva with proved to be effective in reducing tick feeding and the transmission of Borrelia to the mammalian host[63]. Other efforts targeting tick feeding are being pursued, including those from the ANTIDotE consortium in Europe. Despite the setbacks, a recent survey[64] has shown that the willingness to receive a vaccine against Lyme borreliosis is high. This study recommends targeting hesitancy by effective communication from clinicians addressing safety and other vaccine parameters to those that are uncertain on whether to be vaccinated in high-risk areas. According to the CDC, a pre-exposure prophylactic (PrEP) human monoclonal antibody is expected to be tested in clinical trials soon[65]. Both the CDC and clinicaltrials.gov[66] also list an ongoing phase III clinical trial with a vaccine candidate, called VLA15, which is currently listed as enrolling 6,000 participants, 5 years of age and up, in the US (Connecticut, Maine, Maryland, Massachusetts, New Jersey, Pennsylvania, Rhode Island, Vermont and Wisconsin) and Europe (Finland, Germany, Netherlands, Poland and Sweden). VLA15 is being tested by Valneva in collaboration with Pfizer[67]. VLA15 is a multivalent protein subunit vaccine based on OspA that covers the 6 most common serotypes present in Borrelia burgdorferi *sensu lato*, and that are the cause of the vast majority of Lyme borreliosis cases in the US and Europe. The epitope that is weakly homologous to LFA-1 has been eliminated from the sequence in an unnecessary extra step. The results from the second phase of the clinical trial showed persistent antibody titers after six months of completing a 3- or 2-dose schedule in both children and adults, which shows potential as a preventative and perhaps definitive driver to slow down the progression of this infection.

Twenty years have passed since LYMErix was retired from the market. Even though many attempts have been made by numerous researchers over the years to find suitable alternatives, it is finally clear that the original formulation, albeit with required refinements, was an efficient means to prevent infection with Borrelia burgdorferi. The journey has circled us back to OspA. These encouraging results, plus other complementary/alternative approaches (anti-tick vaccines, monoclonal antibodies…) will may take us closer to an effective control of Lyme borreliosis, the most prevalent vector-borne infection in several parts of the world.

Chemical Tools for Synthetic Adjuvant and Vaccine Development

Alberto Fernández-Tejada, *Chemical Immunology Laboratory*

Current subunit vaccine approaches based on homogeneous, structurally defined antigens are less immunogenic than traditional whole-pathogen vaccines. As such, they require the use of an adjuvant, a substance that itself is not necessarily immunogenic, but increases the immunogenicity of the antigen, leading to more robust and durable immune responses. However, not many adjuvants show sufficient potency and acceptable toxicity for human use; in addition, their mechanisms of action are poorly understood[68]. Therefore, novel adjuvants and adjuvanting strategies are required to advance more effective subunit vaccines to the clinic, especially for those based on weakly immunogenic glycan/peptide antigens. The development of new adjuvant–antigen conjugates, in which both components are covalently linked in the same molecule, is emerging as a promising approach for rational vaccine design, enabling a safer and more precise targeting of the immune system that can lead to more effective vaccine induced immunity. However, despite decades of investigation and clinical research, mixed and not entirely successful results have been obtained, with no fully synthetic peptide- and/or carbohydrate-based vaccine being yet commercially available.

In this framework, the primary research program in the Chemical Immunology Laboratory has a dual far-reaching mission based on developing and exploiting new chemical approaches to tackle the above clear gaps in the adjuvant/vaccine field. Notably, important review articles and issues have been published in 2022 focused on adjuvants and glycoconjugates and their applications in the development of vaccines and immunotherapeutics[69],[70]. The past year has brought an important breakthrough in vaccine adjuvants based on saponin natural products with the FDA approval of the NuvaxovidTM Covid-19 vaccine under Emergency Use Authorization[71]. This milestone adds to the approval by the European Medicines Agency in 2021 of this protein-based vaccine, which is adjuvanted with the saponin-based nanoparticulate system Matrix-M[72]. In this regard, a book chapter has been published in 2022 describing the formulation of analogous, naturally-derived ISCOM-matrices nanoparticles for use as vaccine adjuvants[73]. Despite these advances, the development of novel adjuvants that are not extracted from natural sources but can instead be obtained via more practical, alternative routes is necessary. Of note, we have recently developed new synthetic saponin adjuvants using chemical strategies, providing potent lead compounds that induced superior immune responses than our previously reported QS-21 variants[74]. Additionally, other semisynthetic saponin-based adjuvants prepared from different plant species (Momordica saponins) have been developed via extensive structure-activity relationships studies[75] and some have been applied as part of a glycoconjugate pneumococcal vaccine formulation[76]. Furthermore, a recent study has provided important insights into the mechanism of action of saponin adjuvants, revealing that in addition to lipid body formation, activation of the PERK pathway is critical for saponin-induced dendritic cell cross-presentation and immune stimulation[77]. Finally, notable progress in synthetic self-adjuvanting vaccines has taken place during 2022, as exemplified by a novel tricomponent design that included the tumor-associated Tn antigen covalently linked to two adjuvant molecules (MPLA and a-GalCer). This fully synthetic vaccine construct elicited strong Tn-specific antibodies that were able to bind and kill cancer cells and to protect mice from tumor growth[78]. Our lab is also contributing to this field and has recently completed an important study on the development of self adjuvanting synthetic vaccines based on a streamlined saponin adjuvant and the tumor-associated mucin1 antigen [manuscript under revision]. Additional self-adjuvanting designs have been developed this year as potential Covid-19 vaccine candidates, which consisted of different adjuvant–peptide/protein conjugates that induced potent and protective immunity against SARSCoV-2[79],[80]. On a separate topic, we have recently written a review article focused on chemical biology strategies to probe the role of the b-O-N-acetylglucosamine (b-O-GlcNAc) glycosylation of proteins in immunity[81]. Notably, this is an area of great interest in the Chemical Immunology Lab and a research line that will be developed further in my group in the coming year.

Going forward, 2023 is expected to bring important advances in the field, with new self adjuvanting vaccine constructs that may have potential for clinical translation as well as novel molecular insights into the mechanisms of saponin immunopotentiation. Moreover, additional chemical tools will be developed to decipher the structural and functional consequences of the protein O-GlcNAc modification.

Multiscale Computational Modeling Strategies for Tissue Rejuvenation

Antonio del Sol and Sacha Jung, *Computational Biology Laboratory*

Aging is a multifactorial process that is characterized by a progressive physical and functional decline of tissues over time. As such, it is a major risk factor for debilitating conditions, including neurodegeneration, cardiovascular disease and metabolic syndrome. In order to counteract disease onset and increase healthspan, therapeutic strategies are required that delay or revert the aging process. However, to date, therapies are mainly aimed at treating aging-related symptoms after their onset. Although great efforts are being devoted to the identification of strategies for tissue rejuvenation, a number of challenges remain to make the tantalizing prospect of a substantially increased healthspan a reality in humans. In particular, many approaches only increase a limited number of tissue functions and thus fail to produce a holistic rejuvenation. Moreover, even though experimentation in animal models has led to the discovery of many therapeutic disease interventions, the aging process in humans is molecularly remarkably different from that of widely used model systems. Therefore, the vast majority of strategies for tissue rejuvenation show only limited efficacy in humans and cannot be translated to the clinics. Multiscale computational modeling can help addressing these issues by generating predictions to guide experimental efforts in the design of therapeutic interventions for tissue regeneration. Indeed, this modeling approach integrates different levels of biological organization, including intracellular gene regulation. cell-cell interactions and inter-organ communication.

During this year, the development of strategies for tissue rejuvenation gained pace and increasingly enabled the restoration of tissue function. In this regard, a computational cell-cell communication model characterized senescent cells as key contributors in the generation of an “aged-like” muscle niche, which is mediated by pro-inflammatory CD36 signaling cascades[82]. Indeed, elimination of these senescent cells, or reduction of their inflammatory secretome through CD36 neutralization, restores proper muscle functioning and regeneration in young and geriatric mice. In addition, restoration of a nourishing brain tissue niche has been shown to consolidate long-term memory and increased the proliferative capacity of oligodendrocyte precursor cells in aged mice[83]. Strikingly, the improvement of memory function solely relied on a single transcription factor, SRF, that can be activated by supplementing the niche with FGF17. Despite these success stories, the paradigm of counteracting the detrimental effect of the aged niche as a therapeutic strategy for tissue rejuvenation is hindered by the identification of suitable target signaling molecules. To address this issue, a computational model has recently been developed that enables the systematic identification of small molecules to specifically target defined sets of genes, including signaling molecules[84]. In the context of aging, it has been successfully predicted the inhibitors of NF-kB mediated inflammation, a hallmark of aging, and, more generally, is expected to enable the discovery of niche component mimetics to restore a youthful environment. In addition to the inter- and intracellular layer of regulation, genetic modification of the transcriptional regulatory network was successfully employed to revert molecular aging patterns as well as age-related biological processes in the pancreas, liver, spleen and blood[85]. In particular, the reverted molecular features included several metabolites, which suggests potential cross-tissue effects on non-modified cells.

In the following years, we expect an increasing effort towards the design of rejuvenation strategies that can be translated to the clinics. In the era of big data, multiscale computational modeling can greatly help in this endeavor by providing a platform for simulating the expected effects of interventions both in animal models and human tissues. Thus, it can gear experimental efforts towards the most promising interventions at the interorgan, intercellular and intracellular regulatory level. Based on the recent emergence of novel modeling frameworks, we expect to witness an expansion of computational methods that are expected to accelerate aging research and particularly the discovery of health span extending interventions significantly.

Mechanisms leading to aging and chronic associated liver disease

Malu Martinez-Chantar, *Liver Disease Laboratory*

The liver is the primary organ in charge of metabolic homeostasis and xenobiotic transformation in the organism. This organ has an extraordinary regenerative capacity to respond to acute insult or partial hepatectomy in order to maintain an appropriate organ-to-body weight ratio (PH). Maintaining hepatic homeostasis is critical for proper liver function, and this requires adequate nutrition with macro- and micronutrient intake. We have focused on magnesium in the Liver Disease lab due to its critical role in energetic metabolism[86] and metabolic and signalling pathways[87] that maintain liver function and physiology throughout its lifetime. For our lab, the effects of cation on liver disease have been studied[88],[89], with GalNAc siRNAs exerting a therapeutic and regenerative effects. In line with this, and focusing on mitochondria as essential biosynthetic, bioenergetic, and signalling organelles, we have discovered how improving bioenergetics affects hepatic metabolism[90] and the interplay between different cells that drive the liver regenerative process in aging and metabolic compromises preclinical animal models[91].

One of today's technological cornerstones is the resolution of complex intracellular organization through compartmentalization of metabolic processes into organelles in native tissue. Güneş Parlakgül et al. used enhanced Focused Ion Beam Scanning Electron Microscopy (FIB-SEM) imaging, followed by deep-learning-based automated image segmentation and 3D reconstruction have achieve a high resolution of the 3-dimensional organelle structural organization in large volumes of intact liver tissue. They also conducted a comparison of subcellular structures in liver tissue from lean and obese animals[92]. According to their findings, structural regulation is required for metabolic programming, and such a regulatory circuit could lead to new insights into endocrine and metabolic homeostasis, such as responses to hormonal or nutritional cues to determine metabolic outcomes. In this context, technology is one of the key drivers of biomedical innovation. It takes years for a precompetitive technological development to reach the level of maturity required to be risk-free enough for clinical application, industrial adoption, and upscaling. In the case of spatial metabolomics, the technology has now matured to the point where it can be used in many areas of biomedicine. Due to its well-defined spatial-functional organization, the liver is well suited to apply this technology in cancer as well as in the common NAFLD/NASH, providing a unique opportunity to achieve milestones in diagnosis, prognosis, and treatment[93]. Advanced

liver diseases have a very high morbidity and mortality rate, and liver transplantation is the only current therapeutic option. However, due to global donor shortages, a new alternative encouraging a translational perspective and implying future clinical potential has emerged.

The current scenario is moving toward a better understanding of the physiology of liver regeneration, stem cells, and 3D scaffolds for tissue engineering, which will hasten the race to effective therapies for liver failure. The successful development of whole bioengineered livers and their clinical application in place of liver transplantation.

Extracellular Vesicles as carrier of molecular information

Juan M. Falcón-Perez, *Exosomes Laboratory*

Extracellular Vesicles (EVs) including exosomes and microvesicles are intercellular communication entities that provide information about organs, tissues, and cells via liquid biopsy, and are an ideal diagnostic window into the human body. This hypothesis is backed by the fact that there are two FDA-approved EV-based diagnostic tests: Bio-Techne's ExoDx Prostate IntelliScore EPI-CE IVD Test for prostate cancer and Guardant's 360 CDx test for non-small cell lung cancer[94]. With almost 25.000 scientific publications EVs in the last 20 years and 7800 in 2022, and more than 100M€ privately invested to develop EV based diagnostics and therapeutics products, the EV market forecast is expected to reach $2.28 billions by 2030. Globally, there are 204 clinical trials on exosome studies, of which 114 trials are evaluating exosome-based therapeutics and 74 are testing exosome-based diagnostic tests. These numbers reflect the great expectations that academic and pharma industry researchers have in EVs, however there are many challenges that need to be solved before EV research impact significantly in the society.

In 2022, in collaboration with a cell-therapy-based group we have compared exosomes secreted by hair-follicle and adipocyte-derived mesenchymal stem cells and demonstrated their functional possibilities in therapy[95]. We have also collaborated in the characterization of EVs from oral progenitor cell line[96] and the glycan-surface profiling of EVs from mesenchymal stromal cells[97]. From a diagnostic point of view, we have identified a miRNA signature capable to inform about the best moment to perform the embryo implantation in the woman providing a tool to improve IVF[98]. Furthermore, our group has reported the characterization of EVs secreted by hepatocytes under obesity-resembling conditions and showed their metabolic impact on adipocytes[99]. In addition, also through collaborative efforts, the Exosomes Laboratory has participated in the characterization of the role of EVs in other pathological conditions including Thyroid cancer[100] and advanced hepatocellular carcinoma[101].

Apart of therapeutics and diagnostics an area of interest in the group is the study of the role that EVs play in Neurological diseases[102]. Particularly, in this area of EVs and neurodegeneration there have been reported several interesting articles. Huang et al. analysed the proteome of brain derived EVs from brain tissue of individuals with Alzheimer's disease (AD) (n=24) and controls (n=10). To this end, they digested the brain tissue samples with collagenase-3 and subjected them to differential centrifugation and filtration. After centrifugation at 10,000 x g, they kept the pellet (10k) for analysis and subjected the supernatant to size exclusion chromatography to isolate the EVs. Brain homogenate, 10k pellet and EVs were analysed by mass spectrometry. PCA of the EV proteome showed a separation of AD and control groups, while no such separation was observed for 10k or brain homogenate. This indicated that AD pathology most prominently affects the proteome of purified EVs in brain tissue. ELISA analysis showed that total tau, phosphorylated tau and peroxiredoxins 1 and 6 were significantly increased in brain derived EVs. ECL immunoassay analysis indicated that different cell types (e.g., astrocytes, endothelia, microglia, and neurons) contribute to the secretion of brain derived EVs and several cellular origin surface markers were increased in AD patients. These cell surface markers could be further tested in peripheral samples for brain derived EV capture[103]. Zhu et al. analyzed the mechanisms underlying TREM2-dependent modulation of tau pathology. Trem2 deletion enhanced transfer of human P301L tau from the medial entorhinal cortex to the hippocampal dentate gyrus region in mouse brain, which correlated with increased tau pathology, reduced synaptic transmission, and impaired spatial and fear memory. Using a microfluidic 3-chamber assay system, they showed that Trem2 KO microglia can transfer tau between isolated neuronal layers. Internalized tau trafficked into pre-exosomal compartments in Trem2 KO microglia which could be suppressed by treatment with an exosome inhibitor. Purified exosomes from Trem2 KO microglia exposed to tau oligomers displayed elevated levels of tau and showed enhanced tau-seeding competency in vitro. Furthermore, these exosomes induced enhanced pathological tau phosphorylation in WT mouse brain compared to exosomes from WT microglia. Together, these results implicate Trem2 as a suppressor of pathological tau dispersion and demonstrate that Trem2 deletion can aggravate pathological tau spreading through microglial extracellular exosomes[104].

Remarkably, blood-derived exosomes, a subclass of extracellular vesicles (EVs), have emerged as a peripheral biomarker source for Alzheimer's disease (AD). In a neuropathological context, beta-amiloid peptide (Aβ), Tau protein and phosphorylated Tau species (P-Tau 181), synaptic and lysosomal proteins, inflammatory mediators, growth factors, and specific microRNAs, among others, have presented distinct expression patterns in exosomes from AD cases when compared to Controls. Analyzing serum-derived exosomes from Control and AD cases with mass spectrometry (MS), Soares-Martins and collaborators found a set of 9 proteins which presented significantly different abundance levels between both groups of study: APOC3, APOH, C4BPα, CO3 and KV230 were significantly decreased in ADs; whereas AACT Isoform 1, CO9, IGHM Isoform 2 and K2C6A were increased in the exosomes of ADs. Therefore, this study hypothesized that AACT and C4BPα, two Aβ-binding proteins, could represent putative AD exosomal biomarker candidates, being those results also validated in individuals from independent cohorts using Western blot and ELISA approaches. Nonetheless, some inconsistencies have been reported in this article about these two potential exosomal diagnostic biomarkers: they might also represent early or late-stage biomarkers that alter with disease progression; they need to be further validated in a higher number of samples; and its levels also need to be evaluated in other neuropathologies to assess the potential in discriminating AD from other forms of dementia. These findings contribute for AD diagnosis due to the identification of novel blood-based exosomal biomarker candidates, which are needed in clinical practice as a more accessible form of diagnosis[105]. Exosomes constitute a way for propagation of pathological molecules. In this context chronic endoplasmic reticulum (ER) stress in skeletal muscle is one of the consequences of lipotoxicity, one of the pathogenic mechanisms in metabolic diseases. Unidentified signals from cells undergoing ER stress propagate paracrine and systemic unfolded protein response (UPR), which limits protein synthesis to prevent cytotoxicity. Inducing ER stress by palmitate, increases long-chain ceramide 40:1 and 42:1 secretion, regulated by the gene ceramide synthase 2 (CERS2) in a de novo synthesis pathway, regulated by the kinase Perk (involved in the UPR). Interestingly, since those ceramides are packaged into extracellular vesicles, they are secreted and induce UPR activation in naïve myotubes through dihydroceramide accumulation. Therefore, the transport

of ceramides by EVs in propagating cell non-autonomous UPR activation further highlights the important role of extracellular vesicles in skeletal muscle endocrine and paracrine signaling. This offers potential targets for therapeutic intervention in lipotoxicity-associated metabolic diseases[106].

The mechanisms of biogenesis and loading cargo of EVs is also another intense area of research in which our group has contributed[107] and where this year Garcia-Martin and colleagues did a significant contribution that has been published in Nature[108]. Different cell types secrete small Extracellular Vesicles (sEV) that carry different composition of miRNAs. Also, the composition of miRNAs in sEV is often different from their secreting cells. This implies a sorting mechanism inside the cells that may decide the uploading and enrichment of some miRNAs in sEVs. It has been seen that this mechanism could involve the presence of different sequences in the miRNAs that lead to a higher or lower uploading in sEV: CELLmotifs or EXOmotifs. Also involved are two RNA-binding proteins, Fus and Alyref, which export the miRNAs with the strongest EXOmotifs. This biased sorting is important because the miRNAs loaded by EXOmotifs that are delivered into distal recipient cells can inhibit target genes there, which can be a very different organ from the origin.

Ubiquitin-like Modifications in Health and Disease

Rosa Barrio, *Ubiquitin-likes and Development Laboratory*

We are interested on the regulation of developmental processes and diseases by post-translational modifications by the Ubiquitin-like (UbL) family. UbLs, like SUMO, are attached to target proteins altering their function, thus regulating cellular processes like proliferation and transcriptional regulation. Study of UbLs modification is challenging due to the small amounts of a given modified protein and the transient nature of the modification. We developed new proximity proteomics biotin-based strategies, including SUMO-ID, to identify modified proteins and interactors of proteins when modified, applicable to any UbL. We are developing strategies to identify E3 ligases modified targets in a subcellular manner. Importantly, targeted protein degradation (TPD) is more and more relevant in biomedicine. We focus on the Spalt-like (SALL) family of transcription factors, necessary for numerous biological processes. Mutations in SALL1 are associated to Townes-Brocks Syndrome (TBS), a rare disease-causing kidney defects, deafness and polydactyly. Mutant SALL1 interferes with the function of cilia, which opens new opportunities of intervention. We are investigating the role of the UbL family in the phenotypes associated to TBS.

TPD has undergone tremendous development in the last year, being adopted by major pharma companies. Many biotech companies are adding TPD technology to their portfolios. New licensing for clinical phase II estrogen receptor PROTAC from Arvinas-Pfizer, pushed the stakeholders and market. Among the latest advances is the development of BacPROTACs[109], which uses bacterial ClpC:ClpP (ClpCP) protease, present in gram-positive bacteria and in myco-bacteria, to direct degradation of target proteins in bacteria by means of bi-functional adaptors. These strategies might be used for testing new antibiotics and to fight bacterial infections. Many publications are devoted to the use of new E3 ligases into TPD. Till the moment, hijacking TPD molecules for only a handful E3 ligases out of approximately 600 is possible, i.e. CRBN, von Hippel–Lindau (VHL), murine double minute 2 (MDM2), inhibitors of apoptosis proteins (IAPs), among others. The search for new strategies to find approachable ligases is intense, as well as the study of the effect of degraders on E3 targets[110],[111],[112],[113],[114]. New molecules that bind CUL3 complexes were developed[115]. New approaches for preclinical target validations and comparing the different tag-TPD systems are also being developed[116]. Interestingly, XL01126 has been discovered, a VHL-based PROTAC able to degrade Leucine-rich repeat kinase 2 (LRRK2), one of the most promising targets for Parkinson's disease, which can be administrated orally and is able to cross the Blood−Brain Barrier[117]. This important development might open the door for new drug developments.

New cases of TBS have been reported, some of them with mutations in the SALL1 hot-spot region[118],[119], one case with mutations in the DACT1 gene, associated to TBS2[120].

It is expected that the TPD field continue growing. Given the limited types and number of E3 ligases used in TPD, expanding the E3 toolbox is of upmost importance. To characterize E3 ligases that are cell-type, biological process, organ and organelle-specific, will increase the specificity of the TPD applications. New strategies based on biotin proximity proteomics are very promising given their high specificity and sensitivity.

Three-dimensional cryo-electron microscopy of dynamic macromolecular complexes

Mikel Valle, *Cryo-EM of biologicals macromolecules Laboratory*

Structural Biology discipline has experienced an enormous transformation in recent years. First, the improvement in cryo-electron microscopy (cryoEM) has put a large number of biological structures resolved at atomic resolution available. More recently, the development of AI based structural prediction of protein structure has shaken the field, since accurate structures can be obtained just starting with the protein sequence. It is not clear whether structural predictions can reveal complexes between different proteins (or other biomolecules) or their dynamic behavior. Nowadays we are in a crossroad where predictions and experimental structural data can be combined.

In the last year, the prediction of protein structure has taken a large step forward by the creation of a database[121],[122] with over 200 million protein structure predictions by AlphaFold. These numbers are way beyond the previously known structures, and the database has the potential to replace the experimental work in most of the structural studies of protein complexes.

In our team we explore the structure and dynamics of large enzymes while their functioning, and this information is still out of reach for structural predictions. In the last year we have explored the ability of cryoEM to solve multi-pathway enzymatic reactions with pyruvate carboxylase as a model. Intensive classification techniques can be used in cryoEM to separate different steps of the reaction at the level of structural domains and at atomic resolution. This approach has been used by our group[123] and by others[124] revealing a complex reaction and the allosteric regulation that governs this enzymatic reaction.

Another interest in our team is related to the structure of amyloids of prion protein. In this field a large number of amyloid helical structures have been reported, including a breakthrough of high-throughput image processing of this type of samples by cryoEM[125]. This approach and the availability of a high resolution 300 kV electron microscope (Titan Krios G4) at the biophysics unit in the BREM, will push this line of research in the near future.

The following years will clarify the interplay between experimental and predictive Structural Biology approaches. Dynamic behavior of biological macromolecules, interaction between proteins, or between proteins and small ligands still require experimental approaches since these features cannot be resolved by AI based predictions. But is it changing in the near future? Do we need to integrate predictive methods in our

pipeline of experimental techniques? These and other aspects will shape our field profoundly in the incoming years.

Major advances regarding prion strain phenomena coincide with first-in-human treatment

Joaquín Castilla and Hasier Eraña, *Prion Research Laboratory*

2022 has brought some key advances in the field of transmissible spongiform encephalopathies (TSE) or prion diseases. This group of rare and invariably fatal neurodegenerative disorders includes Creutzfeldt-Jakob disease (CJD), Fatal Familial insomnia (FFI) and Gerstmann-Sträussler-Scheinker syndrome (GSS) in humans and scrapie in sheep and goats, bovine spongiform encephalopathy (BSE) in cattle and Chronic Wasting Disease (CWD) in deer. The main event underlying disease development is the misfolding of the cellular prion protein (PrPC) into an alternative pathogenic conformation called PrPSc. This aberrantly folded form, also known as prion, presents the capacity to induce its conformation on the cellular counterpart, forming fibrillary aggregates that accumulate in the brain and causes neuronal death in the process.

After achieving the resolution at an atomic level of the three-dimensional structure of the first brain-derived prion strain in 2021, a milestone in prion research, this 2022 has brought the resolution of other prion strains from different species, confirming their distinct biological properties respond to slightly different structural features[126],[127],[128]. Notably, two independent groups solved the structure of distinct murine prion strains, including also two versions of the same strain in a wild type model[129] and a GPI-less mouse model[130], lacking posttranslational modifications (PTM), and therefore showing that the three-dimensional structure is conserved regardless of these. In addition, the first human brain-derived prion structure was also solved this year, that of a form of GSS[131], as well as that from a recombinant human prion[132]. Likely, in the following years, the collection of prion strain structures will keep growing, getting as closer to the final goal of establishing an structure-activity relationship for these kind of amyloidogenic pathogens. Together with the structural evidences, another study was published clarifying the non-relevant role of prion protein PTM in prion strain characteristics[133], arguing for the usefulness of recombinant prions, lacking PTMs, to faithfully model prion diseases[134],[135]. Nonetheless, other prion disease models have been also of relevance during the past year, such as organotypic cultured brain slices[136],[137], cerebral organoids[138], which were key to discover new aspects of prion mediated neurotoxicity or screen for environmental prion-misfolding inducing factors among others. Finally, animal models of disease, still necessary to explore disease mechanisms need to be mentioned, which have been employed to understand glial involvement in prion diseases[139], to find new early disease events[140], shed light on the role of mitochondrial dysfunction in neurodegeneration[141], explore the consequences of prion protein deficiency[142] or model specific atypical forms of disease for which animal models were lacking[143]. Along with the advances in disease models that are contributing to the understanding of the molecular mechanisms of disease, in 2022, genome-wide studies or the use of –omics technologies have been of importance[144],[145146], as well as cutting edge technologies such as single-cell RNA sequencing[147]. As our knowledge on disease mechanisms increases, improvements in diagnostic are made and new therapeutic strategies designed. In terms of therapeutic approaches, the main highlight from 2022 is the publication of the first-in-human trial of an immunotherapy[148] which caused a modest inhibition of the neurotoxicity of prions[149]. Therapies based on the use of small inhibitory molecules have also kept arising[150],[151],[152] with a special focus on chaperones and disaggregases[153],[154]. In terms of advances in therapies for TSE, as prion protein silencing strategies and other prophylactic approaches are getting close to human trials, it is noteworthy that the target population is being set in the genetic forms of TSE as the optimal to evaluate efficacy, as shown by the observational assay being carried with FFI mutation carriers and doxycycline[155], several tools are being developed in anticipation, such as the analysis of non-human primate models for preclinical studies[156] and the determination of regional prion protein concentration variability[157]. Finally, other related approaches, such as the introduction of dominant negative forms of the protein by CRISPR/Cas9 are being explored[158], which will likely grow in importance in the near future. When it comes to diagnostics, the discovery of early biomarkers in plasma and CSF has been relevant during 2022, confirming the utility of some found previously, such as NfL or β-synuclein[159],[160],[161] or presenting new ones such as SNAP-25[162] and microRNA-146a[163]. These improvements in diagnostic have definitively prompted the report of cases with unusual features[164],[165],[166] or from the least frequent disease forms[167], since as rare diseases reporting the clinical characteristics of such cases is particularly relevant both, for diagnosis and treatment. In addition, given the future treatments will likely be evaluated in genetic forms of disease, controlling all potential carriers has gained relevance, as shown by recent publications on genetic counseling[168],[169], based on cohorts of mutation carriers[170] or studies trying to estimate carrier numbers[171]. Finally, regarding animal prion diseases, CWD from cervids is currently the main concern due to its spread in America and the description of new strains in Europe, being their transmissibility and zoonotic potential the most studied issues in 2022[172],[173],[174].

Altogether, the research field of prion diseases is facing a very exciting future, mainly due to the increasing collection of high-resolution prion fiber structures, which could soon provide the key to establish structure-activity relationship for these misfolded proteic pathogens, and the development of new and more sophisticated models, shedding light on molecular mechanisms completely unknown until now, such as the neurotoxicity mechanisms that will provide new therapeutic targets. In addition, improvements in diagnostic and the proximity of some treatments to human trials warrant an optimistic near future, in which gene therapy-based strategies may rapidly gain importance, as well as novel strategies that cannot be envisaged yet, but will definitely arise when the molecular mechanisms underlying these disorders are better understood.

Endosomal Trafficking. The Retromer Complex

Aitor Hierro, *Membrane Trafficking Laboratory*

The lab is devoted to the study of the molecular and structural mechanisms that govern intracellular transport via coated membrane vesicles. In particular how self-assembling cage scaffolds can generate unique geometries on and around membranes to sort cargo proteins into discrete trafficking pathways that are relevant in several neurodegenerative diseases such as Alzheimer's and Parkinson's disease. The group combines X-ray crystallography, cryo-electron microscopy (cryo-EM) and biochemical reconstitution to characterize the mechanisms by which coat proteins, adaptor proteins and other regulatory molecules determine the itinerary of a cargo protein within the cell.

The number of scientific papers published in the field of Retromer during 2022 (updated Dec 13th, 2022), according to PubMed, is 81. In particular some highlights regard a work[175] where the authors developed an adult-onset model for human Tau (hTau) toxicity in Drosophila to facilitate evaluation of factors that may contribute to Tau-dependent neurodegeneration. The

authors revealed that reduction of retromer activity induces a potent enhancement of hTau toxicity upon synapse loss, axon retraction and lifespan through a specific increase in the production of a C-terminal truncated isoform of hTau. Another work[176] reports a reduction in the retromer proteins in patients with amyotrophic lateral sclerosis (ALS) and in the ALS model provided by transgenic mice expressing the mutant superoxide dismutase-1 G93A. Restoring retromer expression aggravates the paralytic phenotype, whereas reducing retromer expression has the opposite beneficial effect. The study sheds light on a role for retromer in neurodegeneration of pathogenic and therapeutic importance. In another study[177], the authors determined the ACE2-PBM/SNX27-PDZ complex structure, and, through a series of functional analyses, they found that SNX27 plays an important role in regulating the homeostasis of ACE2 receptor. The authors demonstrated that SNX27, together with retromer complex (the core component of the endosomal protein sorting machinery), prevents ACE2/virus complex from entering lysosome/late endosome, resulting in decreased SARS-CoV-2 viral entry.

The process of tubular endosomal budding and trafficking is responsible for the subcellular localization of hundreds of cargo proteins such as signalling receptors, nutrient transporters, ion channels and adhesion molecules. Not surprisingly, numerous mutations affecting cargo recognition and recycling have been associated to a variety of diseases, most of which affect the nervous system. In the future, detailed description of how disease mutations compromise tubule-based endosomal sorting is certainly required. This knowledge is of pivotal importance for future therapeutic intervention as it might provide new targets for the development of scaffolding drugs that either promote or inhibit protein-protein interactions to interfere with specific trafficking pathway(s) without affecting others.

TECHNOLOGIES

Zooming into single-cell genomics

Ana M Aransay, *Genome Analysis Platform*

Over the past ten years, the development of high-throughput technologies to study DNA, RNA and DNA/RNA-protein interactions has revolutionized the ways in which cells and biological systems can be investigated. Within a relatively short time frame, research has changed from studying individual DNA sequence variants or single genes to analyzing entire genomes and/or transcriptomes. This breakthrough at breakneck speed makes genomic core-facilities being constantly learning and tuning up new techniques. This year, the Genome

Analysis Platform of CIC bioGUNE has been helping mainly with transcriptomic strategies, from mRNA to all non-coding RNAs (long-non-codingRNAs, smallRNAs, microRNAs, etc.), and has set-up specific RNA-protein interactions characterizations according to the needs of the particularprojects. These approaches have been carried out both, in bulk tissues and at single-cell level.

Tremendous efforts have been done to identify non-invasive biomarkers within different biofluids such as serum, plasma and urine, being microRNAs the most promising molecules for this aim. As a reflect of this search, a huge demand of smallRNA characterization services have been requested to our platform in 2022. Being aware of the potential technical biases and challenges intrinsic to this kind of projects, in the lab, we are comparing and trying to understand the different results obtained by using several commercial kits, that follow different strategies for the microRNA sequencing library preparation. Preliminary results point to inconsistent results among the outcomes, which is very valuable for us to advice researchers depending on their target, sample and/or aim. In addition, we believe that these divergencies should be considered when performing metanalysis of microRNAseq published data. As mentioned above we have studied the transcriptomic composition and mRNA levels in pools of single-cells, using 10x Chromium System's sc-mRNAseq protocols from 10x Genomics, in several live systems originated from human, mouse, rat, chicken and gecko species. In some of those projects we have run additionally to the single-cell_mRNAseq, bulk tissue mRNAseq, and other epigenetic characterization such as ATACseq (Assay for Transposase-Accessible Chromatin using sequencing), DNA methylation status and microRNAseq. In the coming months, we have the challenge of integrating all these data and interpreting the biological interactions that these stories tell us about (Integrative Biology).

Furthermore, we continue to bet on the adjustment of processes (i) for multiome characterization at single-cell level, which makes possible to study in parallel the mRNA and the genome-wide chromatin accessibility (ATAC) in each single-cell of the studied systems, and (ii) for spatial transcriptomics, being very interested in setting-up this protocol for FFPE-tissues, for which we will try a new kit recently offered by 10x Genomics company for this purpose. All these techniques and new set-ups will allow us to successfully collaborate in several system biology and personalized medicine projects with CIC bioGUNE research groups, CIBERehd collaborators, as well as in a European project related to biomarkers discovery. Moreover, the platform keeps on carrying out several successful collaborations that have yielded remarkable scientific reports in genomic studies[178], [179], [180].

Latest advances in Mass Spectrometry for Proteomics

Felix Elortza, *Proteomics Platform*

The complexity of the proteome forces the proteomic community to continuously improve the analytical methods employed, to face in better position the enormous task of deciphering the molecular events that underlie behind diseases. Recently several advances in mass spectrometry related area have been presented that shall help getting further in biomedical research.

In proteomics, one of the major aims is to compare proteomes from samples of interest such as healthy vs diseased tissue, biofluid etc. The final objective is to identify as much proteins as possible, detect which proteins are differentially expressed and to quantify these differences. Mass spectrometry (MS) is still one of the most popular methods for this purpose. Since liquid chromatography was coupled on-line by electro-spray to a tandem mass spectrometer to perform the so called "shot-gun" proteomics, data dependent acquisition (DDA) has been the most popular method of acquisition. In this method, peptides resulting from a digestion of a complex protein matrix are resolved according to hydrophobicity in a chromatographer. The eluting peptides are first scanned at precursor level (MS); and in a second step they are selected for tandem mass spectrometry (MS/MS) in a signal intensity dependent manner, being prioritized according to their intensity. Nevertheless, a new method of data acquisition is gaining momentum lately: the data independent acquisition (DIA). With this new method, all peptides detected during the first MS cycle can be fragmented in the second round. This acquisition occurs in parallel across peptides and resulting spectra are highly multiplexed eventually keeping all information within the acquired data. Among the pros, developers of this method claim that overall offers higher precision and maintain that it will help quantifying proteins in complex mixtures over a large dynamic range, thereby overcoming the challenge of under sampling when using DDA. At CIC bioGUNE´s proteomics platform we have been working to implement the DIA-NN method within our pipeline[181]. By using the TIMS Tof Pro (Bruker), the dia-PASEF technology uses ion mobility separation to reduce signal interferences and increase sensitivity in proteomic experiments. The novel approach is further optimized by employing deep neural networks for the searches.

Another major challenge in proteomic analysis is the study of the post-translational modifications (PTMs). PTMs are covalent modifications of proteins that can range from small chemical modifications to addition of larger molecules such as glycans, and even entire proteins (e.g., ubiquitylation, SUMOylation, Neddylation etc.). PTMs contribute to regulation of protein function, location, association etc. and thereby greatly increase the functional diversity of the proteome. Saying that, the analysis of PTMs is extremely challenging due to their chemical diversity. Furthermore, PTMs alter the behaviour of a given peptide in chromatography and mass spectrometry analysis. Thus, technical advances that shall help in getting more information about PTMs are always welcome. In this regard we can mention the Synchro-PASEF approach, which allows precursor-specific fragment ion extraction and interference removal in DIA acquisition[182]. The obtained additional confidence in fragment identity can be useful for studying post-translational modifications, where each additional fragment can be crucial in pinpointing the modification site. In this context, another step forward is the omnitrap platform: a versatile segmented linear ion trap for multidimensional multiple-stage tandem mass spectrometry[183]. This platform can be used together with orbitrap or time-of-flight analyzers and shall provide extra information for PTM analysis of proteins. As it could be expected, there is an increasing need for the development of bioinformatic tools that improve glycoprotein/glycopeptide mass spectrometry obtained complex data analysis. As an example, some generic analytical pipelines and useful software became lately available making possible the precise analysis of intact glycopeptides in a robust manner[184].

Altogether, we can affirm that during last years, we are witnessing an accelerated optimization of mass spectrometry related techniques and methodologies that shall help paving new avenues to decipher the enormous complexity of the human proteome.

Cell and Plant Metabolomics

Juan Manuel Falcón and Sebastiaan van Liempd, *Metabolomic Platform*

We supported numerous research projects in their metabolite analysis by high resolution liquid chromatography coupled with high resolution mass spectrometry (hrLCMS) to target the metabolites of interest. Our core assays are focused

on metabolic members of the methionine cycle, polyamide synthesis, transsulfuration pathway and the tricarboxylic acid (TCA, Krebs) cycle. During this year we have incorporated the steroid hormones into the possibilities of the platform[185]. We have also started to approach Plant metabolomics by analyzing how ammonium nutrition interacts with iron homeostasis in Brachypodium distachyon as experimental model for improving resources sustainability of the society[186]. We have also implemented stable-labeled flux analysis methods to our portfolio. The essence of these methods is the tracking of stable labelled atoms from a precursor metabolite throughout a metabolic pathway. In our platform we use 13C5-methionine and 13CD3-methionine to probe the methionine cycle while we determine fluxes trough the TCA cycle with 13C6-glucose and 13C5-glutamine. However, if a labeled precursor is available, we can track most pathways with our hrLCMS setup. The great advantage of these stable labelled experiments is the absence of background signals of endogenous metabolites. We are currently setting-up assays focused on drug metabolism. Especially cytochrome P450 inhibition, reactive intermediate trapping and liver toxicity have our interest. These assays are part of the pre-clinical drug discovery pipeline and thus of great importance to the pharmaceutical industry.

From the in vitro to the *in situ* macromolecules structure

Isaac Santos and **Adriana Rojas,** *Electron Microscopy Platform*

The technological advances in cryo-electron microscopy (Cryo-EM) have driven an exponential growth in the number of molecular structures solved. However, more than a snapshot is needed to understand a biological process as a whole. Knowing all the substrates, reaction intermediates, and the cellular environment is necessary. Only this will allow us to conclude the mechanism and function played by the molecule under study. Therefore, although the number and quality of the particles to be analyzed are no longer a limitation, the heterogeneity of some molecules and the preparation of samples are still a limitation. In this sense, greater success has been demonstrated in obtaining orientations and intermediate states of target molecules thanks to the use of the inkjet mechanism. This technique uses picolitre-sized sample solution droplets sprayed on the grid, capturing precise reaction moments. Another improvement has been reducing dwell time (spot-to-plunge time) to a few msec, thus minimizing adsorbed air-water interface (AWI) particles in fine ice[187].

Today, the Cryo-electron tomography (Cryo-ET) improvements have allowed ambitious challenges as addressed protein complexes closer to the native cellular environment. Instead of the usual "in vitro" purification, Cryo-ET has become the method of choice for high-resolution 3D structural studies of complexes in cells[188]. This technique, combined with cryo-focused-ion-beam (cryo-FIB milling), is one of the most promising methods to know molecular structures in high resolution in their natural state. Unfortunately, the commonly used gallium beam tends to contaminate the generated lamellae, which also undergo ice contamination during transfer to the TEM microscope[189]. For this reason, new strategies are being developed to generate less damaged lamella and more preserved structures. The use of gases such as oxygen and argon can generate plasma, which, used as a beam, generates much larger and cleaner sections. This new technique is called plasma cryo-focused ion beams (cryo-PFIB). Using this new beam, the structure of the human 80S ribosome was solved to a resolution of ~4.9 Å, and the well-ordered regions close to 3.8 Å[190].

The appearance of new tools can disrupt established workflows and generate new ones. This could be the case with scanning transmission electron microscopy with integrated differential phase contrast (iDPC-STEM). (https://doi.org/10.1038/s41592-022-01586-0). So far, the determination of protein structures at 3.5 Å resolution has been achieved. Still, developers assure that, with further updates, it will be possible to achieve resolutions like those obtained today by cryo-EM.

Regarding the biomedical field, cryo-EM continues shading light on our knowledge about cancer. In the last year, key proteins in prostate and breast cancer, or leukemia, have been solved at the atomic level. In this sense, remarkable is the structure of the 11KD and flexible KIX domain of CREB-binding protein (CBP), a potential therapeutic target for acute myeloid leukemia solved at 2.6 Å resolution. This was possible thanks to the development of an ingenious device in which the protein of interest was trapped in a double-shell sandwich formed by apoferritin as the inner shell and maltose-binding protein as the outer shell[191].

Furthermore, the impact caused by the high fidelity of machine learning in predict the molecules 3D structure has also changed the mentality of researchers, since they now have a new tool on which to work and verify their hypotheses[192].

Nuclear Magnetic Resonance Methodological Advances

Tammo Diercks, *NMR Platform*

The CiC bioGUNE NMR facility supports research in a very broad range of research applications, demanding a very broad scan of literature such that only selected topics and important novelties can be listed below:

NMR methodology: This wide category comprises all enabling tools that facilitate novel or enhanced NMR experiments in a general way. For instance, gradient based (tomographic) slice selection enables broadband homonuclear decoupling to collapse multiplet fine structures. Recently, this tool was implemented in the J-resolved experiment to suppress the characteristic phase twist of its correlation signals[193]. Appropriate sets of NMR experiments can be concatenated without the lengthy interscan recovery delay by the “No Relaxation Delay (NORD)” concept to save important total experiment time[194,195]. A new set of adiabatic inversion pulses with high band-selectivity was derived analytically[196]. An isotope filter-editing scheme (Ac-FIND) was presented to monitor N-acetylation of proteins even in live cells[197]. An approach to achieve efficient TOCSY mixing with far reduced RF power requirements, called HORRENDOUS NMR, was transferred from solid to solution state NMR[198]. CA band-selective 13C,15N heteronuclear TOCSY was used to simultaneously exploit and transfer both 13CA and 13N polarisations between amino acids in intrinsically disordered protein[199]. The RODA spin state selection module[200], is just a new name for our previously published qTROSY scheme.

Isotope labelling: Strategies to include NMR active isotopes via isotope exchange (e.g., 13C, 15N, 2H) or chemical modification (e.g., 31P, 19F – discussed separately) are fundamental for advanced NMR studies with high sensitivity, selectivity, and information content. New methods for 17O labelling in amino acids were presented[201] 13C,15N labelled phospholipids for specific lipid studies in complex membranes were obtained from yeast grown in enriched medium[202]. Cost efficient isotope labelling in mammalian cells was accomplished under glutamine-free conditions by co-expressing glutamine synthase[203].

Protein structure analysis: Advances here rely on measuring more, and more exact, conventional NMR restraints. Thus, sampling of initial NOE build-up rates (“exact = eNOEs”)

may yield distance restraints with up to 0.1 Å precision, which allowed to resolve the dual states of an allosteric protein[204],[205] using PDBcor.

Molecular interaction analysis: Solution state NMR is uniquely suited to elucidate molecular interactions rapidly, with atomic resolution, and down to the lowest affinities (mM range). This continuous to be exploited for elucidating the molecular mechanisms of SARS-CoV-2 infection, often by STD NMR for epitope mapping[206],[207], [208],[209]. Chemical shift perturbation (CSP) in protein spectra titrated with a ligand readily reveal affinity constants, Kd, where a recent study showed that residue specific deviations in Kd may occur and correlate with ligand proximity suggesting a rigid contact surface[210].

Chemical/conformational exchange studies: Selective isolated labelling with 13C allows to measure exact 1H,13C cross-correlated relaxation rates to identify and quantify exchange line broadening (Rexch) from slow conformational motions, e.g. in RNA[211],[212]. The thermodynamics and kinetics of chemical exchange can be derived from NMR lineshape analysis, for which an exhaustive, detailed report was presented[213]. Similarly, residue specific thermodynamic and kinetic data of a polypeptide in slow two-state exchange were derived by conventional EXSY and by time-zero HSQC using a novel Π analysis[214].

Computational NMR data analysis: Novel algorithms and methods of artificial intelligence (AI) continue to revolutionise NMR data analysis with great potential. The fundamentals and aspects of deep neural networks for spectral peak picking were discussed in a critical review[215]. These offer important benefits for quantitative metabolite recognition in 1H NMR, as by the COLMARq web server[216] distance correlation, signal shifts in 2D spectra induced by ligand binding were compacted and quantified in a single parameter[217]. The combination of NMR experimental data with molecular dynamics (MD) simulations continues to yield atom and time resolved information on conformational space, as in a recent study on monosaccharides[218]. The CRENSO algorithm was used to compute conformation specific 13C chemical shifts of macrocyclic molecules, allowing to deduce a conformer ensemble consistent with NMR data[219].

NUS processing: Non-uniform NMR data sampling (in the time domain) has by now, and duly, become standard to massively increase the resolution of multidimensional NMR experiments within a given (short) sampling time. The critical aspect lies less with the NMR data sampling itself, but rather with reliable, artefact-free reconstruction of the incompletely sampled spectrum. High-quality NUS data reconstruction by the low rank Hankel matrix (LRHM) approach was greatly accelerated by using a sliding window together with parallel computation[220].

DNP: Dynamic Nuclear Polarisation is a (experimentally still demanding and limited) method to enhance NMR sensitivity by up to 3 magnitudes; however, the conferred nuclear spin hyperpolarisation holds for just a single scan. A detailed protocol to boost NMR sensitivity for biomolecules harbouring labile protons by virtue of their exchange with hyperpolarised H2O[221]. In the absence of exchanging protons, solution-state DNP can typically enhance only heteroatoms, where a new study shows that 13C DNP can be passed via 1JCH coupling to bound 1H for enhanced detection[222]. A novel mechanism akin to CIDNP was presented, called J-driven DNP, to increase the efficiency of DNP in high-field solution state NMR via J-coupling (not dipolar coupling) mediated polarisation transfer[223].

Paramagnetic NMR (PRE, PCS): Paramagnetic agents may not only enhance polarisation, as exploited by DNP, but produce structurally and dynamically important NMR restraints such as paramagnetic relaxation enhancement (PRE) or pseudocontact shifts (PCS). Thus, insertion of a lanthanide-binding peptide sequence into a loop of an extracellular receptor glycoprotein (hRobo1) allowed to measure long-range PCS on 13C labelled methyl groups to gauge conformational changes in the bound and free receptor[224]. Similarly, a new rigid lanthanide-binding aptamer allows to derive PCS restraints for nucleic acids[225].

19F NMR: The absence of natural background signals in combination with excellent NMR sensitivity and dispersion makes fluorine (100% 19F) a unique NMR probe introduced via chemical labelling. Thus, specific in cell studies by 19F NMR are facilitated by the use of novel fluorinated agents[226]. Such in cell studies may also rely on measuring 19F PRE on proteins including fluorinated amino acids[227],[228]. In another application, 19F NMR was used to study the conformation-activity relationship and logp lipophilicity for fluorinated amides[229],[230]. Fluorotryptophan introduction into galectins also allowed highly specific studies of their glycan recognition by 19F NMR (Lete et al. 2022,). The new 4'-SCF3 label facilitates interaction studies in the minor groove of DNA[231],

In cell NMR: Besides 19F, in cell NMR can also exploit protein overexpression for 13C or 15N isotope enrichment above the limit for specific detection. This can be used to elucidate protein structures inside living cells[232] or to study their molecular interactions, as shown for a multi-state protein by 15N SOFAST-HMQC[233]. 31P NMR was similarly used for in cell quantification of phosphohistidine[234].

Diffusion NMR: The biasing coupling between rotational and translational diffusion in case of large, anisotropic molecules was analysed in detail to refine, e.g., aggregation studies of protein fibrils[235]. By applying the NUS concept to both indirect evolution time and diffusion encoding gradient strength, time-resolved Tr-DOSY spectra can be recorded, e.g., to monitor photopolymerisation reactions[236]. An ultrafast new DOSY experiment allows to monitor reactions under flow conditions[237],[238] .

In the near future, NMR methodological advances may be scarce given the increasing maturity of this most versatile analytical technique. Yet, surprise breakthroughs are always possible and most relevant due to their wide, enabling potential. The most dynamically developing sub-category in NMR, however, appears to now be on the side of computational advances in NMR data processing and analysis where, in particular, algorithms of artificial intelligence / machine learning can be expected to find more, broader, and most valuable use in NMR.

Front Page Image:
Naroa Goikoetxea

Editing:
Donatello Castellana

Middle Page Image:
https://www.freepik.es/vectores/tecnologia'>Vector de Tecnología creado por starline - www.freepik.es

REFERENCES

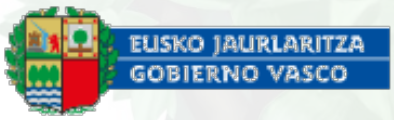
1 Doi: 10.1002/anie.202201432
2 Doi: 10.1002/chem.202202614
3 Doi: 10.1101/2022.09.14.507904
4 Doi: 10.1021/jacs.1c10329
5 Doi: 10.1021/acsinfecdis.2c00046
6 PMID: 34954141
7 PMID: 36284108
8 PMID: 34931806
9 PMID: 36525649
10 PMID: 35682991
11 PMID: 35017217
12 PMID: 35148986
13 PMID: 35087074
14 PMID: 35605245
15 PMID: 35075112
16 PMID: 35283189
17 PMID: 35111385
18 PMID: 35105978
19 doi.org/10.1038/s44160-022-00171-9
20 PMID: 36313169
21 PMID: 36280670
22 doi: 10.1038/s41586-021-03819-2
23 Doi: 10.1038/s41586-021-03828-1
24 doi: 10.1038/s41592-021-01362-6
25 doi: 10.1093/nar/gkab1061
26 doi: 10.1101/2022.07.20.500902
27 doi: 10.1038/s41592-022-01685-y
28 doi: 10.1038/s41592-022-01645-6
29 doi: 10.1101/2021.11.08.467664
30 doi: 10.1038/s41467-022-33729-4
31 doi:10.1038/s41586-022-05402-9
32 doi:10.1016/j.eururo.2021.12.039
33 doi:10.1038/s41586-022-04875-y
34 doi:10.1038/s41586-022-05435-0
35 doi:10.1016/j.immuni.2022.07.016
36 PMID: 36513066
37 PMID: 35175798
38 PMID: 36175792
39 PMID: 35175605
40 doi: 10.1200/JCO.2022.40.17_suppl.LBA501
41 PMID: 36077734
42 PMID: 36156078
43 PMID: 35710981
44 PMID: 35227606
45 PMID: 36347255
46 PMID: 35948782
47 PMID: 34957435
48 PMID: 34741163
49 PMID: 34186061
50 PMID:34615726
51 PMID: 30658996
52 PMID: 34957435
53 PMID: 34642484
54 PMID: 35798503
55 PMID: 35764745
56 PMID: 36264237
57 PMID: 36322632
58 PMID: 34763193
59 PMID: 36350351
60 PMID: 35087055
61 PMID: 35836340
62 PMID: 36102656
63 PMID: 34788080
64 PMID: 34895785
65 https://www.cdc.gov/lyme/prev/vaccine.html
66 clinicaltrials.gov: Code Trial NCT05477524
67 https://www.businesswire.com/news/home/20220808005463/en/
68 doi: 10.1007/978-1-0716-1892-9_8
69 doi: 10.1021/acs.chemrev.1c01032
70 doi: 10.3389/fimmu.2022.941474
71 https://www.fda.gov/media/159902/download
72 https://ir.novavax.com/2021-12-20-European-Commission-Grants-Conditional-Marketing-Authorization-for-Novavax-COVID-19-Vaccine
73 doi: 10.1007/978-1-0716-2185-1_15
74 doi: 10.3389/fimmu.2022.865507
75 doi: 10.1021/acs.jmedchem.2c01087
76 doi: 10.3389/fimmu.2022.1079047
77 doi: 10.1007/s00018-022-04253-x
78 doi: 10.1016/j.apsb.2022.05.028
79 doi: 10.1021/acs.jmedchem.1c02000
80 doi: 10.1038/s41421-021-00370-2
81 doi: 10.3389/fimmu.2022.1089824
82 DOI: 10.1038/s41586-022-05535-x
83 DOI: 10.1038/s41586-022-04722-0
84 DOI: 10.1093/nar/gkac862
85 DOI: 10.1111/acel.13578
86 PMID: 33513920
87 PMID: 25540137
88 PMID: 33571553
89 PMID: 36433951
90 PMID: 35921199
91 PMID: 34510498
92 PMID: 35264794
93 PMID: 34608315
94 cancer.gov
95 PMID: 35395929
96 PMID: 35716044
97 PMID: 36076936
98 PMID: 36029522
99 https://doi.org/10.1002/jex2.32
100 PMID: 35849996
101 PMID: 36078082
102 doi: 10.1016/j.semcancer.2022.11.004
103 PMID: 36213994
104 PMID: 36056435
105 doi: 10.1007/s12035-022-02762-1
106 PMID: 35365625
107 PMID: 24921245
108 PMID: 34937935
109 PMID: 35662409
110 PMID: 35857806
111 PMID: 36430693
112 PMID: 36414666
113 PMID: 36329119
114 PMID: 35013300
115 PMID: 36328355
116 PMID: 35440107
117 PMID: 36007011
118 PMID: 35446975
119 PMID: 36252910
120 PMID: 36066768
121 PMID: 34791371,
122 https://alphafold.ebi.ac.uk/
123 PMID: 36261450
124 Doi: 10.1016/j.molcel.2022.09.033
125 Doi: 10.1039/D2FD00034B
126 doi: 10.1007/s00441-022-03676-z
127 doi: 10.1371/journal.ppat.1010594
128 doi: 10.1371/journal.ppat.1010947
129 doi: 10.1038/s41467-022-30457-7
130 doi: 10.1038/s41467-022-30458-6
131 doi: 10.1007/s00401-022-02461-0
132 doi: 10.1038/s41594-022-00833-4
133 doi: 10.1371/journal.ppat.1010900
134 doi: 10.3390/v14091940
135 doi: 10.3390/biom12111694
136 doi: 10.1038/s41594-022-00814-7
137 doi: 10.1111/bpa.13130
138 doi: 10.1371/journal.pone.0277051
139 doi: 10.1002/glia.24244
140 doi: 10.1016/j.nbd.2022.105834
141 doi: 10.1371/journal.pone.0267720
142 doi: 10.1186/s12915-021-01203-0

143 doi: 10.1186/s40478-022-01477-7
144 doi: 10.3389/fvets.2022.824677
145 doi: 10.1007/s00401-022-02481-w
146 doi: 10.15252/embj.2022112338
147 doi: 10.1186/s40478-022-01450-4
148 doi: 10.1016/S1474-4422(22)00082-5
149 doi: 10.1038/s41598-022-13455-z
150 doi: 10.1038/s41598-022-12049-z
151 doi: 10.3390/molecules27227935
152 doi: 10.1016/j.ejmech.2022.114923
153 doi: 10.3389/fnins.2022.902600
154 doi: 10.1093/brain/awac144
155 doi: 10.1080/19336896.2022.2083435
156 doi: 10.1371/journal.ppat.1010728
157 doi: 10.1172/jci.insight.156532
158 doi: 10.1371/journal.pone.0269342
159 doi: 10.1099/jgv.0.001764
160 doi: 10.3390/ijms23137182
161 doi: 10.1007/s12035-022-02891-7
162 doi: 10.1136/jnnp-2021-328646
163 doi: 10.5114/fn.2022.113561
164 doi: 10.1007/s00415-022-11467-3
165 doi: 10.1016/j.clinph.2022.08.025
166 doi: 10.1186/s40478-022-01415-7
167 doi: 10.1007/s00415-022-11051-9
168 doi: 10.1016/j.gim.2022.06.003
169 doi: 10.3233/ADR-210067
170 doi: 10.1093/jnen/nlac113
171 doi: 10.1038/s41431-022-01132-8
172 doi: 10.1007/s00401-022-02482-9
173 doi: 10.3390/v14071578
174 doi: 10.1093/infdis/jiab033
175 doi: 10.1038/s41467-022-32683-5
176 doi: 10.1073/pnas.2118755119
177 doi: 10.1073/pnas.2117576119
178 PMID: 36029522
179 PMID: 35705557
180 PMID: 35430283
181 doi: 10.1101/2021.03.08.434385
182 doi: 10.1101/2022.11.01.514654
183 doi: 10.1021/jasms.2c00214
184 doi: 10.1074/mcp.R120.002090
185 doi: 10.3390/metabo12080714
186 doi: 10.1093/jxb/erab427)
187 doi: 10.3389/fmolb.2022.892459
188 doi: 10.1016/j.sbi.2022.102484
189 doi: 10.1038/s41596-020-0320-x
190 doi: 10.1101/2022.08.01.502333
191 doi: 10.1021/acscentsci.1c01090
192 doi: 10.1093/bib/bbac308
193 doi: 10.1016/j.jmr.2022.107267
194 doi: 10.1016/j.jmr.2022.107297
195 doi: 10.1039/d1cc06663c
196 doi: 10.1016/j.jmr.2022.107298
197 doi: 10.1016/j.jmr.2022.107169
198 doi: 10.1016/j.jmr.2022.107176
199 doi: 10.1016/j.jmr.2022.107166
200 doi: 10.1016/j.jmr.2022.107177
201 doi: 10.1002/chem.202203014).
202 doi: 10.1002/chem.202202472
203 doi: 10.1007/s10858-021-00387-5
204 doi: 10.1007/s10858-022-00392-2
205 doi: 10.1038/s41467-022-33687-x
206 doi: 10.1021/jacs.2c05603
207 doi: 10.1002/chem.202202614
208 doi: 10.1002/chem.202202599
209 doi: 10.1002/anie.202201432
210 doi: 10.1007/s10858-022-00402-3).
211 doi: 10.1016/j.jmr.2022.107245
212 doi: 10.1021/acs.chemrev.1c00845
213 doi: 10.1038/s41598-022-20136-4
214 doi: 10.1007/s10858-022-00396-y
215 doi: 10.1007/s10858-022-00393-1
216 doi: 10.1021/acs.analchem.2c00891
217 doi: 10.1016/j.jmr.2022.107165
218 doi: 10.1002/chem.202201544
219 doi: 10.1002/anie.202113905
220 doi: 10.1016/j.jmr.2022.107283
221 doi: 10.1038/s41596-022-00693-8
222 doi: 10.1016/j.jmr.2022.107304
223 doi: 10.1039/d1cp04186j
224 doi: 10.1038/s41598-022-18769-6
225 doi: 10.1002/chem.202202114
226 doi: 10.1002/anie.202213495
227 doi: 10.1016/j.str.2021.09.009
228 doi: 10.1002/anie.202201097
229 doi: 10.1002/chem.202202208
230 doi: 10.1002/anie.202114862
231 doi: 10.1002/anie.202201848
232 doi: 10.1038/s42003-022-04251-6
233 doi: 10.1002/anie.202213976
234 doi: 10.1371/journal.pone.0273797
235 doi: 10.1016/j.jmr.2022.107303
236 doi: 10.1021/jacs.2c05944
237 doi: 10.1002/chem.202201175
238 doi: 10.1016/j.pnmrs.2022.01.002

CIC bioGUNE
MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

www.cicbiogune.es

Funding Agencies:



Fundación Jesús Gangoiti

BRTA
BASQUE RESEARCH & TECHNOLOGY ALLIANCE