

<http://proteoaix2023.sciencesconf.org>

 @2023ProteoAix

**MAP**  
MARSEILLE PROTEOMIQUE



# ProteoAix

en Provence

3rd joint Meeting of Spanish,  
Portuguese and French  
Proteomics Societies

20-23 June 2023

# Abstract Book



# Summary

- Welcome
- Organizing and Local Committees
- Scientific Committee
- Sponsors
- Exhibition map
- Scientific Programme
- Symposium abstracts
  - FPS thesis prize
  - Plenary Conferences
  - Oral presentations
  - Flash Presentations
  - Posters communications





# Welcome

On behalf of the organizing committee, it is a pleasure to welcome you to Aix-en-Provence or Ais de Prouvènço to attend ProteoAix 2023. The town was founded in 122 BC by the Roman Republic Consul Caius Sextius Calvinus under the name of Aquae Sextiae (the water of Sextius). Aix-en-Provence subsequently became the capital of Provence. This explains why this town has an important cultural heritage. It is still known as the city of a hundred fountains. You will have the opportunity to admire this city, home to the famous Post-Impressionist painter Paul Cezanne.

This congress will be the third joint meeting of Spanish, Portuguese and French Proteomics Societies. The meeting was organized to cover all aspects of proteomics. For this, we have invited 8 plenary speakers who will share their work and experiences, bringing the opportunity to discuss and share opinions. We will also have around 30 oral communications, 14 flash communications and more than 60 posters. In addition, there will be numerous oral communications from sponsors and 11 booths.

We are very happy with the high number of registrations (more than 160) and we would like to thank you for your participation. The scientific level of the communications received was very high, and this complicated the task of selection by the scientific committee.

We would also like to thank all of the invited speakers and the support of all sponsors and partners. This conference would not be possible without them.

Now we wish you to enjoy ProteoAix and make this conference a success.

We look forward to welcoming you to Aix-en-Provence.

**Luc Camoin**

Chair ProteoAix Organizing committee



### Welcome address by the President of the Spanish Proteomics Society (SEProt)

From the Spanish Society of Proteomics (SEProt) we want to welcome you to the third edition of the Joint Meeting with the proteomics societies of France and Portugal that is hosted this time in the beautiful setting of Aix-en-Provence.

We're really excited about the opportunity this conference presents. It's a great chance for us to share ideas, learn from one another, and continue pushing the boundaries of what we know about proteomics. It's through our combined efforts that the field continues to evolve and grow.

Big thanks to our guest speakers. We appreciate you taking the time to come and share your knowledge with us. We can't wait to learn from you. We would also like to acknowledge the organizers for all their hard work in putting this conference together. And of course, let's not forget everyone who's made it to the conference. Your work, your curiosity, your ideas - these are what will make this conference a success. We're really glad you're here.

See you in ProteoAix and enjoy the congress! "

Cheers

**Montserrat Carrascal** – President of SeProt

### Welcome address by the President of the Portuguese Proteomics Society (ProCura Omics)

Traditions are built in small steps and over time.

It is a pleasure to collaborate on the organization of the third joint meeting with our friends from the French and Spanish proteomics societies. A special thanks to the French Proteomics Society, which this year was responsible for organizing this event. The dynamism of science requires knowledge sharing, peer collaboration and the pleasure of doing what we love most. See you at Aix-en-Provence in what we all hope will be enriching for each of us and memorable for all.

Sincerely yours

**Francisco Amado** - President of PROCURA

### Welcome address by the President of the French Proteomics Society

Dear colleagues,

It is a pleasure for the French Proteomics Society to co-organize this third edition of the joint meeting with our friends of the Spanish Proteomics Society (SEProt) and of the Portuguese Proteomics Society (Rede ProCura). The event will occur in our beautiful city of Aix-en-Provence, a city that can be reached in few hours by plane from the main cities in Spain and Portugal. Last year, ProteoVilamoura was a great success and we tried to keep the same format, giving a maximum exposure to our younger colleagues. Once again, sponsors reacted very positively to this trinational event and EuPA offered several travel grants. Our three societies are very active in the field of proteomics and it should be a great moment of science with a high quality program. We wish you many fruitful exchanges and many new collaborations.

Sincerely yours,

**Franck Vandermoere** - President of the French Proteomics Society, FPS



# Organizing and Local Committees

- **Organizing Committee**

- **Samuel Granjeaud, CRCM**
- **Patrick Fourquet, CRCM**
- **Pascal Mansuelle, IMM**
- **Luc Camoin, CRCM**

- **Local Committee**

- **Stéphane Audebert, CRCM**
- **Alexia Fataka, CRCM**
- **Yves Toiron, CRCM**
- **Emilie Baudelet, CRCM**
- **Sofia Bekdouche, CRCM, IMM, MCT**
- **Théo-Bob Muller, MCT**
- **Maya Belghazi, IMM**
- **Christophe Verthuy, IMM**
- **Manel Khelil-Berbar, IMM**



# Scientific Committee

- **Francisco Amado** University of Aveiro, Portugal
- **Ana Varela Coelho** ITQB NOVA, Lisboa, Portugal
- **Bruno Manadas** Center for Neuroscience and Cell Biology – University of Coimbra, Portugal
- **Montserrat Carrascal** Institute of Biomedical Research of Barcelona - CSIC, Barcelona, Spain
- **Felix Elortza** Proteomics Platform, CIC bioGUNE, CIBERehd, ProteoRed-ISCI, Bizkaia Science and Technology Park, Derio, Spain
- **Carmen Duran Ruiz** University of Cádiz, Spain
- **Mariette Matondo** Institut Pasteur, Paris, France
- **Nicolas Desbenoit** Centre de Génomique Fonctionnelle de Bordeaux (CGFB), Bordeaux, France
- **Franck Vandermoere** Institut de Genomique Fonctionnelle, Montpellier, France
- **Daniel Lafitte** Aix Marseille université, Marseille, France
- **Régine Lebrun** Institut de Microbiologie de la Méditerranée FR 3479, Marseille, France
- **Luc Camoin** Centre de Recherche en Cancérologie de Marseille, Marseille, France





# Sponsors

## PLATINUM



## GOLD



SILVER

EVUSEP

PREOMICS

SHIMADZU

SCIEX

LNT

Eurisotop

BRONZE

Cell Signaling  
TECHNOLOGY®

PARTNER



ProFI  
PROTEOMICS



BIOGNOSYS  
NEXT GENERATION PROTEOMICS



RESYN BIOSCIENCES

PROCURA  
OMICS<sup>PT</sup>

fps  
40 years

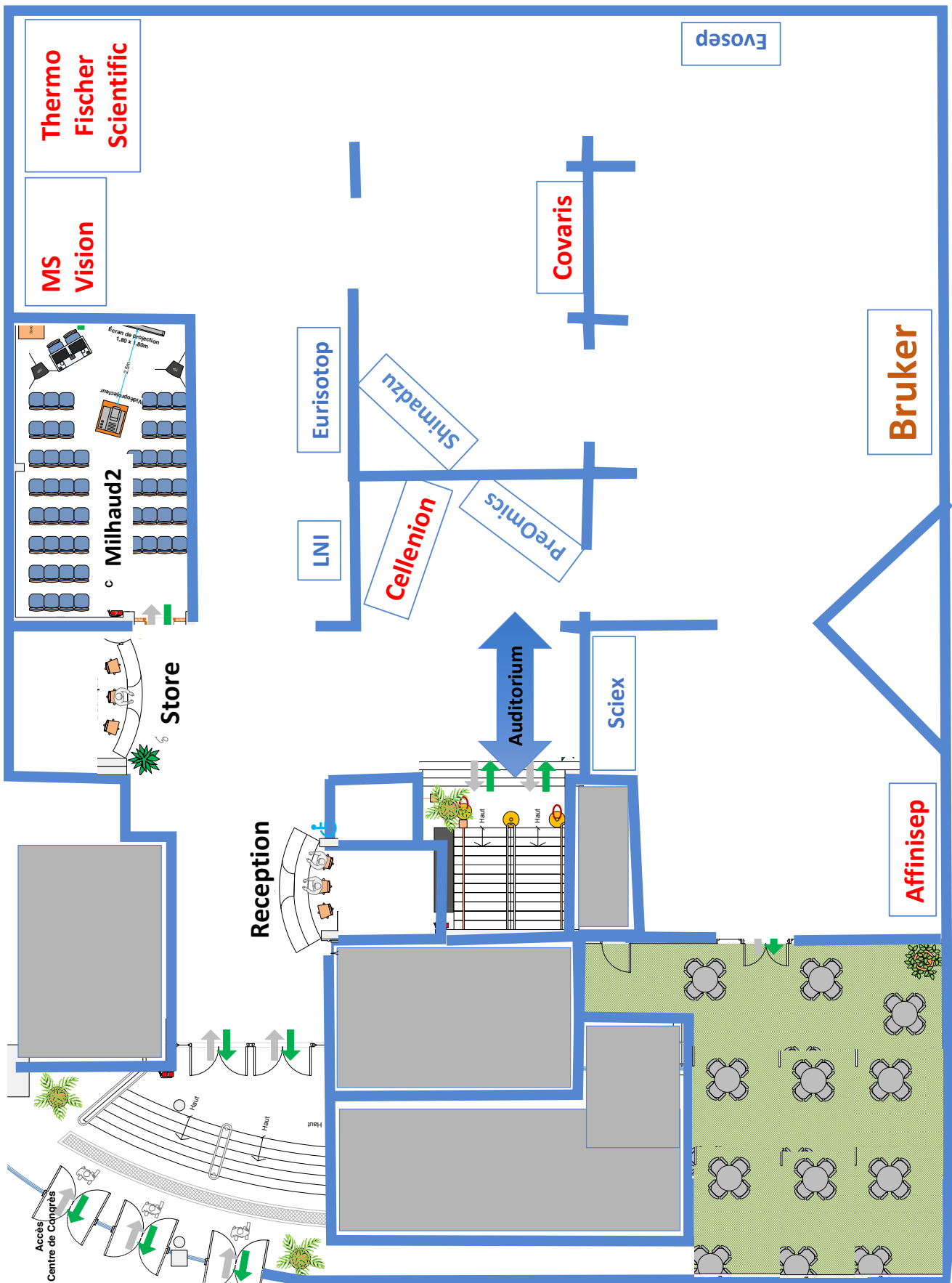


SEProt  
Sociedad Española  
de Proteómica

With the support of



# Exhibition map

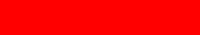











# Scientific Program



## THEMES

	Clinical and disease proteomics, Cancer
	Bioinformatics
	Post-translational modifications
	Proteomics of model organisms, microbial proteomics
	Interactomics and protein structure
	Structural proteomics
	Proteomics in cell biology; cellular and functional proteomics
	Mass spectrometry imaging

## Tuesday 20 June 2023

<b>9:30 - 14:30 Registration</b>		
13:00 - 14:30	Coffee and refreshments welcome	
<b>14:30 - 15:00 Opening Ceremony - Aix-en-Provence Convention Center</b>		Audit
	Moderator: Franck Vandermoere	Audit
<b>15:00 - 15:30 FPS Thesis Prize</b>		
15:00	<b>Charlotte Brun</b> - Development of quantitative analytical strategies for the study of proteins, their phosphorylation and glycation	
<b>15:30 - 17:00 Plenary Lectures 1-2</b>		Audit
	Moderator: Daniel Lafitte	
15:30	<b>Charles Pineau</b> - Imaging mass spectrometry in drug development and environmental toxicology	PL1
	Moderator: Montserrat Carrascal Perez	
16:15	<b>Javier Munoz Peralta</b> - Spatial Proteomics re-defines precise cargoes of extra-cellular vesicles	PL2
	Moderator: Pascal Mansuelle	
17:00	<b>Exhibitor Lecture 1: Covaris - Nicolas Autret</b> - Scalable Sample Preparation Methods for LC-MS Assays	SP1
17:30	<b>Exhibitor Lecture 2: Thermo Fisher Scientific - Shakir Shakir</b> - Pushing Frontiers of High-throughput High-resolution Analysis: Orbitrap Technology Unites with a New Star	SP2
18:00	Welcome cocktail	Hall
<b>19:30 END</b>		

## Wednesday 21 June 2023

8:30 - 9:15	<b>Plenary Lecture 3</b> Moderator: Régine Lebrun <b>Julien Marcoux</b> - Structural MS of proteasome complexes: a test-case study	Audit PL3
9:15 - 9:40	<b>Flash Posters 1</b> Moderators: Bastien Arnaud and Adrien Brown	
9:15	<b>Alice Capuz</b> - Immunoglobulins expression by non immune cells	FL1
9:18	<b>Hugo Gizardin-Fredon</b> - Going against the grain: development of denaturing mass photometry for fine-tuning of protein-protein cross-linking reactions	FL2
9:21	<b>Anne Gonzalez de Peredo</b> - Deciphering the signaling mechanisms of the C-type lectin receptor Dectin-1 in myeloid cells using high throughput proteomics	FL3
9:24	<b>Manuelle Ducoux-Petit</b> - Characterization of abundant periprostatic adipose tissue (PPAT) in prostate cancer cells: from inflammation-free expansion to over-secretion of endotrophin	FL4
9:27	<b>Madisson Chabas</b> - Flash MS/MS identification of microbial isolates	FL5
9:30	<b>Alexandre Burel</b> - Pathway Grabber: Exploring proteomics data with KEGG pathways	FL6
9:33	<b>Marta Rojas Torres</b> - Comparison of tissue homogenization protocols for optimization of quantitative proteomics and phosphoproteomic analysis	FL7
9:36	<b>Exhibitor Lecture 3: PreOmics</b> - <b>Quentin Enjalbert</b> - Solution to move to proteomics sample prep standardization for routine analysis	SP3
9:45 - 10:45	<b>SESSION 1</b> Moderator: Mariette Matondo	Audit
9:45	<b>Bertrand Fabre</b> - In Depth Exploration of the Alternative Proteome of <i>Drosophila melanogaster</i>	OL1
10:00	<b>Léo Schlosser-Perrin</b> - Constitutive proteins of lumpy skin disease virion assessed by next-generation proteomics	OL2
10:15	<b>Anne-Aurélié Raymond</b> - Spatial characterization of beta-catenin-mutated hepatocellular adenomas subtypes by proteomic profiling of the tumor rim.	OL3
10:30	<b>Blandine Chazarin</b> - LF-SCardioP: Label-Free approach for the analysis of Single Cardiomyocyte Proteome	OL4
10:45	Coffee break	Hall
11:30 - 12:15	<b>Plenary Lecture 4</b> Moderator: Felix Elortza <b>Juri Rappsilber</b> - Chemical microscopy: at the frontier of proteomics	PL4
12:15	<b>Exhibitor Lecture 4: Bruker</b> - <b>Pierre Olivier Schmit</b> - The 4D-Proteomics™ ecosystem : enabling large-scale personalized medicine research	SP4
12:45	Lunch Break	Hall

14:15 - 15:15	Session Posters	Hall
15:15 - 16:30	<b>SESSION 2</b>	Audit
	Moderator: Ana Varela Coelho	
15:15	<b>Rajesh Durairaj</b> - Sequence analyses and structural annotation of gustatory receptors from <i>Aedes aegypti</i>	OL5
15:30	<b>Cartas Cejudo Paz</b> - Early and progressive neuropathological-stage dependent proteomic alterations at the level of the olfactory tract in human Alzheimer's disease	OL6
15:45	<b>Laurine Lagache</b> - New method machine process to unravel tissue biomarkers and validate it by multiplex MALDI IHC	OL7
16:00	<b>Tristan Cardon</b> - Extending the vision of the proteome: the Ghost proteome	OL8
16:15	<b>Elisabete Morais</b> - Integration of proteomics and metabolic modelling to study the pH-dependent response of pathogenic and commensal <i>Staphylococcus epidermidis</i> strains	OL9
16:30	Coffee break	Hall
17:15 - 18:30	<b>SESSION 3</b>	Audit
	Moderator: Delphine Pflieger	
17:15	<b>Thibaut Leger</b> - Deciphering the hepatotoxic mechanisms induced by chlordecone and its metabolite chlordecol by multi-omics	OL10
17:30	<b>Avais Daulat</b> - Proteomic approach to characterize and target MINK1-PRICKLE1 signaling pathway in triple negative breast cancer	OL11
17:45	<b>Franck Vandermoere</b> - Nanobody-based APMS analysis of endogenous metabotropic glutamate receptor 2 interactome from mice brain	OL12
18:00	<b>Emmanuelle Com</b> - Parallel accumulation - serial fragmentation combined with data-independent acquisition for the successful identification of plasma proteins correlated with entry into sexual maturation in the rainbow trout	OL13
18:15	<b>Exhibitor Lecture 5: Evosep - Stoyan Stoychev</b> - Enabling high-throughput and standardised proteomics using Evosep One and end-to-end, automated sample preparation	SP5
18:30	Annual General Meeting FPS	Audit
19:30	<b>END</b>	



## Thursday 22 June 2023

<b>8:30 - 9:15</b>	<b>Plenary Lecture 5</b> Moderator: Luc Camoin <b>Vadim Demichev</b> - DIA proteomics: advances in technology and data processing towards sensitivity and speed	Audit PL5
<b>9:15 - 9:40</b>	<b>Flash Posters 2</b> Moderators: Maryam Aboutiman and Mehdi Cherkaoui	
9:15	<b>Sarahi Jaramillo Ortiz</b> - A novel DIA label-free proteomics workflow to study protein glycation in plasma samples	FL8
9:18	<b>Vera Mendes</b> - Identification of potential biomarkers related to A $\beta$ deposition through proteomic profiling in patients with Mild Cognitive Impairment	FL9
9:21	<b>Pauline Perdu-Alloy</b> - Optimization of sample preparation and nanoLC-MS/MS methods for single cell proteomic analysis	FL10
9:24	<b>Rui Vitorino</b> - Promising Salivary Protein Markers for Head and Neck Squamous Cell Carcinoma Prognosis	FL11
9:27	<b>Bastien Arnaud</b> - LC-MS/MS analysis of penguin tissues and customization of protein database to identify novel antimicrobial peptides	FL12
9:30	<b>Yasmine Boughanmi</b> - Short modified peptides or dimeric peptide isolated from animal venoms: Are LC-MS-MS, Edman sequencing enough to determine their sequences?	FL13
9:33	<b>Sarah Alilat</b> - Comparison of Bacillus thuringiensis virulence potential on human Caco-2 intestinal cells using multi-omics investigation	FL14
9:36	<b>Exhibitor Lecture 6: Cell Signalling Technology - Maxime Jacquet</b> - Improved immunoaffinity enrichment methods for arginine methylation	SP6
<b>9:45 - 10:45</b>	<b>SESSION 4</b> Moderator: Franck Vandermoere	Audit
9:45	<b>Marie Locard-Paulet</b> - Pitfalls of gene-level annotation applied to functional analysis of protein groups	OL14
10:00	<b>Madalena Monteiro</b> - Proteome profiling of tissue-derived extracellular vesicles and soluble secretome reveals non-invasive candidate markers involved in colorectal cancer proliferation and progression	OL15
10:15	<b>Inês Caramelo</b> - Proteomics characterization of neurons upon oxygen and glucose deprivation – neuronal modulation induced by the secretome of physiologically primed MSCs	OL16
10:30	<b>Florencia Cristina Picchi Figueira</b> - MACHINE learning approach for the stratification of rheumatoid arthritis patients based on circulating proteomic signatures associated with their synovial pathotype	OL17
10:45	Coffee break	Hall
<b>11:30 - 12:15</b>	<b>Plenary Lecture 6</b> Moderator: Francisco Amado <b>Ana Varela Coelho</b> - Commensalism vs pathogenicity in Staphylococcus epidermidis - uncovered by an integrated omics approach	Audit PL6
12:15	<b>Exhibitor Lecture 7: AffiniseP - Florine Hallez</b> - New kits for proteomics sample purification	SP7
12:45	Lunch Break	Hall15

14:15 - 15:15		Session Posters	
15:15 - 16:30		<b>SESSION 5</b>	Audit
		Moderator: Bruno Manadas	
15:15	<b>Chiara Guerrera</b>	- Pushing DIA proteomics analyses of neat plasma to 1000 protein groups ID/h	OL18
15:30	<b>Tsung-Heng Tsai</b>	- Statistical inference of protein turnover <b>CANCELLED</b>	OL19
15:45	<b>Océane Girard</b>	- Multi-OMICS integration analysis for assessing human stem cell models of human peri-implantation development	OL20
16:00	<b>Aurélie Hirschler</b>	- Boosting immunopeptidomics identification using an optimized workflow	OL21
16:15	<b>Exhibitor Lecture 8: Sciex - Heather Chassaing</b> - Comprehensive peptide mapping of proteins using a complementary approach of electron activated dissociation (EAD) and Collision Induced Dissociation (CID) fragmentation techniques		SP8
16:30	Coffee break		Hall
17:15 - 18:30		<b>SESSION 6</b>	Audit
		Moderator: Montserrat Carrascal Perez	
17:15	<b>Bruno Manadas</b>	- Unmasking hidden systemic effects of neurodegenerative diseases	OL22
17:30	<b>Sara Ceccacci</b>	- High-sensitivity phosphoproteomics with dia-PASEF	OL23
17:45	<b>Noelia De Lama</b>	- Development of an orthogonal method for host cell protein characterization via immunoaffinity enrichment and ms analysis	OL24
18:00	<b>Sandrine Frelon</b>	- Tree frogs living in Chernobyl Exclusion Zone exhibit different protein expression patterns than in the control areas	OL25
18:15	<b>Exhibitor Lecture 9: Resyn Biosciences - Stoyan Stoychev</b> - Mag-Net: deep plasma profiling using streamlined and single-shot membrane vesicle enrichment enabled by MagReSyn® hyper-porous magnetic microparticles		SP9
18:30		<b>END</b>	
18:30 - 20:00		Guided tour of the city of Aix-en-Provence	
20:00		ProteoAix Dinner at AQUABELLA HOTEL	

## Friday 23 June 2023

8:30 - 9:15	<b>Plenary Lecture 7</b> Moderator: Maria Carmen Duran Ruiz <b>Ana Martinez Del Val</b> - Hybrid-DIA: Intelligent Data Acquisition for Online Targeted and Discovery Proteomics Applied to Phosphoprotein signaling in Colorectal Cancer	Audit PL7
9:15	<b>Exhibitor Lecture 10: CELLENION</b> - <b>Fabiana Izaguirre</b> - cellenONE for multiplexed and label-free single cell proteomics sample preparation	SP10
9:45 - 10:45	<b>SESSION 7</b> Moderator: Francisco Amado	Audit
9:45	<b>Baptiste Mouysset</b> - Deciphering the molecular mechanism of action of repurposed drugs using Thermal Proteome Profiling	OL26
10:00	<b>Jean Armengaud</b> - Taxonomical exploration of complex samples by proteomics-derived proteotyping	OL27
10:15	<b>Lucia Grenga</b> - Deep proteomic profiling reveals the dynamics of the host- microbiota interactions from cystic fibrosis sputum samples	OL28
10:30	<b>Alexandra Moreira Pais</b> - Exploring prostate carcinogenesis progression through urine proteome profiling: an animal model study	OL29
10:45	Coffee break	Hall
11:30 - 12:15	<b>Plenary Lecture 8</b> Moderator: Daniel Lafitte <b>Albert Heck</b> - What do we know about (your) antibodies? Novel insights from novel techniques in mass spectrometry	Audit PL8
12:15	<b>Exhibitor Lecture 11: MS Vision</b> - <b>Steven Daly</b> - Teaching an old dog new tricks: Extending the utility of your mass spectrometer	SP11
12:45 - 13:15	<b>CLOSING SESSION / AWARDS</b>	Audit

# Symposium Abstracts





# FPS Thesis Prize

## **Development of quantitative analytical strategies for the study of proteins, their phosphorylation and glycation**

Charlotte Brun, CEA DAM, Bruyères-le-Chatel, France

Proteomic analysis using mass spectrometry enables the study of complex biological systems by identifying and quantifying proteins while characterizing their structure. This approach is also applicable to investigating post-translational modifications (PTMs) of proteins, which add another layer of complexity to the proteome. My thesis focused on developing analytical methods to analyze two specific PTMs: phosphorylations and glycations.

Firstly, a quantitative method was developed to analyze phosphorylations in bear muscle tissue. A complete semi-automated approach was established to efficiently analyze phosphorylations on a large scale. This involved optimizing the various steps of proteomic analysis, including sample preparation, LC-MS/MS analysis, and data processing. By implementing automated hpH RP fractionation, IMAC enrichment, HCD fragmentation, and the use of MaxQuant software, a robust identification and quantification of numerous phosphopeptides was achieved. This method was subsequently applied to studying muscle regulations during bear hibernation. To enhance the biological interpretation of results in non-model species, tools were set up to search for homologous human phosphorylation sites and leverage information from functional databases.

The second focus of my thesis was the study of glycations, a complex family of PTMs that have been underexplored in label-free proteomic analysis of biological fluids. An LC-MS method was developed to analyze glycations in plasma and zebra finch hemolysate, enabling the determination of glycation rates on major proteins. Additionally, a SEC-MS method was implemented to analyze the structure of the haemoglobin complex. These techniques pave the way for future analysis of plasma and hemolysates from various bird species, as well as bats. The ultimate goal is to explore the extent of glycation phenomenon in these flying species, which maintain high blood sugar concentrations without experiencing the typical consequences observed in mammals, such as diabetes. However, analyzing whole proteins provides only a limited view of the modified proteome, offering insight into only a few major glycated proteins without pinpointing the exact glycation events on protein sequences. Consequently, a label-free quantitative proteomic analysis method was developed as a second step. It involved automated cartridge-based BAC enrichment and peptide fractionation protocols for analyzing glycations in plasma. Although automated BAC enrichment proved ineffective with prototype cartridges on plasma samples, better performance is anticipated with larger capacity cartridges. Moreover, the addition of hpH RP peptide fractionation prior to BAC enrichment of plasma peptides improved method repeatability and now allows for high-throughput studies of glycations in the complex plasma fluid.

Throughout this thesis, automated methods were developed for global proteome or modified peptide analysis. Considering the benefits of automation, such as repeatability and throughput, it is evident that automated sample preparation will be essential in the future. This work also emphasized that no universal sample preparation protocol exists, highlighting the importance of adapting the different preparation steps to the sample type, available quantity, and desired analysis approach (global or targeted on specific protein families or modifications).



# Plenary Conferences



## Imaging mass spectrometry in drug development and environmental toxicology

Charles Pineau <sup>1,2</sup>

<sup>1</sup> Inserm UMR\_S1085, IRSET, Univ Rennes, 35000 Rennes, France

<sup>2</sup> Protim Core Facility, BIOSIT - UAR 3480 CNRS - US18 Inserm, Univ Rennes, 35000 Rennes, France

✉: [charles.pineau@inserm.fr](mailto:charles.pineau@inserm.fr)

Among the numerous applications of mass spectrometry, MALDI imaging mass spectrometry (IMS) is a truly blooming field in biological research. IMS is widely recognized as a valuable tool for identifying a variety of compounds including metabolites, lipids and proteins from tissue sections, as it also provides spatial and quantitative information about the analytes *in situ*. Hence, IMS has naturally emerged as a robust and versatile technique for drug development pharmacokinetics and toxicological studies. It is particularly powerful for distribution studies as it offers several advantages over conventionally used LC-MS and whole-body autoradiography (QWBA). Indeed, IMS does not require radioactive labeling and can be used for the simultaneous monitoring of a drug and its metabolites directly in tissue sections, thereby preserving the crucially needed spatial information.

Despite the increasing improvements of IMS, the low detection sensitivity of some compounds remains an important challenge to overcome. We will discuss our developments ranging from on-tissue chemical derivatization (OTCD) that consists in modifying the chemical structure of analytes in order to improve their ionization yields to the use and benefits of three-dimensional IMS. Examples will be given on the use of the technology for the development of pharmaceutical drugs and for understanding the mechanisms that underlie the toxicity of environmental pollutants.



**SPATIAL PROTEOMICS RE-DEFINES PRECISE CARGOES OF EXTRA-CELLULAR VESICLES**Javier Muñoz<sup>1,2</sup>

<sup>1</sup>Cell Signaling and Clinical Proteomics Group. Biocruces Bizkaia Health Research Institute.  
48903, Barakaldo, Spain

<sup>2</sup>Ikerbasque, Basque foundation for science, 48011 Bilbao, Spain.

Virtually all cells secrete extra-cellular vesicles (EVs) containing a wide range of intracellular biomolecular material with key regulatory functions in cell-cell communication processes. The encapsulated content of EVs seems to reflect the cell of origin and, consequently, EVs have a huge potential as biomarkers for cancer and many other diseases. However, current approaches for purification of EVs possess inherent limitations and technical biases, resulting in heterogeneous preparations contaminated by other EVs subtypes and non-vesicular nanoparticles. This is especially problematic when specific regulatory functions are attributed to EVs subtypes.

Here, we report on the development of a spatial proteomic strategy enabling the deconvolution of the protein cargo of small EVs from non-vesicular particles. Crude EVs pellets (purified by differential ultra-centrifugation) are separated in a density gradient, analysed by LC-MS/MS and unambiguously assigned to different compartments by protein correlation profiling. Using this approach, we have systematically analyzed a panel of 14 cancer cell lines providing a highly curated catalogue of pan-human and ubiquitous proteins present in these vesicular entities. Our results have significant implications in the field of EVs, including the mechanisms of cargo selection. In addition, we demonstrate the applicability of our approach to analyze dynamic cellular responses and also to identify biomarkers in EVs purified from human biofluids.





## Structural MS of proteasome complexes: a test-case study

Julien Marcoux, IPBS, Toulouse, France

Structural Mass Spectrometry (MS) approaches are now important tools for integrative structural biology. They give access to a wide range of information, from protein primary to quaternary structure, dynamics and interactions. Top-Down MS [1-2] provides a thorough description of proteoforms present in the sample. Native MS [3] and Mass Photometry [4] of entire complexes inform about their stoichiometry as well as ligand binding. Cross-linking MS [5] can identify new binding partners and provides distance restraints that can be further used for molecular modeling. Finally, Hydrogen-Deuterium eXchange MS (HDX-MS) investigates solvent accessibility, providing information on dynamics and highlighting binding interfaces or long-range allosteric changes [6].

This talk will take advantage of the wide diversity of proteasome complexes to illustrate how these innovative techniques can be applied, complementarily to shotgun proteomics and high-resolution structural methods, to gather precious and orthogonal structural information.

### References:

- [1] Živković D, Sanchez Dafun A, Menneteau T, Schahl A, Lise S, Kervarrec C, Toste Rêgo A, da Fonseca PCA, Chavent M, Pineau C, Burlet-Schiltz O, Marcoux J\*, Bousquet MP\* (2022) “Proteasome complexes experience profound structural and functional rearrangements throughout mammalian spermatogenesis” PNAS 119(15):e2116826119.
- [2] Dafun AS, Živković D, Leon-Icaza SA, Möller S, Froment C, Bonnet D, Almeida de Jesus A, Alric L, Quaranta-Nicaise M, Ferrand A, Cougoule C, Meunier E, Burlet-Schiltz O, Ebstein F, Goldbach-Mansky R, Krüger E, Bousquet Dubouch MP\*, Marcoux J\* (2023) “Establishing 20S proteasome genetic, translational and post-translational status from precious biological and patient samples with Top-Down MS” Cells 12, 844
- [3] Moscovitz O, Ben-Nissan G, Fainer I, Pollack D, Mizrahi L, Sharon M (2015) “The Parkinson's-associated protein DJ-1 regulates the 20S proteasome” Nat Commun. 6:6609.
- [4] Sonn-Segev A, Belacic K, Bodrug T, Young G, VanderLinden RT, Schulman BA, Schimpf J, Friedrich T, Dip PV, Schwartz TU, Bauer B, Peters JM, Struwe WB, Benesch JLP, Brown NG, Haselbach D, Kukura P (2020) “Quantifying the heterogeneity of macromolecular machines by mass photometry” Nat Commun. 11(1):1772.
- [5] Wang X, Chemmama IE, Yu C, Huszagh A, Xu Y, Viner R, Block SA, Cimermancic P, Rychnovsky SD, Ye Y, Sali A, Huang L. (2017) “The proteasome-interacting Ecm29 protein disassembles the 26S proteasome in response to oxidative stress” JBC. 292(39):16310-16320.
- [6] Lesne J\*, Locard-Paulet M\*, Parra J, Zivković D, Menneteau T, Bousquet MP, Burlet-Schiltz O, Marcoux J (2020) “Conformational maps of human 20S proteasomes reveal PA28- and immunodependant inter-ring crosstalks” Nat Commun. 11(1):6140.

## Chemical microscopy: at the frontier of proteomics

Juri Rappsilber, Technische Universität, Insitute für Biotechnologie, Berlin

Accurately modelling the structures of proteins and their complexes using artificial intelligence is revolutionizing molecular biology. Remaining challenges result from the limited information available on specific protein states including protein-protein interactions. We present how proteomics data enable systematically modelling novel protein assemblies and how proteomics data can inject information about in-cell states of proteins into the modelling workflow. Taking a decisively in-cell approach allows discovering aspects of biology that are lost when lysing cells. Indeed, we propose novel structural models of 153 dimeric and 14 trimeric protein assemblies in the model organism *Bacillus subtilis*. We report and validate novel interactors of central cellular machineries that include the ribosome, RNA polymerase, and pyruvate dehydrogenase, assigning function to several uncharacterized proteins. Our approach uncovers protein-protein interactions inside intact cells, provides structural insight into their interaction interfaces, and is applicable to genetically intractable organisms, including pathogenic bacteria. Facile generation of pseudo-atomic models of PPIs provides a plausible shortcut to the improved functional understanding of the many currently uncharacterised proteins.



## DIA proteomics: advances in technology and data processing towards sensitivity and speed

Vadim Demichev, Charité – Universitätsmedizin, Institute of Biochemistry, Berlin

Recently, data-independent acquisition (DIA) proteomics has gained significant popularity for a wide range of applications. The strengths of DIA include high proteomic depth, data completeness and quantitative accuracy. Novel DIA-based technologies are being actively developed, addressing many of the historical shortcomings of DIA. In this talk, I will discuss the factors that determine the performance of DIA methods and their optimisation for the specific experiment, as well as how recent progress in technologies and computational data analysis can boost the performance of DIA proteomics and enable new applications.



## Commensalism vs pathogenicity in *Staphylococcus epidermidis* - uncovered by an integrated omics approach

LG Gonçalves<sup>1</sup>, E Morais<sup>1</sup>, F Magalhães<sup>1</sup>, S Santos<sup>1</sup>, LP Gomes<sup>1</sup>, J Armengaud<sup>2</sup>, M Zimmermann-Kogadeeva<sup>3</sup>, A Gil<sup>4</sup>, M Miragaia<sup>1</sup>, AV Coelho<sup>1</sup>

<sup>1</sup>ITQB NOVA, Oeiras, Portugal, <sup>2</sup>SPI, Université Paris-Saclay, CEA, INRAE, Bagnols-sur-Cèze, France, <sup>3</sup>EMBL, Heidelberg, DE, <sup>4</sup>Univ Aveiro, Portugal

SE is part of skin microbiota and contributes to its homeostasis and protection against pathogens. However, it is the most frequent cause of medical device-associated infections. Skin isolates belonging to clonal complex 2 (CC2) lineage are the major colonizers sharing their ecological niche with other minor genetic backgrounds (non-CC2). From genomic and proteomic data it was possible to identify relevant differences between the metabolic and biological processes of both lineages. Additionally, the intracellular metabolome and proteome associated with both lineages under pH environmental changes, mimicking the transition from the skin (pH 5.5) to blood (pH 7.4) were evaluated and showed specific responses [1]. The CC2 strain seems more prepared to survive in blood and to promote adhesion to medical-devices. The obtained results were complemented with time-course exometabolomic data during bacterial growth, which were also integrated in genome-scale metabolic models. Following a proteogenomic approach, obtained proteomics data are being used to refine the annotation of both strains genomes.

[1] L G Gonçalves, S Santos, L P Gomes, J Armengaud, M Miragaia, A V Coelho "Skin-to-blood pH shift triggers metabolome and proteome global remodelling in *Staphylococcus epidermidis*" *Front Microbiol.*, 2022, 13:1000737. doi: 10.3389/fmicb.2022.1000737

This work was partially supported by project PTDC/CVTCVT/ 29510/2017 funded by Fundação para a Ciência e Tecnologia; Projects LISBOA-01-0145-FEDER-007660 (Microbiologia Molecular, Estrutural e Celular) and UID/Multi/04378/2019) funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI); by ONEIDA project (LISBOA-01-0145-FEDER- 016417) co-funded by FEEI - "Fundos Europeus Estruturais e de Investimento" from "Programa Operacional Regional Lisboa2020"; by national funds through FCT; by Transnational Cooperation Portugal/France – Pessoa Program, supported by FCT/Campusfrance; and by SymbNet Project (IGC-UE-Horizon 2020-952537). LGG was financed by an FCT contract according to DL57/2016, (SFRH/BPD/111100/2015).





## Hybrid-DIA: Intelligent Data Acquisition for Online Targeted and Discovery Proteomics Applied to Phosphoprotein signaling in Colorectal Cancer

Ana Martinez Del Val, University of Copenhagen, Denmark

Achieving sufficient coverage of regulatory phosphorylation sites by mass spectrometry (MS)-based phosphoproteomics for signaling pathway reconstitution is challenging, especially when analyzing tiny sample amounts. To address this, we present an innovative hybrid data-independent acquisition (DIA) strategy (hybrid-DIA) that combines targeted and discovery proteomics through an Application Programming Interface (API) to dynamically intercalate DIA scans with accurate triggering of multiplexed tandem mass spectrometry (MSx) scans of predefined (phospho)peptide targets. By spiking-in heavy stable isotope labeled phosphopeptide standards covering seven major signaling pathways, we benchmarked hybrid-DIA against state-of-the-art targeted MS methods (i.e. SureQuant) using EGF-stimulated HeLa cells and found the quantitative accuracy and sensitivity to be comparable while hybrid-DIA also profiled the global phosphoproteome. To demonstrate the robustness, sensitivity and biomedical potential of hybrid-DIA, we profiled chemotherapeutic agents in single colon carcinoma multicellular spheroids and evaluated the phospho-signaling difference of cancer cells in 2D vs 3D culture. Finally, we show the applicability of this methodology to find novel kinase targets in clinical samples.



**What do we know about (your) antibodies?  
Novel insights from novel techniques in mass spectrometry**

Albert J. R. Heck<sup>1,2</sup>

<sup>1</sup> *Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.*

<sup>2</sup> *Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands.*

In our body we produce every day huge amounts of antibodies, of which many end up in circulation. It has been estimated that humans can make about  $10^{15}$  distinct antibody clones, all exhibiting a slightly different sequence. This huge number has so far refrained many from charting whole serum antibody, or immunoglobulin (Ig), repertoires. We recently developed an LC-MS based antibody repertoire profiling method for studying immunoglobulins in a quantitative manner. We analyzed a variety of samples from both healthy as well as diseased donors and made some paradigm-shifting observations. Firstly, circulating Ig repertoires are much simpler than anticipated, dominated by a few hundred clones. Second, the clonal repertoires are entirely unique for each donor, both for IgG1 and IgA1 we found virtually no overlap between individuals. Conversely, longitudinal samples from the same (healthy) donor showed a far-reaching overlap even when samples were taken months apart. In the repertoires of severely ill patients, more plasticity was observed. Charting the Ig repertoires in diseased donors allowed the selection of Ig clones of interest, i.e., those emerging after the onset of disease. We demonstrate that these latter serum clones can be fully *de novo* sequenced by combining top-down and bottom-up analysis and iterative software algorithms to connect these layers of data. In this manner antigen-directed antibodies could be identified and developed into novel therapeutics.

All (sub)classes of immunoglobulins have unique structural features. In our work we also investigate the structures of IgA1 and IgM, and reveal that they are not always as described in text-books. I will present work through which we redefine the molecular composition of circulatory IgM. Using single-particle charge-detection mass spectrometry, mass photometry, proteomics, and immunochemical assays, we reveal that circulatory IgM is (re)defined by the universal presence of an additional protein component. We study the covalent attachment of this protein and evaluate its effect on the binding of IgM to several receptors. Lastly, our data reveal the distinctiveness of the circulatory and secretory IgM.



# Oral presentations



## In Depth Exploration of the Alternative Proteome of *Drosophila melanogaster*

Bertrand FABRE<sup>a</sup>, Sebastien CHOTEAU<sup>b</sup>, Carine DUBOÉ<sup>c</sup>, Carole PICHEREAUX<sup>d</sup>, Audrey MONTIGNY<sup>c</sup>, Dagmara KORONA<sup>e</sup>, Michael DEERY<sup>f</sup>, Mylène CAMUS<sup>d</sup>, Christine BRUN<sup>b</sup>, Odile BURLET-SCHILTZ<sup>d</sup>, Steven RUSSELL<sup>e</sup>, Jean-Philippe COMBIER<sup>c</sup>, Kathryn LILLEY<sup>f</sup>, Serge PLAZA<sup>c</sup>

<sup>a</sup> Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, <sup>b</sup> Turing Centre for Living Systems, Marseille, FRANCE, <sup>c</sup> Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, Auzeville-Tolosane, FRANCE, <sup>d</sup> Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, Toulouse, FRANCE, <sup>e</sup> Cambridge Systems Biology Centre and Department of Genetics, University of Cambridge, Cambridge, UK, <sup>f</sup> Cambridge Centre for Proteomics, Cambridge Systems Biology Centre and Department of Biochemistry, University of Cambridge, Cambridge, UK

Recent studies have shown that hundreds of small proteins were occulted when protein-coding genes were annotated. These proteins, called alternative proteins, have failed to be annotated notably due to the short length of their open reading frame (less than 100 codons) or the enforced rule establishing that messenger RNAs (mRNAs) are monocistronic. Several alternative proteins were shown to be biologically active molecules and seem to be involved in a wide range of biological functions. However, genome-wide exploration of the alternative proteome is still limited to a few species. We recently performed a deep analysis of the alternative proteome of *Drosophila melanogaster* which enabled the identification of 401 alternative proteins. Subcellular localization, protein domains, and short linear motifs were predicted for 235 of the alternative proteins identified and point toward specific functions of these small proteins. Several alternative proteins had approximated abundances higher than their canonical counterparts, suggesting that these alternative proteins are actually the main products of their corresponding genes. Finally, we observed 14 alternative proteins with developmentally regulated expression patterns and 10 induced upon the heat-shock treatment of embryos, demonstrating stage or stress-specific production of alternative proteins.

## Constitutive proteins of lumpy skin disease virion assessed by next-generation proteomics

Léo SCHLOSSER--PERRIN<sup>a</sup>, Philippe HOLZMULLER<sup>a</sup>, Philippe CAUFOUR<sup>a</sup>, Bernard FERNANDEZ<sup>a</sup>,  
Guylaine MIOTELLO<sup>b</sup>, Jean ARMENGAUD<sup>b</sup>, Stéphane BERTAGNOLI<sup>c</sup>, Aymeric NEYRET<sup>d</sup>, Nouredine DAHMANI<sup>a</sup>

<sup>a</sup> UMR ASTRE, CIRAD, INRAE, <sup>b</sup> Département Médicaments et Technologies pour la Santé (DMTS), Université Paris Saclay, CEA, INRAE, <sup>c</sup> IHAP, Université de Toulouse, INRAE, ENVT, <sup>d</sup> CEMIPAI, University of Montpellier

Lumpy skin disease virus (LSDV) is the causative agent of an economically important cattle disease which is notifiable to the World Organization for Animal Health. Over past decades, the disease has spread at an alarming rate throughout the African continent, the Middle East, Eastern Europe, the Russian Federation and many Asian countries. While multiple LSDV whole genomes have made further genetic comparative analyses possible, knowledge on the protein composition of the LSDV particle remains lacking. Furthermore, this first incursion within the *Capripoxvirus* genus proteome represents one of very few proteomic studies beyond the sole *Orthopoxvirus* genus, for which most of the proteomic studies have been performed on Vaccinia Virus.

The large and complex enveloped LSD virus (LSDV) particle encloses a double-stranded linear DNA genome of 151 kbp, comprising 156 predicted ORFs, together with a variety of proteins that are not yet identified. In parallel with a recent widespread expansion of LSD, many LSDV whole genomes have been sequenced but knowledge on protein composition of the LSDV particle remains missing. In this study, LSD mature virions (MV) from strain KSGP-0240 were purified through a multistep ultracentrifugation process. The protein composition of LSD virions was then analysed using label-free shotgun proteomics, based upon nano-liquid chromatography and tandem mass spectrometry. This procedure resulted in the identification of a total of 112 LSDV proteins and 1,929 host proteins. Considering that this first MV proteome extended beyond packaged proteins into the field of contaminants, an analytical methodology was developed and made it possible to select 66 viral proteins and 66 host proteins as candidates for packaging into MV. These viral and host proteins were analysed comparatively with proteins previously demonstrated to be constitutive of the Vaccinia virus MV particle. Offering for the first time a comprehensive proteomic analysis of a LSDV strain, this study contributes to our understanding of the structural features of infectious LSDV MV particles and paves the way for further systematic proteomic characterization of other LSDV strains.

## Spatial characterization of beta-catenin-mutated hepatocellular adenomas subtypes by proteomic profiling of the tumor rim.

Sylvaine DI TOMMASO<sup>a</sup>, Cyril DOURTHE<sup>a</sup>, Jean-William DUPUY<sup>b</sup>, Nathalie DUGOT-SENANT<sup>c</sup>, David CAPPELLEN<sup>d</sup>, H el ene CAZIER<sup>e</sup>, Valerie PARADIS<sup>e</sup>, Jean-Fr ed eric BLANC<sup>f</sup>, Brigitte LE BAIL<sup>g</sup>, Charles BALABAUD<sup>h</sup>, Paulette BIOULAC-SAGE<sup>h</sup>, Frederic SALTEL<sup>a</sup>, Anne-Aur elie RAYMOND<sup>a</sup>

<sup>a</sup> Univ.Bordeaux, Inserm UMR1312 BoRdeaux Institute of onCology (BRIC), F-33000 Bordeaux, France AND Oncoprot Platform, TBM-Core US 005, F-33000 Bordeaux, France, <sup>b</sup> Univ.Bordeaux, Proteomics Platform, F-33000 Bordeaux, France, <sup>c</sup> Histology Platform, TBM-Core US 005, F-33000 Bordeaux, France, <sup>d</sup> Univ.Bordeaux, Inserm UMR1312 BoRdeaux Institute of onCology (BRIC), F-33000 Bordeaux, France AND Bordeaux University Hospital Center, Tumor Bank and Tumor Biology Laboratory, F-33600 Pessac, France, <sup>e</sup> Pathology Department, H opital Henri Mondor, Cr eteil, France, <sup>f</sup> Univ.Bordeaux, Inserm UMR1312 BoRdeaux Institute of onCology (BRIC), F-33000 Bordeaux, France AND Department of Hepatology and Oncology, Bordeaux University Hospital, INSERM CIC 1401 F-33000 Bordeaux, France, <sup>g</sup> Univ.Bordeaux, Inserm UMR1312 BoRdeaux Institute of onCology (BRIC), F-33000 Bordeaux, France AND Department of Pathology, Bordeaux University Hospital, F-33000 Bordeaux, France, <sup>h</sup> Univ.Bordeaux, Inserm UMR1312 BoRdeaux Institute of onCology (BRIC), F-33000 Bordeaux, France

**Background and aims:** Hepatocellular adenomas (HCA) are rare liver benign tumors classified at the clinico-pathological, genetic and proteomic levels. The b-catenin activated (b-HCA) subtypes have several types of mutations in the  $\beta$ -catenin gene (*CTNNB1*) associated with different risks of malignant transformation or bleeding. Glutamine synthetase (GS) is a surrogate marker of  $\beta$ -catenin pathway activation linked to the risk of malignant transformation. Recently we revealed the GS overexpression in the rim of b-HCA exon3 S45 and b-HCA exon7/8 compared to the rest of the tumor. A difference in vascularization was associated with this rim, revealed by a diffuse CD34 labeling only at the center of the tumor.

In this study, we sought to characterize this tumor heterogeneity in order to better understand its physio-pathological involvement.

**Methods:** Using Mass Spectrometry (MS) imaging, genetic and proteomic analyses combined with laser capture, we compared the center with the tumor rim and with the adjacent non-tumor area.

**Results:** First, the tumor rim displayed the same mutation as the tumor center and was therefore part of it. MS Imaging showed differential spectral profiles between the rim and the rest of the tumor. Proteomic profiling revealed 40 proteins significantly differentially expressed between the rim and the center of the tumor. A large majority of these proteins were associated with metabolism with deregulations comparable to a perivenous expression profile in a normal hepatocyte context.

**Conclusions:** The differential phenotype of the tumor rim of b-HCA exon3 S45 and exon7/8 don't depend on the mutational status of CTNNB1. In a context of sinusoidal arterial pathology, this tumor heterogeneity at the rim has perivenous characteristics, and might be rather due to a functional peripheral venous drainage.

## LF-SCardioP: Label-Free approach for the analysis of Single Cardiomyocyte Proteome

Dr Blandine CHAZARIN<sup>a</sup>, Aleksandra BINEK<sup>a</sup>, Simion KREIMER<sup>a</sup>, Alexandre HUTTON<sup>b</sup>, Jonathan KRIEGER<sup>c</sup>, Joshua CANTLON<sup>d</sup>, Jesse MEYER<sup>b</sup>, Anja KARLSTAEDT<sup>e</sup>, Jennifer VAN EYK<sup>a</sup>

<sup>a</sup> ACBRI, Cedars Sinai Medical Center, <sup>b</sup> Cedars Sinai Medical Center, <sup>c</sup> Bruker Daltonics, <sup>d</sup> Scienion, <sup>e</sup> Karlstaedt lab, Cedars Sinai Medical Center

### (Introduction)

Single cell transcriptomics is a well-established method helping to decipher cell-to-cell heterogeneity. Nowadays, single cell proteome analysis is emerging as a powerful tool to identify potential subpopulations of cells into the same cell type. This approach was allowed by the technological developments for the cell isolation, the sample preparation, and the sensitivity of the mass spectrometers.

### (Material and method)

We are presenting a study in which we successfully quantify the proteome of 701 single cardiomyocytes obtained from 4 individual wild-type mice. The cell isolation and the sample preparation were performed on CellenONE platform (Scienion) ensuring a unique quality control based on the pictures of the isolated cells and limiting the manipulation of the samples. After digestion, the peptides were resuspended by U3000 HPLC system (Thermo) coupled to a timsTOF scp mass spectrometer (Bruker Daltonics) for the label-free quantification of the peptides. The stability of the LC-MS platform was monitored in real time with repeated injections of HeLa digest. Data analysis was performed using the novel DIANN-PASER software (Bruker). Data interpretation was obtained by using widely accepted databases as KEGG Pathway and Gene Ontology. To decipher potential sub-populations of cardiomyocytes, after filtering, UMAP analysis was performed and proteome signatures of 3 potential sub-populations was characterized.

### (Results and Interpretation)

We were able to establish efficient quality controls using the pictures collected by CellenONE platform and based on the quantified proteome from individual cells. Finally, we were able to quantify the proteome of 701 single cardiomyocytes. Most of the proteins are associated with Mitochondrion, Membrane, Cytosol and very few with Nuclear Matrix (GO terms). This result is in line with the cell biology of cardiomyocytes i.e. a large cytosol containing a large number of mitochondria to produce and to process ATP for heart muscle contraction. Also, we noticed that among the most represented Biological Processes (GO terms) a majority was associated to ATP production/processing. Furthermore, we obtained representation of Negative Regulation of Apoptotic Process comforting us that we capture the proteome of live cardiomyocytes.

Comparing the 3 cell clusters, the distribution of the metabolic pathways (KEGG Pathway) highlights differences between clusters with a higher representation of proteins involved in Oxidative Phosphorylation metabolism in one cell cluster and a metabolism driven by TCA cycle and Carbon metabolism for the 2 other cell clusters. Furthermore, we were able to establish major change in contractile proteins playing a key role in cardiology.

This method was applied to the analysis of Human Cardiomyocytes obtained from healthy transplants. The data interpretation is currently ongoing.

### (Conclusion)

We were able to establish a robust method for the analysis of Single Cardiomyocyte Proteome analysis able to point up 3 potential sub-populations with proteome and metabolic signatures.

## Sequence analyses and structural annotation of gustatory receptors from *Aedes aegypti*

Samia LAIB<sup>a</sup>, Rajesh DURAIRAJ<sup>b</sup>, Aurore LAMY<sup>b</sup>, Cécile BIENBOIRE-FROSINI<sup>c</sup>, Patrick PAGEAT<sup>d</sup>

<sup>a</sup> Department of Bioinformatics and Chemical Communication (D-BICC), Research Institute in Semiochemistry and Applied Ethology (IRSEA), Quartier Salignan, <sup>b</sup> Department of Bioinformatics and Chemical Communication (D-BICC), Research Institute in Semiochemistry and Applied Ethology (IRSEA), Quartier Salignan, Apt, FRANCE, <sup>c</sup> Department of Molecular Biology and Chemical Communication (D-BMCC), Research Institute in Semiochemistry and Applied Ethology (IRSEA), Quartier Salignan, Apt, FRANCE, <sup>d</sup> Research and Education Board, Research Institute in Semiochemistry and Applied Ethology (IRSEA), Quartier Salignan, Apt, FRANCE

The hematophagous female *Aedes aegypti* (Aeag) mosquitoes transmit a wide range of pathogens responsible for diseases such as dengue fever. The mosquitoes are attracted to compounds released by mammalian hosts, including humans and horses. Among attractive semiochemical signals, CO<sub>2</sub> activates gustatory receptors (Gr's) expressed in the mosquito cpA neurons in sensilla. Mosquitoes detect CO<sub>2</sub> and other host-seeking behaviours via Gr's (Gr 1/2/3), which are seven transmembranes (7TM) GPCR proteins. However, the sequence analysis and the experimental structures of Gr's are not yet reported. We employed a computational strategy to examine Gr sequences, orthologues, and evolutionary diversity using ClustalO, BLASTp, and MEGA tools respectively. The membrane topology and signal peptides were predicted using various servers, such as Topcons, TMHMM, etc. We analysed the similarity templates from iTasser, and HHpred for confirming the comparative structural model of AeagGr's constructed by ColabFold. Additionally, we evaluated the model quality using ProFunc, and its stability using GROMACS molecular dynamics simulations (MDS) within a suitable membrane. The results revealed a high sequence identity (>70%) of mosquito Gr's towards intra- and inter-species (such as flies, worms, and insects) orthologs. The phylogenetic tree showed 3 deviated clades for AeagGr's and we identified highly conserved sites between the similar orthologs among 10 species. Most of the selected Gr's showed 7TM internal (IN) N-terminal topology and were modelled properly with ColabFold. The AeagGr 1/2/3 have more than 90% of residues that appeared in the mixture of favoured and allowed regions in the Ramachandran plot. Nevertheless, the dynamics of AeagGr structures showed very stable conformations around 20 ns to 50 ns and few fluctuations were observed in the 7TM regions. Therefore, these AeagGr monomer models are useful for building accurate heterodimers to study dimer interface properties and receptor-ligand interactions.



## Early and progressive neuropathological-stage dependent proteomic alterations at the level of the olfactory tract in human Alzheimer's disease

Cartas-Cejudo PAZ<sup>a</sup>, Lachén-Montes MERCEDES<sup>a</sup>, Cortés ADRIANA<sup>a</sup>, Ferrer ISIDRE<sup>b</sup>, Fernández-Irigoyen JOAQUÍN<sup>a</sup>, Santamaría ENRIQUE<sup>a</sup>

<sup>a</sup> Clinical Neuroproteomics Unit, Navarrabiomed, HUN, UPNA, IdiSNA, <sup>b</sup> Department of Pathology and Experimental Therapeutics, UB, IDIBELL, CIBERNED-ISCIII, Barcelona, Spain

**Introduction:** Smell impairment is one of the earliest features in Alzheimer's disease (AD). Late AD is associated with a large amount of amyloid beta and neurofibrillary tangles composed by hyperphosphorylated Tau aggregates (NFTs) in the olfactory tract (OT). However, the molecular mechanisms associated with the olfactory dysfunction in early neuropathological stages are unknown.

**Methods:** We applied sequential window acquisition of all theoretical fragment ion spectra mass spectrometry (SWATH-MS) in 34 postmortem OTs derived from non-demented (n=17) and AD (n=17) subjects, classified according to the degree of pathology: Braak I-II (with no discernible or a few isolated NFTs) (n=4), Braak III-IV (with low-moderate numbers of NFTs) (n=4) and Braak V-VI (with a huge quantity of NFTs) (n=9). Data analysis was performed using Perseus and RStudio software. AlzData, Metascape and Ingenuity Pathway Analysis were considered for data interpretation.

**Results :** Around 11% of the quantified proteins were differentially expressed. A progressive proteomic impairment was observed across Braak stages. 116, 144 and 335 differential expressed proteins (DEPs) were exclusively detected in Braak I-II, Braak III-IV and Braak V-VI stages, respectively (FC 30%; p-value<0.05). Interestingly, a stage-dependent deregulation of specific pathways was observed, revealing aberrant proteostasis involved in synaptic transmission, axon guidance, and metabolism of amino acids and lipids as part of the functional modules that are disrupted across AD grading. Moreover, our preliminary data point out neuropathological stage-dependent commonalities and differences in terms of olfactory proteostasis, pathway modulation and protein networks. Our workflow is also currently being complemented with secretability analysis in order to increase our knowledge about the AD progression at the olfactory level. Part of the differential datasets showed a gene expression correlation with AD pathology in amyloid beta and/or Tau line AD mouse models.

**Conclusions :** Early AD-dependent protein imbalance was evidenced at the OT level. This proteomic dyshomeostasis was more severe in advanced stages of the disease. Some DEPs are differential features in AD mouse models before AD pathology, suggesting a relevant role during the initial stages of neurodegenerative process.

**Keywords:** Alzheimer's disease, Olfactory tract, Neuropathological stages, SWATH-MS, Systems biology.

## New method machine process to unravel tissue biomarkers and validate it by multiplex MALDI IHC

Laurine LAGACHE<sup>a</sup>, Yanis ZIREM<sup>a</sup>, Alexandre GOOSSEN<sup>a</sup>, Zoltan TAKATS<sup>a</sup>, Isabelle FOURNIER<sup>a</sup>, Michel SALZET<sup>a</sup>

<sup>a</sup> PRISM - INSERM U1192

Integrating machine learning tumour heterogeneity in the cancer diagnosis process is an essential issue to tackle the complexity of the pathology with precision and efficiency. Specific omic molecules identifications for each heterogeneous tumour subpopulations are especially important in order to characterize them in depth. In this way, distinguishing and detecting one clone from another, whatever the molecules analyzed, is an easy and fast process which can be carried out during the diagnosis. Moreover, protein information from each subpopulations are particularly interesting for drug discovery to tailor optimized therapy to the heterogeneous tumour.

For this purpose, the implementation of the concept was first optimized on rat brain (RB). Lipid, protein and peptide MALDI mass spectrometry imaging (MSI) were performed on 5 cerebellum area rat brain replicates. A first MatLab script was developed to process imaging data, providing unsupervised clustering, independently for each data set. 3 clusters were identified with the same spatial location for each images, whatever the molecule analysed, corresponding to the Granular Layer (GL), Molecular Layer (ML) and the White Matter (WM) region of the cerebellum RB. A second Python script provided the lipid, protein and peptide discriminant cluster ions. Microproteomic analysis, and Back side analysis using SPIDERmass technology, were performed to identify the protein ions previously listed. MS/MS analysis were also performed with SPIDERmass to identify the discriminant lipids. Finally, the multi MSI made it possible to demonstrate the presence of various heterogeneous subpopulations, each presenting a specific localized lipid and protein network. This information will make it possible to produce a computer prediction model, which will make it possible to identify the presence of any cluster on any tissue according their specific molecular fingerprint, and validate it by machine learning.

In order to validate the identification of a subpopulation, a multiplex immunohistochemical (IHC) MALDI-MSI technique, using TAG mass technology, is being developed to identify specific biomarkers of the latter on tissue. To do so, antibodies, targeting the discriminating proteins of interest, are modified to add a photo-cleavable group as well as reporters, via a click chemistry reaction. After recognition of the target by each probe, the complex is placed under UV radiation in order to release photo-cleavable tags whose mass to charge ratio are known. Thus, they can be easily detected and assayed by MALDI-MSI in place of the probe. This diagnostic tool aims to be fast, very sensitive and specific, with the possibility of possibility of simultaneous detection of several markers by MS, allowing the identification of several cluster at the same time.

## Extending the vision of the proteome: the Ghost proteome

Dr Tristan CARDON<sup>a</sup>, Diego Fernando GARCIA-DEL RIO<sup>a</sup>, Isabelle FOURNIER<sup>a</sup>, Michel SALZET<sup>a</sup>

<sup>a</sup> PRISM Inserm U1192

For several years we have been working on the identification of an extended proteome. Indeed, most proteomic studies consider a proteome reduced to the proteins referenced (RefProt) in the databases allowing the analysis of data from "bottom-up" approaches. However, there is another vision of the proteome, the "non-referenced or Ghost proteome" one, which contains proteins called "Alternative proteins" (AltProt). These proteins are derived from regions described as non-coding in the annotation of transcripts such as the 5' & 3' UTR regions, a shift in +1 or +2 of the reading frame but also in ncRNAs (nc-, lnc-, circRNA). Not referenced in classical databases, they are therefore very difficult to identify and their function is still poorly understood.

In order to highlight this Ghost proteome, we have developed a methodological pipeline allowing "Bottom-up" approaches integrating AltProt prediction databases (OpenProt) in combination with the classic UniprotKB database. In order to obtain as much information as possible on new AltProt, the non-targeted protein-protein interactions (PPIs) identification strategy: crosslink-MS (XL-MS) was used. This method forms chemical bridges between proteins close of 5 to 30Å, this highlight PPIs between AltProt and RefProt. These interactions allow us to link for the first time and without apriority AltProt to signaling pathways via the application of a Gene Ontology enrichment of the referenced interaction partners. In order to improve the detection of crosslinked peptides in the samples we also combined the XL-MS approach with subcellular fractionation of the cell, allowing to kill two birds with one stone: improving the detection of crosslinked peptides by a de-complexification of the sample and assigning a subcellular localization for the identified AltProts. This approach allowed us to identify 112 AltProts and the detection of 220 PPIs in the cell including 16 involving an AltProt with a RefProt. In the absence of biomolecular tools (antibodies, etc.) allowing targeted studies, the probability of interaction observed by XL-MS was modeled by 3D construction of the proteins coupled to a docking of the AltProt on the RefProt, allowing to confirm the possibility of interaction, while having a coherent distance [5 to 30Å] between the amino acids observed in interaction during the XL-MS analysis. This analysis showed the possibility of interaction with MHC class I, through the identification of a crosslink interaction between HLA-B and an AltProt from the 5'UTR region of the FAM227B gene transcript, in place of the B2M partner of MHC.

Cardon T\*, Garcia-del Rio DF\*, Eyckerman S, Fournier I, Bonnefond A, Gevaert K, Salzet M. Employing non-targeted interactomics approach and subcellular fractionation to increase our understanding of the ghost proteome. *iScience*. Elsevier; 2023 Feb 17;26(2):105943.

## Integration of proteomics and metabolic modelling to study the pH-dependent response of pathogenic and commensal *Staphylococcus epidermidis* strains

Elisabete MORAIS<sup>a</sup>, Susana SANTOS<sup>a</sup>, Elena MATVEISHINA<sup>b</sup>, Eva-Maria GEISSEN<sup>b</sup>, Jean ARMENGAUD<sup>c</sup>, Maria MIRAGAIA<sup>a</sup>, Maria ZIMMERMANN-KOGADEEVA<sup>b</sup>, Luís G GONÇALVES<sup>a</sup>, Ana V COELHO<sup>a</sup>

<sup>a</sup> Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal, <sup>b</sup> European Molecular Biology Laboratory, Heidelberg, Germany, <sup>c</sup> Département Médicaments et Technologies pour la Santé, Université Paris-Saclay, Bagnols-sur-Cèze, France

*Staphylococcus epidermidis* (SE), a common colonizer of the human skin, is often the cause of infections associated with medical devices in immunocompromised patients. Strains that retain pathogenic and commensal potential coexist in human skin belonging to clonal complex 2 (CC2) and non-CC2, respectively <sup>1</sup>. Improved knowledge of the pathogenicity mechanisms of SE is needed to design more effective prevention and treatment strategies. We aim to use our proteomic data as flux constraints to build strain-specific genome-scale metabolic models (GEM) to understand SE metabolic responses to relevant biologic pH states, such as those found in skin colonization and blood infection (pH 5.5 and pH 7.4, respectively).

Specific GEMs for representative CC2 and non-CC2 SE strains were built using as template a manually curated GEM of a reference SE strain, RP62A <sup>2</sup>. Manual curation was performed based on the Best Bidirectional Hits (BBH) for Blastp results of each representative strain against the reference strain annotated genome: 1) enzymatic reactions were maintained if the protein was found in the BBH results; 2) additional enzymatic reactions were added based on strain specific protein capacity to catalyse a known reaction; 3) exchange reactions were updated to include the translocation of metabolites, previously measured in the extracellular medium (unpublished results), across the membrane. Intracellular proteomic data under different pH conditions <sup>3</sup> will be further used to constrain the metabolic fluxes of each strain-specific GEM. First, undetected proteins will be inactivated in the GEM, assuming that their associated processes are not active in each experimental condition under study. Second, for the differentially abundant proteins their fold-change will be used to further constrain the reaction fluxes. To simulate the metabolism of the two strains at different pH, we will perform a flux balance analysis using biomass production as the optimisation parameter. This integrative approach of omics data with metabolic modelling generates more reliable GEMs in relevant biological conditions per strain becoming a powerful tool to interpret other omics datasets.

**References:** 1. D. Espadinha et al., *Front. Microbiol.* 10, 1-17 (2019); 2. T. D. Calvo et al., *Metabolites.* 12, 136 (2022); 3. L. G. Gonçalves and S. Santos et al., *Front. Microbiol.* 13, 1-22 (2022).

**Acknowledgments:** Funding projects: ONEIDA (LISBOA-01-0145-FEDER-016417); LISBOA-01-0145-FEDER-007660 (MOSTMICRO); Transnational Cooperation Portugal/France - Pessoa Program/2019 "Identification of putative bacterial targets for the development of precision antimicrobials against pathogenic *S. epidermidis* - a strategy to prevent medical device associated infections", FCT/CampusFrance; AAC 01/SAICT/2016.; PD/BD/150980/2021; SymbNET Genomics and Metabolomics in a Host-Microbe Symbiosis Network.

## Deciphering the hepatotoxic mechanisms induced by chlordecone and its metabolite chlordecol by multi-omics

Dr Thibaut LÉGER<sup>a</sup>, Sarah ALILAT<sup>b</sup>, Léonie DEC<sup>b</sup>, Kevin HOGVEEN<sup>b</sup>, Pierre-Jean FERRON<sup>c</sup>, Sylvie HUET<sup>b</sup>, Yoann DEVRIENDT-RENAULT<sup>d</sup>, Julien PARINET<sup>d</sup>, Valérie FESSARD<sup>b</sup>, Ludovic LE HÉGARAT<sup>b</sup>

<sup>a</sup> ANSES, Toxicology of Contaminants Unit, <sup>b</sup> ANSES, Toxicology of Contaminants Unit, Fougères, FRANCE, <sup>c</sup> Univ Rennes, INSERM, INRAE, CHU Rennes, NuMeCan institute, Rennes, FRANCE, <sup>d</sup> ANSES, Pesticides and Marine Biotoxins (PBM) unit, Maison-Alfort, FRANCE

Chlordecone (CLD) is a pesticide used to control the banana weevil in the French West Indies from 1970 to 1993. Despite its ban in the 90s, the highly persistence of CLD and contamination of the food webs raised tremendous concern for public health [1]. CLD can induce reprotoxic, neurotoxic and hepatotoxic effects in humans and is involved in prostate cancer [2] and liver fibrosis potentiation [3]. In liver, although CLD is metabolized into the phase I metabolite chlordecol (CLD-OH), it is highly biopersistent. CLD has been described as an estrogen receptor agonist and an inhibitor of mitochondrial Mg<sup>2+</sup> ATPases. Nevertheless, the cellular hepatic effects of these 2 compounds have been poorly described. Therefore, we applied an original approach coupling metabolomics and proteomics to investigate in vitro toxicity responses of CLD and CLD-OH in liver cell models (HepaRG cell line and primary human hepatocytes, PHHs).

A concentration-response study after a 24h treatment and a kinetic-response study between 3 and 48h have been performed in differentiated HepaRG cells exposed to selected concentrations of CLD (1, 5 µM) and CLD-OH (1, 5, 10 µM). As well, PHHs were exposed to CLD (5 µM) and CLD-OH (10 µM) during 48 hours. Cell culture media and cell lysates were analysed using metabolomics and proteomics approaches [4].

HepaRG cells and PHHs were more sensitive to CLD than CLD-OH. Metabolomics showed that CLD was very little metabolized into a glucuronide form unlike CLDOH. This metabolite was thus more excreted and CLD more sequestered in hepatocytes. Perturbations of mitochondrial activities were observed by proteomics and ATP levels were slightly decreased similarly for both compounds in HepaRG cells but restored at 48 hours indicating the existence of a toxicity mechanism independent of ATPases inhibition. The main finding was a decrease of septin abundances by CLD treatment, in both HepaRG cells and PHHs, that was not observed with CLD-OH. These key cytoskeleton-proteins are implicated in numerous cellular functions, such as intracellular vesicular trafficking, and modification of their expressions is associated with human liver diseases. Docking investigation pointed out that CLD and CLD-OH show a good binding affinity for septins. Moreover, we demonstrated that septins were degraded by the proteasome. Besides, CLD induced the disturbance of lipid-droplet localization and bile acid secretion and favored long chain dicarboxylic fatty acids production, key events in NAFLD progression. In conclusion, our original study highlighted that CLD exposure promote hepatotoxicity and steatohepatitis by targeting septins and that these effects can be significantly overcome when CLD is converted to CLD-OH.

### Références:

Resiere, D., et al. (2023). Chlordecone (Kepone) poisoning in the French Territories in the Americas. *Lancet* 401, 916.

Multigner L., et al. (2010). Chlordecone exposure and risk of prostate cancer. *J Clin Oncol* 28, 3457-3462.

Tabet E., et al. (2016). Chlordecone potentiates hepatic fibrosis in chronic liver injury induced by carbon tetrachloride in mice. *Toxicol Lett* 255, 1-10.

Leger T, et al. (2023). Fate and PPARgamma and STATs-driven effects of the mitochondrial complex I inhibitor tebufenpyrad in liver cells revealed with multi-omics. *J Hazard Mater* 442, 130083.

## Proteomic approach to characterize and target MINK1-PRICKLE1 signaling pathway in triple negative breast cancer

Dr Avois DAULAT<sup>a</sup>, Giulia DELACOURT<sup>a</sup>, Jean LY<sup>a</sup>, Stéphane AUDEBERT<sup>a</sup>, Luc CAMOIN<sup>a</sup>, Patrice DUBREUIL<sup>a</sup>, Jean-Paul BORG<sup>a</sup>

<sup>a</sup> CRCM

Triple negative breast cancer (TNBC) is the most aggressive cancer among breast cancer. It is characterized by molecular heterogeneity, resistance to standard treatment and lack of available targeted therapies. Upregulation of Wnt/PCP components is observed in many cancers and is associated with cancer development at early and late stages. We recently showed that PRICKLE1, a core Wnt/PCP component, is overexpressed in triple negative breast cancer (TNBC) and associated with poor prognosis. PRICKLE1 is a scaffold protein. MINK1 is a serine/threonine kinase which phosphorylates PRICKLE1. Using interactomic approach, we identified several PRICKLE1 associated proteins. We then cross-analyzed our dataset with data generated from a phosphoproteomic approach designed to identify putative MINK1 substrates. Among the putative substrates, we characterized LL5 $\beta$  which is required for focal adhesion disassembly and cell migration. LL5 $\beta$  is associated with PRICKLE1. At the molecular level, we demonstrated that PRICKLE1 is a kinase substrate adaptor and MINK1 phosphorylates sequentially PRICKLE1 and LL5 $\beta$ . Analysis of gene expression database showed that the concomitant up-regulation of PRICKLE1 and LL5 $\beta$  mRNA levels encoding MINK1 substrates is associated with a poor metastasis-free survival for TNBC patients. Altogether, our results suggest that MINK1 may represent a potential target in TNBC. We are currently characterizing a potential MINK1 inhibitor and developing antibodies targeting identified MINK1 phosphorylation site to generate an acute tool to measure MINK1 activity *in cellulo* and *in vivo*.

## Nanobody-based APMS analysis of endogenous metabotropic glutamate receptor 2 interactome from mice brain

Dr Franck VANDERMOERE<sup>a</sup>, Clémentine Eva PHILIBERT<sup>a</sup>, Candice DISDIER<sup>a</sup>, Pierre-André LAFON<sup>a</sup>, Mathieu OOSTERLAKEN<sup>a</sup>, Anne PIZZOCCARO<sup>a</sup>, Julie KNIAZEFF<sup>a</sup>, Jean-Philippe PIN<sup>a</sup>, Philippe RONDARD<sup>a</sup>, Philippe MARIN<sup>a</sup>

<sup>a</sup> Institut de Génomique Fonctionnelle, Univ Montpellier, INSERM, CNRS

Nanobodies are Lama single chain antibodies whose small size is well-suited for accessing small and hydrophobic pockets in proteins. This technology is growing fast in the field of seven-transmembrane domain receptors (GPCRs), the largest family of proteins expressed at the cell surface targeted by nearly 40% of prescribed drugs. Type 2 metabotropic glutamate receptors (mGlu2) are GPCRs targeted by a new generation of antipsychotics that show unprecedented efficacy against negative and cognitive symptoms of schizophrenia. Despite a clear advantage on actual antipsychotics, signalling triggered by mGlu2 activators in the brain remains poorly characterized. That is the reason why we performed an affinity purification mass spectrometry (APMS) analysis of the interactome of the endogenous mGlu2 in the prefrontal cortex of Wild-type mouse. This APMS strategy took advantage of an in-house developed nanobody highly specific for mGlu2 with sub nanomolar affinity.

From 3 independent experiments on three different mice cortex, statistical analysis of MS1 label free quantitative data showed a significant enrichment of 149 proteins in pull down with mGlu2 nanobody vs control pull down with non-relevant nanobody. String analysis of the list of potential interactors highlighted a highly connected network and an enrichment in gene ontology terms such as "glutamate receptor signalling pathway", "adenylate cyclase-inhibiting G protein-coupled glutamate receptor signalling pathway", "regulation of synaptic transmission, glutamatergic" or "presynaptic active zone".

One of the candidates particularly attracted our attention as it is a receptor well-known to regulate the physiological mechanisms affected during negative and cognitive symptoms of schizophrenia. Proximity ligation assays (PLA) and Bioluminescence Resonance Energy Transfer (BRET) assays confirmed a close (if not direct) and specific interaction between the two receptors in vivo, in primary culture of neurons and in transfected cells. Activity readouts indicated a reciprocal positive cross talk between the two receptors. Finally, behavioural studies on a preclinical mice model for schizophrenia indicated that antagonist of the interacting receptor strongly inhibits the effect of glutamatergic antipsychotics in tasks related to negative and cognitive symptoms.

In conclusion, the ability of nanobodies to recognize small and hydrophobic epitopes invisible to classical antibodies is an important advantage to target GPCRs. They can be used in affinity purification of endogenous GPCRs at a scale compatible with mass spectrometry analysis of in vivo interactome. Using this strategy, we identified a large number of mGlu2 receptor interacting proteins that are relevant to the brain and that might be key to the activity of next generation antipsychotics.



## Parallel accumulation - serial fragmentation combined with data-independent acquisition for the successful identification of plasma proteins correlated with entry into sexual maturation in the rainbow trout

Régis LAVIGNE<sup>a</sup>, Aurélien BRIONNE<sup>b</sup>, Emmanuelle COM<sup>a</sup>, Charline DANTHU<sup>b</sup>, Blandine GUÉVEL<sup>a</sup>, Jean-Jacques LAREYRE<sup>b</sup>, Charles PINEAU<sup>a</sup>

<sup>a</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35042 Rennes cedex, France; Univ Rennes, CNRS, Inserm, Biosit UAR 3480 US\_S 018, Protim core facility, F-35000 Rennes, France, <sup>b</sup> INRAE LPGP (Laboratoire de Physiologie et Génomique des Poissons), F-35042 Rennes cedex, France

In fish farms, precocious sexual maturation can occur in one- or two-year-old male rainbow trout much before fish reach commercial weight that is otherwise attained at 3-year-old. Unfortunately, sexual maturation in salmonids reduces growth, alters flesh quality and increases animal susceptibility to diseases. The present study aimed to identify plasma proteins correlated with sexual maturation that could be used to develop sterile fish and/or non-invasive screening tests for the detection of early puberty. We combine Parallel Accumulation Serial Fragmentation with a DIA approach (diaPASEF) for an in-depth identification and quantification of plasma proteins whose expression profile is correlated with sexual maturation in rainbow trout.

We compared in-solution and stacking gel protein digestion with trypsin/LysC, but also the performance of different diaPASEF methods, used to identify and quantify digest of HeLa, and individual trout plasma protein samples. For HeLa proteins digest, we generated a spectral library from a previously published dataset (Meier et al., Nat Methods 2020: 24 fractions). For the trout protein samples, we generated a spectral library from 4 pools of individual trout plasma based on the different population types (immature females, immature males, sexually maturing males and spermiating males) and separated on a 1D gel electrophoresis. The separated protein lanes (80 different fractions) were acquired with DDA-PASEF technology. All DDA and diaPASEF analysis were performed on a nanoElite liquid chromatography system, and a timsTOF Pro mass spectrometer (Bruker Daltonics). Instrument's methods were adapted to perform data-independent isolation of multiple precursor windows within a single 100ms TIMS separation. We tested multiple schemes for precursor selection window size and placement in the m/z-ion mobility plane. Data were analyzed using a library-based workflow with Spectronaut 16 software (Biognosys) compared to a DIA direct approach. 79 samples were then successfully analyzed requiring close monitoring using quality controls and iRTs.

Using a simple stacking gel protein digestion and a diaPASEF method defined with 32 DIA windows (ion mobility plane with narrow 25 m/z precursor windows) and two windows in each 100ms diaPASEF, we could identify 37168 peptides which segregated into 4096 protein groups (PG). 75% of those PG were encoded by a single gene. We completed the annotation of the PG by searching annotated orthologous genes in other fish species. This approach enabled us to increase the number of annotated PG from 2380 to 3572 out of 4096 PG. Finally, we identified 681 differentially expressed PG out of 4096, in at least one of the tested conditions.

We detected a large number of proteins whose relative abundance varies during the different stages of sexual maturation of males or females. Analysis of GO terms in males highlighted significant changes occurring at puberty and modifying the relative abundance of certain proteins previously implicated in reproductive function, muscle and skeletal growth, humoral and cellular immunity, renal function, and vision. PG differentially expressed between sexually immature males and females were also identified. Significant differences were observed with GO terms related to alcohol metabolism, vitamin metabolism, vision, renal function and differentiation of primary sexual traits, in particular gonadal sex differentiation.

Based on previously released transcriptomic data, we identified PG showing gonad-specific gene expression profiles. The biological relevance of these plasma proteins which likely originate from the gonads will be presented thanks to a tissue-specificity index to identify ubiquitously or specifically expressed genes in male or female gonads.

In conclusion the power of the diaPASEF approach has allowed the identification of:

- new factors of gonadal origin in plasma with hormonal activity,
- candidate plasma markers for sex genotyping,
- candidate plasma markers for the detection of puberty onset in most farmed salmonids.





## Pitfalls of gene-level annotation applied to functional analysis of protein groups

Dr Marie LOCARD-PAULET<sup>a</sup>, Nadezhda T. DONCHEVA<sup>b</sup>, Lars Juhl JENSEN<sup>b</sup>

<sup>a</sup> Institut de Pharmacologie et de Biologie Structurale (IPBS), Université de Toulouse, CNRS, Université Toulouse III—Paul Sabatier (UPS), <sup>b</sup> Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

Nowadays, protein relative quantities can routinely be measured across samples using high-throughput mass spectrometry (MS). Classically, proteins are digested into peptides before MS analysis, so the set of quantified peptides may not include all necessary gene-specific peptides to report all single-gene products. This is why MS relative quantities are reported per protein groups that gather all the gene products that cannot be distinguished based on the set of detected peptides in a given experiment. Yet, functional annotation is performed at gene level and functional analysis requires single-gene information. The presence of multiple gene products in a single protein group should then be handled carefully during data analysis. However, there is no agreement or best practice on how to deal with protein groups for functional enrichment, protein-protein interaction networks or knowledge graphs.

Here, we investigate the composition of the protein groups identified in 14 published proteomics data sets resulting from different types of proteomics data, including deep proteomes, phosphoproteomics, single-cell analysis and pull downs from different species. Sensitivity strongly impacts protein group composition: more identified peptides lead to more gene-specific peptides, and thus more single-gene protein groups. Thus, experiments with limited input material, such as single-cell proteomics, result in a higher proportion of multiple-gene protein groups. Another factor is the set of sequences (*i. e.* the proteome) used for searching MS data: a well curated proteome such as the UniProt/SwissProt human proteome can result in a high proportion of single-gene protein groups, whereas less-well studied proteomes containing protein accessions with similar sequences result in larger protein groups (multiple genes per group).

We show that functional analysis such as Gene Ontology (GO) enrichment and network analysis can be affected by which single gene per group is selected for downstream analysis, and that this gene selection should not be overlooked. We discuss several strategies of GO-term enrichment and explore new avenues for network analysis that would take into account all the genes contained in protein groups of MS-based data sets.

## Proteome profiling of tissue-derived extracellular vesicles and soluble secretome reveals non-invasive candidate markers involved in colorectal cancer proliferation and progression

Madalena MONTEIRO<sup>a</sup>, Catarina ALMEIDA-MARQUES<sup>a</sup>, Logan BISHOP-CURREY<sup>a</sup>, Valerie DUSSELSORP<sup>a</sup>, Beatriz ESCUDERO<sup>a</sup>, Sander PIERSMA<sup>a</sup>, Thang PHAM<sup>a</sup>, Jaco KNOL<sup>a</sup>, Hanieh SADEGHI<sup>a</sup>, Gerrit MEIJER<sup>b</sup>, Remond FIJNEMAN<sup>b</sup>, Meike DE WIT<sup>b</sup>, Irene BIJNSDORP<sup>c</sup>, Connie JIMENEZ<sup>a</sup>

<sup>a</sup> Department of Medical Oncology, Amsterdam University Medical Center, Amsterdam, Netherlands, <sup>b</sup> Department of Pathology, The Netherlands Cancer Institute, Amsterdam, Netherlands, <sup>c</sup> Department of Urology, Amsterdam University Medical Center, Amsterdam, Netherlands

Secreted proteins significantly impact cancer progression by increasing proliferation, activating angiogenesis, facilitating invasion, enabling metastasis, avoiding immune destruction, and sustaining inflammation. Proteins in the secretome are either secreted via the classical secretory pathway, which requires an N-terminal signal peptide, or through the unconventional pathways involving vesicles or membrane transporters and channels. Both soluble secretome (SS) and extracellular vesicles (EV) are detectable in biofluids, thus being of high interest for non-invasive insights into tumor biology and biomarker discovery. Profiling the protein components of the cancer tissue secretome in an unbiased manner by mass spectrometry is expected to further our understanding of how cancer cells alter the tumor microenvironment. So far most studies have not addressed secretome proteomics using patient tissues.

Tumor and adjacent normal tissue secretome were collected *ex vivo* from 17 patients with colorectal cancer (CRC) and 4 patients with adenoma. Mass spectrometry-based proteomics was performed on tissue, SS and EV fractions isolated using Vn96 peptide followed by protein fractionation by gel electrophoresis and in-depth nanoLC-MS/MS analysis (five 90-min runs/ sample) on a QExactive HF instrument. Statistics was performed using beta binomial test and Spearman correlation. Biological annotations were performed using single sample Gene Set Enrichment Analysis (ssGSEA), Gene Ontology mining, String network connectivity and Cytoscape, and OutCyte database.

The multi-compartment CRC dataset comprised 3745 tissue, 3353 EV and 2238 SS proteins. Unsupervised clustering (almost) completely separated tumour and normal samples in all 3 compartments, emphasising complementary views of cancer phenotype. Analysis using ssGSEA revealed distinct enrichment patterns of cancer hallmarks in each fraction, but a concerted dysregulation in cancer. Specifically, proliferation and DNA damage were significantly up-regulated in tumour samples, while processes related to metabolism, immune, and development were down-regulated. Correlation of tissue protein expression to EV or SS abundance and hierarchical clustering revealed separation between tumor and control groups with tissue-EV correlated proteins related to DNA replication, transcription, extracellular matrix, and membrane trafficking while tissue-SS proteins were related to extracellular matrix and iron transport and storage. Comparative analysis revealed wide-spread regulation of protein abundance ( $p > 0.05$ ;  $|FC| > 1.5$ ; 35% of Tissue, 53% of EV, and 69% of SS proteomes). Interestingly, proteins up-regulated in cancer were predominantly annotated as unconventionally secreted, and down-regulated proteins predominantly annotated as classically secreted. These cancer-associated EV and SS proteins have functions in protein translation, DNA replication, stress granule assembly, unfolded protein response, and cell adhesion. Finally, using stringent filters ( $p < 0.01$ ;  $|FC| > 3$ ) highly discriminatory EV and SS proteins with a trend profile selected as proteins of interest for non-invasive diagnostic applications. A subset was validated in a published genetic mouse model secretome and clinical tissue proteome datasets. Validated proteins with ELISA assays will be followed-up in plasma of patients with different stages of CRC and controls.

Altogether, we report a unique multi-compartment clinical proteomics dataset of CRC and matched normal tissue and secreted proteins. We found concerted dysregulation of cancer hallmarks with distinct contributions from specific proteins and signalling pathways unique to tissue, EVs, or SS and identified novel candidate non-invasive CRC biomarkers for further development.

## Proteomics characterization of neurons upon oxygen and glucose deprivation – neuronal modulation induced by the secretome of physiologically primed MSCs

Inês CAMELO<sup>a</sup>, Vera M. MENDES<sup>b</sup>, Ivan L. SALAZAR<sup>a</sup>, Catarina DOMINGUES<sup>a</sup>, Ana Carolina SANTO MENDES<sup>c</sup>, Margarida GERALDO<sup>b</sup>, Sandra Isabel ANJO<sup>b</sup>, Carla CARDOSO<sup>d</sup>, Alexandra DINIS<sup>e</sup>, Carlos B. DUARTE<sup>f</sup>, Mário GRÃOS<sup>g</sup>, Bruno MANADAS<sup>h</sup>

<sup>a</sup> Center for Neuroscience and Cell Biology - University of Coimbra; Institute for Interdisciplinary Research, University of Coimbra (IIIUC), <sup>b</sup> Center for Neuroscience and Cell Biology - University of Coimbra, <sup>c</sup> Center for Neuroscience and Cell Biology - University of Coimbra; Biocant, Technology Transfer Association, 3060-197 Cantanhede, Portugal, <sup>d</sup> Crioestaminal, <sup>e</sup> Pediatric Intensive Care Unit, Hospital Pediátrico, Centro Hospitalar e Universitário de Coimbra, <sup>f</sup> Faculty of Sciences and Technology, Department of Life Sciences, University of Coimbra, <sup>g</sup> Biocant, Technology Transfer Association, <sup>h</sup> Center for Neuroscience and Cell Biology, University of Coimbra; Institute for Interdisciplinary Research, University of Coimbra (IIIUC)

There has been a growing interest in the therapeutic value of mesenchymal stem cells (MSCs) in recent decades, due to their capability to migrate to the site of injury, replace damaged cells and/or act through paracrine signaling on neighboring cells. Therefore, the study of the secretome of these cells is gaining importance, as its characterization can reveal signaling pathways triggered to promote tissue regeneration. However, a “single dose” of secretome requires extensive expansion *in vitro*, which might compromise their multipotency and therapeutic properties. Therefore, we hypothesized that mimicking the physiological environment of MSCs in culture could modulate the proteome and secretome of these cells. For this purpose, umbilical-cord MSCs were maintained at physiological conditions for 48h or 7-10 days, controlling the stiffness ( $\approx 3\text{kPa}$ ), or oxygen levels ( $\approx 5\%O_2$ ), similar to what they sense in the tissue. First, a proteomic screening of the cellular proteome was performed by comparing these conditions to “standard culture” (plastic ( $\approx 1\text{GPa}$ ) and  $\approx 18\%O_2$ ). Proteins were quantified with DIA (SWATH-MS) and analyzed using multivariate approaches (PLS-DA). Results indicated that more than one-hundred-fifty proteins were modulated by priming the MSCs at physiological conditions for 48h. Interestingly, a gene ontology analysis revealed that both mechanic and oxygen levels stimuli were modulating similar processes, such as protein metabolism, energy pathways, and cellular components, such as ribosomes or mitochondria. So, we hypothesized that the secretome of primed MSCs, and consequently their therapeutic potential, would also be altered upon priming at physiological conditions. Our team established an *in vitro* model of hypoxic-ischemic encephalopathy, a brain injury resulting from a peripartum asphyxia event in the newborn that might result in profound neurophysiological sequelae or even death. Rat cortical neurons were isolated from embryos and were subjected to oxygen and glucose deprivation stimuli (OGD) at DIV8. Then, neurons were incubated with the secretome of standard and primed-MSCs, and the neuronal proteome was quantified using DIA (SWATH-MS). Proteins of interest were selected through univariate and multivariate analysis. More than 200 proteins were identified as significantly different on pos-OGD neurons cultured in the presence of MSCs secretome, compared to not rescued neurons. These proteins are mostly associated with ribosomes and involved in translation processes, such as L13a-mediated translational silencing of ceruloplasmin. In addition, it was found that the oxygen consumption ratio of these neurons varies differently when exposed to primed secretomes. Interestingly, more than 150 proteins were found to be commonly modulated by the three secretomes and are mainly involved in microtubule organization and axiogenesis. In conclusion, our data has identified important pathways for neuronal survival following an OGD event, that could be used as a potential drug target. However, further research is required to elucidate the specific mechanisms triggered by primed secretome and assess other potential applications based on their properties.

This work was financed by the European Regional Development Fund (ERDF) through the COMPETE 2020 - Operational Programme for Competitiveness and Internationalization and Portuguese national funds via FCT, under projects POCI-01-0145-FEDER-029311, POCI-01-0247-FEDER-045311, UIDB/04539/2020, UIDP/04539/2020 and LA/P/0058/2020, POCI-01-0145-FEDER-402-022125 (ref.: ROTEIRO/0028/2013), and individual Ph.D. fellowship SFRH/BD/143442/2019 (Ines Caramelo).

## MACHINE learning approach for the stratification of rheumatoid arthritis patients based on circulating proteomic signatures associated with their synovial pathotype

Florencia Cristina PICCHI FIGUEIRA<sup>a</sup>, Rocío PAZ-GONZÁLEZ<sup>a</sup>, Lucía LOURIDO<sup>a</sup>, Valentina CALAMIA<sup>a</sup>, Patricia FERNÁNDEZ-PUENTE<sup>b</sup>, Rebecca HANDS<sup>c</sup>, Patricia QUARANTA<sup>a</sup>, Costantino PITZALIS<sup>c</sup>, Francisco J. BLANCO<sup>d</sup>, Cristina RUIZ-ROMERO<sup>e</sup>

<sup>a</sup> Unidad de Proteómica, Grupo de Investigación de Reumatología (GIR), Instituto de Investigación Biomédica de A Coruña (INIBIC), Complejo Hospitalario Universitario de A Coruña (CHUAC), Sergas, A Coruña, Spain., <sup>b</sup> Centro Interdisciplinar de Química y Biología (CICA), Universidade da Coruña, Spain., <sup>c</sup> Centre for Experimental Medicine & Rheumatology, William Harvey Research Institute, Queen Mary University of London, London, UK., <sup>d</sup> Grupo de Investigación de Reumatología y Salud (GIR-S), Departamento de Fisioterapia, Medicina y Ciencias Biomédicas, Centro de investigaciones Avanzadas (CICA), Universidade da Coruña (UDC), A Coruña, Spain., <sup>e</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBERBBN), Madrid, Spain

**Background:** Rheumatoid arthritis (RA) is a complex autoimmune disorder characterized by considerable clinical variability. This disease affects millions of people worldwide. Patient stratification is essential for

developing personalized treatments and improving patient outcomes and reducing healthcare costs, specific pathological

phenotypes of synovial tissue (pathotypes) have emerged as potential correlates of diverse clinical progression and therapeutic response. These pathotypes - lymphoid (L), myeloid (M) and fibroid (F) - reflect the different immune cell infiltrates and molecular mechanisms underlying the inflammation and joint damage in RA. In this study, we aimed to detect signatures of circulating proteins that correspond with synovial pathotypes in RA patients, which could be useful as non-invasive biomarkers for patient stratification and the optimization of treatments.

**Methods:** We conducted a comprehensive proteomic analysis on samples from patients participating in the Pathobiology of Early Arthritis Cohort (PEAC), an ongoing observational study designed to uncover the molecular basis of early RA and identify predictive biomarkers for disease outcomes. Synovial biopsies guided by ultrasonography from these patients allowed their classification based on their pathotype. Serum samples were depleted of the 14 most abundant proteins and analyzed via reversed-phase LC-MS/MS employing a SWATH strategy in a tripleTOF MS system (Sciex). The acquired quantitative data were processed using ProteinPilot 5.0.1 and PeakView 2.1 software (Sciex). We used a machine learning workflow (FactorMiner, Factorextra, Caret, e1071 and VIP-Clustering from R software version 4.2.2) for the development and validation of the classifier.

**Results:** The discovery study involved 54 baseline serum samples, corresponding to 22 L patients, 18 M, and 14 F. The proteomic analysis of these samples resulted in the identification and quantification of 229 proteins common between the screening (30 samples) and verification (24 samples) sets. Data underwent PCA pre-processing for dimension reduction, and then they were analyzed using machine learning techniques. In a Divergent Test Model Evaluation (DTME) phase, support vector machines (SVM) demonstrated better performance than other algorithms tested. The SVM model with radial kernel function was used to transform data into a higher-dimensional space, where a hyperplane was determined for the best separation of the different pathotypes. The overall model performance demonstrated 97% accuracy and a substantial Kappa arrangement (0.95), indicating high reliability and reproducibility. This protein signature was subsequently assessed in an external new cohort of 57 patients to validate the training model. Classifier metrics had an overall sensitivity of 90% and a specificity of 85%. Additionally, the VIP-Clustering technique was employed to identify the most important variables and reveal the underlying structure of the dataset. This approach led to the characterization of a proteomic signature of the 3-clinical clusters, based on 11 proteins, and their similarities: Myeloid-Lymphoid vs Fibroid.

**Conclusion:** We successfully identified and validated a circulating proteins signature that is associated with synovial pathotypes in early RA patients. The possibility of reducing the panel to 3-core proteins aims to apply it in clinical trials and personalized medicine. The integration of the protein signature with other 'omics' data could offer a more comprehensive understanding of RA.



SEProt  
Sociedad Española  
de Proteómica

Gomez, E. A., Colas, R. A., Souza, P. R., Hands, R., Lewis, M. J., Bessant, C., Pitzalis, C., & Dalli, J. (2020). Blood pro-resolving mediators are linked with synovial pathology and are predictive of DMARD responsiveness in rheumatoid arthritis. *Nature Communications*, 11(1), 5420. <https://doi.org/10.1038/s41467-020-19176-z>

Nerviani, A., Di Cicco, M., Mahto, A., Lliso-Ribera, G., Rivellese, F., Thorborn, G., Hands, R., Bellan, M., Mauro, D., Boutet, M.-A., Giorli, G., Lewis, M., Kelly, S., Bombardieri, M., Humby, F., & Pitzalis, C. (2020). A pauci-immune synovial pathotype predicts inadequate response to TNF $\alpha$ -blockade in rheumatoid arthritis patients. *Frontiers in Immunology*, 11, 845. <https://doi.org/10.3389/fimmu.2020.00845>



## Pushing DIA proteomics analyses of neat plasma to 1000 protein groups ID/h

Ines METATLA<sup>a</sup>, Kevin ROGER<sup>a</sup>, Sara CECCACCI<sup>a</sup>, Cerina CHHUON<sup>a</sup>, Maria DE MORAES<sup>b</sup>, Anne JAMET<sup>c</sup>, Pierre-Olivier SCHMIT<sup>d</sup>, Manuel CHAPELLE<sup>d</sup>, Ida Chiara GUERRERA<sup>a</sup>

<sup>a</sup> Proteomics platform Necker, Université Paris Cité - Structure Fédérative de Recherche Necker, INSERM US24/CNRS UAR3633, Paris 75015, France., <sup>b</sup> Université Paris Cité, Inserm UMR1151, CNRS UMR8253, Institut Necker Enfants Malades, Team Immunoregulation and Immunopathology, Paris, France, <sup>c</sup> Université Paris Cité, Inserm UMR-S1151, CNRS UMR-S8253, Institut Necker Enfants Malades, Team Pathogenesis of systemic infections. F75015 Paris, France., <sup>d</sup> Bruker Daltonique S.A., Wissembourg, Grand Est, France

### Introduction

Direct proteomics analysis of neat plasma remains a challenge because of the huge dynamic range of the plasma proteins, but is also very appealing because it requires very small volumes and is unexpensive compared to new depletions technology. We focused on workflow optimization of neat plasma analysis using dia-PASEF<sup>®</sup> approach to maximize the number of proteins groups (PG) identified and quantified, while minimizing both gradient time and missing value levels. Here we describe the comparisons of different LC settings, and the optimization of data analysis using DiaNN, PaSER and Spectronaut<sup>®</sup> 17.

We finally report a reference “deep workflow” to quantify up to 990 proteins in less than 1h and a “fast workflow” leading to 425 proteins using a 21 min gradient.

### Methods

The plasma sample cohort consisted in 15 patients affected by a rare genetic disease and 18 age-matched controls. Samples were digested with trypsin using STRAP columns, separated either by nano-HPLC (nanoElute, Bruker Daltonics) using a 25cm column on a 30min gradient (IonOpticks, Australia), or with an EvosepOne ran with the 60 Samples Per Day (60SPD) method (Endurance column, 8cm). Both LC's were connected to a timsTOF Pro<sup>™</sup> instrument (Bruker Daltonics) via its Captive Spray source. The timsTOF Pro was in dia-PASEF acquisition mode: a 19-windows method using 33da windows covering a 0.65-1.3 V/cm<sup>2</sup> mobility ranges over a 400-1050m/z range. DiaNN1.8.1 (Demichev group), a GPU-based version of diaNN (TIMS-DIA-NN on PaSER, Bruker) and Spectronaut<sup>®</sup>17 (Biognosys) were used for data processing.

### Results

To understand the key factors in obtaining higher number of proteins we evaluated two LC systems and three DIA library-free data analysis workflows. We found that the number of identifications correlated well with the number of samples simultaneously searched. Using the nanoElute, a single plasma analysis leads to the identification of ca450 proteins groups, the analysis of 4 samples (enabling MBR), leads to 507<sup>+6</sup> proteins (2% missing values, MV hereafter). Searching 33 samples simultaneously led to 811 <sup>+78</sup> protein groups (33% MV) and up to 978 <sup>+140</sup> (58% MV) if depleted plasma files from independent samples were also added to the analysis. We show a significant effect of the MBR step, as the same analyses run without MBR do not lead to the same improvement in number of IDs (20% increase instead of 70% for the 33 plasma search). Analysis on PaSER lead to 376 <sup>+102</sup> (with 45% MV) on 33 runs. Using Spectraunaut<sup>®</sup>17, we obtained around 460 proteins on single files, 537<sup>+4</sup> (with 5% MV) for 4 plasma search, and 731<sup>+50</sup> (23% MV) on 33 runs.

The analysis of the same cohort using 60SPD EvosepOne methods was 2.7x faster than nanoElute, but led to 2x lower number of IDs using both DiaNN and Spectronaut<sup>®</sup> 17. We obtained 394<sup>+37</sup> (27% MV) and 426<sup>+29</sup> (with 20% MV) on the 33 samples analysis using DiaNN and Spectronaut<sup>®</sup>17 respectively.

We will discuss the trade-off of analysis depth and quantification quality for time, for the tested softwares in both library-free and library-based mode, and will also present the outcome of further optimization of MS methods and LC chromatographic setup.



## Statistical inference of protein turnover

Dr. Tsung-Heng TSAI<sup>a</sup>

<sup>a</sup> KENT STATE UNIVERSITY

Bottom-up proteomics with liquid chromatography coupled with mass spectrometry (LC-MS) has been widely used to quantify changes in protein abundances, structures, and modifications across conditions in biological and clinical research. Besides the snapshot of protein information, LC-MS analysis can be combined with stable isotope labeling (e.g., with a deuterated water tracer) to quantify protein turnover, where the kinetics is modeled with a time series of the incorporation of isotopes into proteins. Accurate quantification of protein turnover allows unbiased interrogation of global changes in proteome dynamics, which provides a comprehensive understanding of the proteome and offers new insights into biomedical research.

The output of bottom-up proteomics experiments is a list of quantified peptide ions or their fragments. Using this indirect evidence to derive protein-level conclusions can be challenging. Turnover rate analysis in proteome dynamics studies is further complicated by the additional dimension of time, which necessitates robust kinetic modeling. Among the goals of proteome dynamics are (1) estimation of turnover rate in one condition, and (2) detection of changes in turnover rate between conditions. Conventional methods fit a one-compartment model to the time-course data of each peptide separately. The peptide turnovers are then averaged to obtain protein-level turnover, and in difference detection, provided as input to a statistical test (e.g., two-sample t-test) to detect changes in turnover rate. While simple, this approach does not fully account for the sources of variations, and it requires a sufficient number of peptides whose turnover rates can be properly quantified.

We propose a statistical approach, which addresses the two goals through statistical modeling and model-based conclusion of kinetic data. In consideration of unique properties of distinct proteoforms, data for native and post-translationally modified proteins are analyzed separately. The turnover rates of native peptides are viewed as repeated measures of the turnover rate of their corresponding protein. Thus, the time-course data of native peptides corresponding to the same native protein are modeled jointly to estimate the turnover rate of the protein. The total labeling data are modeled by a nonlinear regression model representing the one-compartment model for each protein and each condition, while considering the peptide-to-peptide variation for the baseline and plateau enrichments. The nonlinear regression model is fitted to all the time-course data for the protein, which produced the coefficient estimate and standard error of protein turnover rate, among other parameters. To assess the changes in turnover rate across conditions, an F-test was performed, where the full model with separate parameters for each condition is compared to a reduced model under the null hypothesis, with an identical turnover rate shared by both conditions. For post-translationally modified proteins, estimation of turnover rate and test for change between conditions are performed in a site-specific manner. Turnover rate for each modified site is estimated and tested similarly as for native protein, but only the modified peptides covering the same modification site are analyzed and modeled together.

We evaluated the proposed approach by comparing it to the results of the conventional t-test using computer simulations and biological investigations. Compared to the t-test, the proposed approach improved the accuracy and the efficiency of turnover rate estimation, and improved statistical power of detecting changes in turnover rate. The proposed method reduced the negative impact of missing values and outlying measurements and improved the efficiency of turnover rate estimates in such circumstances. The improvements were particularly notable for proteins or modified sites with fewer quantified peptides. This general statistical modeling framework can be extended to handle other statistical tasks in kinetic analysis.

## Multi-OMICS integration analysis for assessing human stem cell models of human peri-implantation development

Constance ONFRAY<sup>a</sup>, Simon CHEVOLLEAU<sup>b</sup>, Eva MOINARD<sup>b</sup>, Kasturi MAHADIK<sup>c</sup>, Océane GIRARD<sup>b</sup>, Régis LAVIGNE<sup>d</sup>, Ryan ALLSOP<sup>e</sup>, Grigorios GEORGIOPOULOS<sup>e</sup>, Ophélie RENOULT<sup>f</sup>, Irene AKSOY<sup>g</sup>, Jean-François OUIMETTE<sup>c</sup>, Charles PINEAU<sup>d</sup>, Thomas FREOUR<sup>h</sup>, Claire PECQUEUR<sup>f</sup>, Vincent PASQUE<sup>e</sup>, Claire ROUGEULLE<sup>c</sup>, Laurent DAVID<sup>i</sup>

<sup>a</sup> Nantes Université, CHU Nantes, Inserm, CR2TI, <sup>b</sup> Nantes Université, CHU Nantes, Inserm, CR2TI, F-44000 Nantes, <sup>c</sup> Université Paris Cité, CNRS, Epigenetics and Cell Fate, F-75013 Paris, France, <sup>d</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail) - UMR\_S 1085, F-35000 Rennes, France; Univ Rennes, CNRS, Inserm, Biosit UAR 3480 US\_S 018, Protim Core Facility, F-35000 Rennes, France, <sup>e</sup> KU Leuven - University of Leuven, Department of Development and Regeneration, Leuven; Institute for Single Cell Omics and Leuven Stem Cell Institute, Herestraat 49, B-3000 Leuven, Belgium, <sup>f</sup> Nantes Université, CNRS, Inserm, CRCI2NA, F-44000, Nantes, France, <sup>g</sup> Univ Lyon, Université Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, 69500 Bron, France, <sup>h</sup> Nantes Université, CHU Nantes, Inserm, CR2TI, F-44000 Nantes, France, <sup>i</sup> Nantes Université, CHU Nantes, Inserm, CR2TI, F-44000 Nantes, France; Nantes Université, CHU Nantes, Inserm, CNRS, BioCore, F-44000, Nantes, France

Understanding human peri-implantation development is essential to appreciate the onset of human life but also for multiple clinical applications such as in vitro fertilization, regenerative medicine and placenta physiopathology.

To overcome the technical challenges and ethical issues that result in limited access to human embryo for research, we studied stem cells lines modelling human embryo development. While opening new horizons for further research, these models require proper benchmarking in order to define their best applications and achieve their full potential. Therefore, we propose a proteomics characterization of pluripotent and trophoblastic stem cell models with a DIA approach using a hybrid library to complete and refine transcriptomic, epigenetic and metabolic approaches. Proteomics coupled with transcriptomic analyses allowed us to compare pluripotent (hNPSCs, hPPSCs, hEPSCs) and trophoblastic (hTSCs) stem cell models. Subsequently, our strategy, once combined with metabolomics could distinguish hEPSCs from hPPSCs modelling the post-implantation epiblast that gives rise to the embryo and confirmed the high metabolic activity in trophoblast stem cells that modelling the post-implantation cytotrophoblast giving rise to the placenta.

Altogether, our work on OMICS contributes to define hallmarks for pluripotent and trophoblast stem cells that are particularly important, when reliable, to ensure proper models are used to tackle hypohesis.

**Keywords:** DIA analysis, peri-implantation development, cell fate, stem cells



## Boosting immunopeptidomics identification using an optimized workflow

Aurélie HIRSCHLER<sup>a</sup>, Perrine SPINNHIRNY<sup>b</sup>, Arthur DECLERCQ<sup>c</sup>, Jeewan Babu RIJAL<sup>a</sup>, Nicodème PAUL<sup>b</sup>, Seiamak BAHRAM<sup>b</sup>, Lennart MARTENS<sup>c</sup>, Raphael CARAPITO<sup>b</sup>, Christine CARAPITO<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse Bio-Organique (LSMBO), IPHC, UMR 7178, Université de Strasbourg; Infrastructure Nationale de Proteomique ProFI - FR2048, 67087 Strasbourg, France, <sup>b</sup> Laboratoire d'ImmunoRhumatologie Moléculaire, Plateforme GENOMAX, INSERM UMR\_S 1109, <sup>c</sup> CompOmics group, VIB-UGent Center for Medical Biotechnology, VIB, Zwijnaarde, Belgium

### Introduction

Immunopeptidomics aims to identify major histocompatibility complex (MHC)-presented peptides on almost all cells that can be used, amongst others, in anti-cancer vaccine development. However, identification of immunopeptides through LC-MS suffers from some major problems and every single steps of this challenging identification workflow needs to be optimised, from the sample preparation, through MS analysis to the data treatment (1). Indeed, the nontryptic nature of immunopeptides hampers not only the acquisition through LC-MS but also the identification of immunopeptides' spectra. Also, the acquisition of immunopeptides spectra is limited by their very low abundances and even more so when searching for neoepitopes. First optimizations were done on a Q-Exactive Orbitrap mass spectrometer and recently we benchmarked different MS methods on a TimsTof Pro instrument. Finally, we evaluated MS<sup>2</sup>Rescore, a very promising AI-based rescoring algorithm using machine learning for peptide fragmentation spectrum (MS2PIP) and retention time (DeepLC) predictions (2).

### Methods

Our optimized sample preparation protocol was tested on 5 different cell types (HL60, PBMC, EBV, HEK, HeLa cells) starting from 5.10e8 cells. LC-MS/MS analyses were performed on a nanoAcquity Q Exactive HFX coupling (Waters, Thermo Scientific). Data were interpreted using Maxquant software for peptides identification and rescored using MS<sup>2</sup>Rescore at 1% FDR at peptide level.

Further MS optimizations were conducted on a nanoLC-TimsTof Pro coupling (Bruker Daltonics) platform on HL60 samples. Following parameters were benchmarked: Ion Mobility range with windows between 0.6 and 1.75 Vs/cm<sup>2</sup>, accumulation time varying from 100 to 200 ms, collision energy (CE) ramp from 10 to 55 eV or from 20 to 59 eV and scans number using either 6 or 10 scans.

### Results

Our immunopeptidomics protocol was developed and validated on 5 different cell types and led to an average of 3054 identified unique immunopeptides on a Q-Orbitrap system. These identification results were significantly boosted using MS<sup>2</sup>Rescore resulting in the identification of an average of 6146 unique peptides. In addition, MS<sup>2</sup>Rescore also overall increased the overlap between replicates from 53 to 61%. Lately, preliminary results using open modification searches also showed to clearly improve the number of peptide identifications. We then benchmarked 8 different MS methods including DDA and DIA-PASEF acquisitions on a Q-TOF instrument, the TimsTof Pro. Collision energy revealed to be the strongest impacting parameter on the output results with 3400 peptides identified with a 10-55 eV CE ramp which increased to 5800 peptides (vs 2500 peptides on HFX) with a 20-59 eV CE ramp, prior to any rescoring. The next step will consist in evaluating MS<sup>2</sup>Rescore on Timstof data to attempt further increasing our chances to detect challenging neo-antigens in cancer tissues.

### Conclusions

Altogether, these results demonstrate a significant improvement of our immunopeptidomics workflow. First, our sample preparation protocol was shown to be robust and repeatable. Then, the acquisition of immunopeptides spectra was improved thanks to an optimized method on the TimsTofPro. Last but not least, MS<sup>2</sup>Rescore boosts peptides identification rates of 50% on average. Overall, the coverage and depth of analysis achieved is very promising to enable detection of neo-antigens in cancer tissues.

### References

- 1) Faridi, P., Purcell, A. W., and Croft, N. P. (2018) In Immunopeptidomics we need a sniper instead of a shotgun. *Proteomics* 18, e1700464



- 2) Declercq,A., Bouwmeester,R., Hirschler,A., Carapito,C., Degroeve,S., Martens,L. and Gabriels,R. (2022) MS2Rescore: Data-Driven Rescoring Dramatically Boosts Immunopeptide Identification Rates. Molecular and Cellular Proteomics, 21, 100266.



## Unmasking hidden systemic effects of neurodegenerative diseases

Prof Bruno MANADAS<sup>a</sup>, Rosado MIGUEL<sup>a</sup>, Duarte DIANA<sup>a</sup>, Mendes VERA M<sup>a</sup>, Baldeiras INÊS<sup>a</sup>, Gomes ANDREIA<sup>b</sup>, Pires DIANA<sup>b</sup>, Santa CATIA<sup>a</sup>, Pinto JOANA<sup>a</sup>, Januário CRISTINA<sup>c</sup>, Santana ISABEL<sup>c</sup>, Mendonça ALEXANDRE<sup>d</sup>, Castelo-Branco MIGUEL<sup>e</sup>, Anjo SANDRA<sup>f</sup>

<sup>a</sup> CNC - Center for Neuroscience and Cell Biology - University of Coimbra, <sup>b</sup> CIBIT, <sup>c</sup> Serviço de Neurologia, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal, <sup>d</sup> Instituto de Medicina Molecular e Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal, <sup>e</sup> Coimbra Institute for Biomedical Imaging and Translational Research (CIBIT), Coimbra, Portugal., <sup>f</sup> CNC - Centro de Neurociências e Biologia Celular, Universidade de Coimbra, Coimbra, Portugal.

**Introduction:** The discovery of novel biomarkers remains a critical aspect of clinical research. Serum is a highly complex sample, and the detection of low-abundance proteins can be challenging due to the presence of highly abundant proteins. High molecular weight (HMW) fractionation can serve as a complementary approach to increase the range of detectable proteins, making it potentially useful in the identification of biomarkers in proteopathies such as Alzheimer's disease (AD) and Parkinson's disease (PD).

**Objectives:** The objectives of this study were twofold: i) to evaluate the applicability of HMW fractionation in the investigation of peripheral biomarkers in neurodegenerative diseases (NDs), specifically AD and PD, and ii) to screen a panel of potential biomarkers for AD and PD and stratify PD patients from atypical PD.

**Methodology:** The study utilized serum samples from three groups: AD (n = 22), PD (n = 24), and a control group (n = 12). Two different strategies were employed for each sample: i) direct analysis of total serum and ii) HMW fractionation via ultrafiltration (HMW serum). Protein quantification was acquired via LC-MS/MS, and the data were analyzed to identify statistically/biologically significant proteins. A discriminant model was constructed to differentiate between groups. PD subtypes were also addressed using PBMCs to differentiate between PD (n=20) and atypical PD (n=17).

**Results:** A total of 186 and 203 proteins were quantified in HMW serum and total serum, respectively, of which 39 and 41 were altered between at least two groups. A discriminant model was constructed using the altered proteins from each dataset, but the best performance was obtained using a combination of both, with a median sensitivity and specificity of 97.4% and 91.7%, respectively. The diagnostic capacity evaluation confirmed the robustness of this model for the classification of NDs (AUC = 0.999), and a panel of 10 proteins was subsequently identified. PD patients were further distinguished from atypical PD using a combined panel of 5 proteins.

**Conclusions:** HMW fractionation can be employed as a complementary approach to direct serum analysis for biomarker discovery, and PD stratification can be achieved using a small protein panel. These biomarkers require further investigation using a validation cohort.

## High-sensitivity phosphoproteomics with dia-PASEF

Dr Sara CECCACCI<sup>a</sup>, Kévin ROGER<sup>b</sup>, Joanna LIPECKA<sup>b</sup>, Inès METATLA<sup>b</sup>, Cerina CHHUON<sup>b</sup>, Ida Chiara GUERRERA<sup>b</sup>

<sup>a</sup> Proteomics platform Necker, Université Paris Cité - Structure Fédérative de Recherche Necker, INSERM US24/CNRS UAR3633, <sup>b</sup> Proteomics platform Necker, Université Paris Cité - Structure Fédérative de Recherche Necker, INSERM US24/CNRS UAR3633, Paris, FRANCE

Protein phosphorylation is a widely studied post-translational modification (PTM) that occurs frequently in eukaryotes and plays a crucial role in cellular signaling pathways and other cellular functions. Increasing sensitivity of mass spectrometry-based phosphoproteomics is key to increase coverage of signalling pathways. Data independent acquisition (DIA) is superior to data dependent acquisition (DDA) in terms of sensitivity, dynamic range coverage, and reproducibility. Recently, dia-PASEF, a combination of DIA and parallel accumulation-serial fragmentation, has further enhanced DIA sensitivity and the ability to distinguish isobaric phosphopeptides. Two recent publications (Bruker Daltonics Application 12-2022, LCMS-202, 1809177; Skowronek et al. *Molecular & Cellular Proteomics* **2022**, *21* (9)) demonstrate the feasibility of analyzing phosphopeptides using dia-PASEF with Spectronaut. In this study, we evaluated the performance of dia-PASEF in phosphopeptide analysis by comparing results obtained from different pipelines of dia-PASEF with DDA-PASEF.

Human breast cancer cell lysates (n=4) were digested following S-Trap micro spin column (Protifi) protocol and 100 µg proteome digest were subjected to TiO<sub>2</sub> phosphopeptide enrichment (Chauvet, M. et al. *Frontiers in Cellular and Infection Microbiology* **2021**, *11*, 637604). Amounts equivalent to 10 µg of protein extract were injected on a nanoElute (Bruker Daltonics) high-performance liquid chromatography (HPLC) system, coupled to a timsTOF Pro (Bruker Daltonics) mass spectrometer. Peptides were separated within 30 minutes gradients at flow rate of 250 nL/min using a packed emitter column (C18, 25 cm × 75µm 1.6µm) (Ion Optics). For DDA-PASEF we employed only one method with broad ion mobility (IM) range (1/K0 = 1.40 to 0.65 Vs cm<sup>-2</sup>), whereas for dia-PASEF we tested two methods with a deep IM range (1.24 to 0.75 Vs cm<sup>-2</sup>) and a broad IM range (1/K0 = 1.39 to 0.66 Vs cm<sup>-2</sup>). DDA-PASEF raw files were processed in PaSER (Bruker). dia-PASEF data were processed with Spectronaut (v17, Biognosys) in library-free (directDIA, Phospho PTM Workflow) and library-based mode (PTMs Significant Workflow), exploiting the human phospho-library published by Skowronek et al. (*Molecular & Cellular Proteomics* **2022**, *21*(9)).

With DDA-PASEF, we quantified ~8400 unique phosphopeptides on average in 4 samples (CV%=1). With dia-PASEF in directDIA, we were able to quantify ~13,200 unique phosphopeptides (CV%=1; data completeness=99%) using a deep IM range and ~12,800 (CV%=1; data completeness=99%) using a broad IM range. Although the number of phosphopeptides was similar, only 42% of these were in common between the two analysis. We then analysed the same dia-PASEF data against the Skowronek et al. phospholibrary, acquired in DDA-PASEF mode within an IM range from 1.51 to 0.6 Vs cm<sup>-2</sup>. Data acquired using the deep IM range led to the quantification of ~13200 unique phosphopeptides (CV%=4; data completeness=76%) with only 33% in common with those quantified in directDIA. However, data acquired using a broader IM range, more aligned with the IM range of the library, led to ~18800 (CV%=5; data completeness=77%).

Our results allow to draw different conclusions:

- i) dia-PASEF phosphopeptide analysis is 1.5 to 2.2 times more sensitive than DDA-PASEF;
- ii) directDIA shows a higher reproducibility and data completeness in comparison to the spectral library-based analysis (obtained on different samples).
- iii) in directDIA, the acquisition in a broader IM range doesn't significantly affect the global number of quantified phosphopeptides, but rather the nature of phosphopeptides identified. However, in the spectral library-based analysis, the widening of IM range can substantially increase the number of quantified features, if a phospho-library acquired in a similar broad IM range to the experimental samples is used.

In conclusion, our preliminary results show that dia-PASEF boosts phosphopeptides detection allowing to explore a wider part of the regulated phosphoproteome.

## Development of an orthogonal method for host cell protein characterization via immunoaffinity enrichment and ms analysis

Noelia DE LAMA<sup>a</sup>, Christine CARAPITO<sup>a</sup>, Diego BERTACCINI<sup>b</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique, IPHC (UMR 7178), CNRS, 67000 Strasbourg, France, <sup>b</sup> Global Analytical Development, Merck KGaA, Corsier-sur-Vevey, Switzerland

### Introduction

Host Cell Proteins (HCP) are process-related impurities that are produced by the host organism during biotherapeutic manufacturing and production. It is crucial to precisely identify and measure HCPs since they may pose a potential risk to patient safety or affect the degradation of the drug substance.

Due to the continuous evolving in regulations surrounding HCP in drug products, there is now a growing demand for the implementation of orthogonal techniques to confirm ELISA results and to address the inherent limitations of this method. Indeed, the ELISA test is not capable of identifying the HCP and is strongly reliant on the specificity of the antibody used.

Mass spectrometry shows promising potential as a viable alternative for this application. Hence, the primary aim of this project is to develop and tailor an immuno-enrichment and mass spectrometry-based test that can validate ELISA assays and provide a comprehensive analysis of HCP coverage.

### Methods

Initially, experiments were conducted to determine the most effective immuno-precipitation support, which included the testing of plate, magnetic beads, and column supports. The selection process involved the use of an in-house antigen and HCP antibody, with the Cygnus couple acting as a positive control, as well as monodimensional and bi-dimensional electrophoresis. Subsequently, the chosen optimal support was further corroborated through mass spectrometry analysis of the obtained samples.

In this regard, an exhaustive library was established for both HCP antigen samples by performing a DDA off-line gel fractionation and the Gas Phase Fractionation (Searle, 2020). The immuno-precipitation samples were prepared following an in-gel digestion protocol, they were then injected on the Orbitrap Eclipse™ Tribrid™ (ThermoFischer) operated in both DDA and DIA mode. The libraries obtained were applied for the analysis of the samples to verify which would be the most suitable to use. The data was processed using the latest version of Spectronaut (Biognosys).

### Results

In terms of building a comprehensive library, Gas Phase Fractionation presents itself as a reliable alternative to off-line gel fractionation, producing a comparable number of protein groups (3940 and 5077 respectively) with significant overlap between the two methods (50%).

For the analysis of samples from immuno-precipitation tests, the complete library of each antigen was used, which was built from both DDA and Gas Phase Fractionation runs. Since successful results were obtained only from the column support, these samples were further analyzed using MS. The samples consisted of three fractions: the input, which contained the immune-complex and non-immunoreactive HCP after incubation; the unbound, which contained only the non-immunoreactive HCP; and the elution, which corresponded to the immune-complex retained by the support. All fractions obtained showed expected behavior in terms of Protein Group identification with a significant overlap. However, a reduced number of Protein Groups were found solely in the unbound fraction.

### Conclusion

After implementing multiple testing approaches, it was concluded that the bead and plate supports were unsuitable as dependable immuno-precipitation supports. In contrast, the column was identified as a highly promising candidate, exhibiting satisfactory results.

The samples taken from the selected support confirm that the immuno-precipitation process was effective, with



significant overlap among the fractions and few unique protein groups found solely in the unbound. These unique protein groups are of particular interest to us because they are a priori undetectable by the ELISA test and thus escaped the radar so far.

## References

1. Vanderlaan et al. 2018 Experience with host cell protein impurities in biopharmaceuticals. *Biotechnol Prog.* Vol 34(No. 4): 828-837.
2. Searle et al. 2020. Generating high quality libraries for DIA MS with empirically corrected peptide predictions.



## Tree frogs living in Chernobyl Exclusion Zone exhibit different protein expression patterns than in the control areas.

Dr Sandrine FRELON<sup>a</sup>, Luc CAMOIN<sup>b</sup>, Stéphane AUDEBERT<sup>b</sup>, Samuel GRANJEAUD<sup>b</sup>, Emilie BAUDELET<sup>b</sup>, Jean-Marc BONZOM<sup>c</sup>, Clément CAR<sup>c</sup>, Léa DASQUE<sup>c</sup>, Elen GOUJON<sup>d</sup>, André GILLES<sup>e</sup>, Jean ARMENGAUD<sup>f</sup>, Olivier ARMANT<sup>c</sup>

<sup>a</sup> Institut de Radioprotection et de Sûreté Nucléaire, <sup>b</sup> MaP-CRCM (Marseille), <sup>c</sup> IRSN/SRTE LECO (St Paul Lez Durance), <sup>d</sup> IRSN/SESANE LRTOX (Fontenay aux roses), <sup>e</sup> UMR RECOVER (Marseille), <sup>f</sup> CEA LI2D (Bagnol sur Cèze)

Multi-omics approaches (e.g., transcriptomics, proteomics) are efficient tools to understand the potential toxic pathways of chemicals on organisms and link them with the physiological organism perturbations. However, such approaches remain limited for wild (non-model) species while the use of these species are necessary, notably to bring more environmental realism and more relevance to ecotoxicological studies in a context of ecological risk assessment. The massive improvement of sequencing methods over the last decade has opened the field of high throughput screening studies to these non-model species enabling drafting their genome and transcriptome assembly for the study of toxic effects through omics and multiomics methods. In order to study the effects of the radioactive fallout due to the Chernobyl nuclear accident on an amphibian living radiocontaminated areas, we developed a proteogenomics approach<sup>1,2</sup> on the tree frog *Hyla orientalis*, relevant sentinel species for environmental studies (sensitive, exposure to pollutants via diverse habitats depending on their life stage)<sup>3</sup>.

De novo sequencing and annotation of *Hyla orientalis* transcriptome as well as proteome analysis of the same tissue has been performed on individuals coming from a control area (Hyla Project, ECCOREV 2019). These allowed notably proteomic analyses of frog muscles of individuals coming from differentially radiocontaminated areas in order to decipher the biological pathways involved in long term effects of radiocontamination. We focus here on the proteomic approach which is described in the following. After protein extraction, concentration and digestion, peptide separation and analysis were performed via nanoHPLC (Ultimate 3000RSLC) and mass spectrometry (Orbitrap Fusion Lumos Tribrid - Thermo Fisher Scientific), using "Data Independent Acquisition" (DIA) mode. The full detected protein set is around 3600 proteins, whereas for data treatment we mostly used the data set where 70% proteins had valid quantitative values in each group, corresponding to a set of 2370 proteins. Analysis is still in progress, but preliminary results highlighted differential protein signature and abundance depending on the area of sampling (highly contaminated in the Chernobyl Exclusion Zone (CEZ) vs control area). Underlying biological pathways will be then studied.

Such comprehensive omics analyses of frogs will complement the genetic studies<sup>4</sup> and the phenotypic effects assessed in the same individuals and will be modelled according to the total dose rate received by them. All these results will enable the assessment of health status of frogs in the CEZ 35 years after the accident and contribute to assess the long-term impact of a nuclear fallout, still under scientific debate because of scarce and patchy studies<sup>5</sup>.

\* Co-first authorships

### REFERENCES:

- 1 Gouveia D. et al., Journal of Proteomics, 2019. 198: 66-77
- 2 Cogne Y. et al., Proteomics, 2020. 20 (10)
- 3 Helbing, C.C., Frontiers in Genetics, 2012. 3
- 4 Car, C et al., Evolutionary Applications, 2021
- 5 Beresford, N.A. et al, Journal of Environmental Radioactivity, 2020. 211

Acknowledgements to G. Orizaola (Oviedo University, Spain), P. Burraco (Doñana Biological Station, Spain) for sampling campaign help; S. Gashack and Y. Gulyaichenko (Chornobyl Center for Nuclear Safety, Radioactive Waste and Radioecology, Ukraine) for sampling campaign help and sample radioactive activity measurements.





## Deciphering the molecular mechanism of action of repurposed drugs using Thermal Proteome Profiling

Baptiste MOUYSSSET<sup>a</sup>, Marc-Antoine GERAULT<sup>a</sup>, Jérémy ARIEY-BONNET<sup>a</sup>, Stéphane AUDEBERT<sup>a</sup>, Raphael BERGES<sup>b</sup>, Samuel GRANJEAUD<sup>a</sup>, Dominique FIGARELLA-BRANGER<sup>b</sup>, Nicolas ANDRÉ<sup>c</sup>, Marion LE GRAND<sup>a</sup>, Luc CAMOIN<sup>a</sup>, Eddy PASQUIER<sup>a</sup>

<sup>a</sup> CRGM, <sup>b</sup> INP, <sup>c</sup> AP-HM

Until very recently, the paradigm in drug development was mainly based on “one drug, one target, one indication”. The concept of poly-pharmacology (= the interaction of a drug molecule with multiple targets) has however emerged and it is now estimated that drugs have on average 11.5 known molecular targets. This number is likely to be underestimated since large-scale studies characterizing the full interactome of drug molecules in human cells are scarce. Yet, this high degree of poly-pharmacology could open therapeutic avenues for drug repurposing, which consists in using existing drugs in new indications.

Mebendazole (MBZ), an anti-helminthic agent, which belongs to the benzimidazole class, has shown anticancer properties in various human brain tumor models leading to ongoing clinical trials. Using an orthotopic xenograft model of glioblastoma, we also showed that MBZ could potentiate the anticancer activity of temozolomide, the gold standard treatment for glioblastoma. However, the mechanism(s) of action involved is still poorly understood and warrants further investigations. Using *in silico* molecular target prediction we unveiled MBZ as a potent MAPK14 inhibitor, but it still doesn't fully explain the activity of MBZ against brain tumors. Revealing the interactome of MBZ using unsupervised proteomics could thus decipher its complete mechanism.

To do so, we used “Thermal Proteome Profiling” (TPP) combined with label-free quantitative mass spectrometry. TPP is a biophysical test based on the principle of the modulation of protein thermostability induced by their interaction with a ligand, allowing the detection of protein interactions with a drug within a cell. We applied TPP on glioblastoma cell lysates treated with vehicle or MBZ. Using a LC-MS/MS strategy coupled with a deep learning algorithm for raw data processing (Data Independent Analysis - Neural Network; DIA-NN), we achieved a high proteomics depth and identified over 6,000 proteins per sample. One of the main challenges in chemoproteomics applied to interactome analysis is to find a way to discriminate between true and false positive interactors. To do so we established a “Stability Rate” score (SR) that reflects how much MBZ treatment impacted the thermal stability of a given protein. For each protein we extracted the fold changes (MBZ vs Vehicle) at every temperature, took the absolute value from the mean, and ordered them in decreasing order. We then fitted a weighted linear regression to these data. We applied weights five times higher for the two biggest fold changes so more importance is given to them while still considering the whole profile. From this regression we extracted the intercept with the y-axis and to keep track if the protein is destabilized or stabilized, we multiplied this value by the sign of the mean of all the fold changes (either -1 or 1). Finally, we applied a z-score normalization on this SR score and ranked the identified proteins. We found well known targets of MBZ such as tubulins as well as recently identified target MAPK14 but also uncovered additional interactors of MBZ including MAP2K4, MAP2K5 and RIPK2, which belong to the neurotrophin signalling pathway together with MAPK14. We now aim to validate those molecular targets by using biophysical approaches and functional genomics to both characterize novel vulnerabilities and decipher the mechanism of action of MBZ in brain tumors.

By using TPP coupled with DIA-LC-MS/MS, we uncovered the interactome of MBZ in glioblastoma cells and found proteins of the neurotrophin signalling pathway as new targets of MBZ. Overall, chemoproteomics appears to be a powerful and versatile tool to elucidate the mechanisms of action of existing drugs and represents a unique opportunity to unveil novel targetable vulnerabilities.

## Taxonomical exploration of complex samples by proteomics-derived proteotyping

Dr Jean ARMENGAUD<sup>a</sup>, Charlotte MAPPA<sup>a</sup>, Guylaine MIOTELLO<sup>a</sup>, Mélodie KIELBASA<sup>a</sup>, Béatrice ALPHA-BAZIN<sup>a</sup>, Olivier PIBLE<sup>a</sup>

<sup>a</sup> CEA

The correct identification of the organisms present in a complex sample is a crucial issue for microbiological diagnostics and research. Proteotyping based on tandem mass spectrometry, a methodology derived from metaproteomics, can help establish an inventory of organisms because the origin of the detected peptides can be traced back. The methodology we developed takes into account taxon-specific peptide sequences, but also the other peptides that are shared between taxa. An unbiased search of a large generic database is followed by a peptide search restricted to the most representative organisms to better delineate taxa. We introduced the term „Taxa-to-Spectrum Matches“ (TSMs) to best describe the composition of any sample. Here, we demonstrate the sensitivity of the approach, as trace amounts of biological material equivalent to a single human cell are sufficient (Mappa et al. 2023a). We document its accuracy with an artificial reference assemblage of 24 bacterial species with several challenges such as very closely related organisms (Mappa et al., 2023b). We illustrate the potential of the methodology to describe biofilms collected from very hostile environments such as nuclear storage pools (Pible et al., 2023). For these biofilms, we report the identification of three genera, namely *Sphingomonas*, *Caulobacter*, and *Acidovorax*. Once identified, their functional characterization by metaproteomics showed that these organisms are metabolically active. The differential expression of the Gene Ontology GOslim terms between the two main microorganisms highlights their metabolic specialization. The methodology can also be applied on biological material from very old samples (Oumarou Hama et al., 2023). As the approach can be performed rapidly on a limited amount of material, we foresee a wide application in microbiological diagnostics and microbiome research (Armengaud 2023).

### References:

Mappa C, Alpha-Bazin B, Pible O, Armengaud J (2023a) Evaluation of the limit of detection of bacteria by tandem mass spectrometry proteotyping and phylopeptidomics. *Microorganisms* 11:1170.

Mappa C, Alpha-Bazin B, Pible O, Armengaud J (2023b) Mix24X, a lab-assembled reference to evaluate interpretation procedures for tandem mass spectrometry proteotyping of complex samples. *International Journal of Molecular Sciences*. In press.

Pible O, Petit P, Steinmetz G, Rivasseau C, Armengaud J (2023) Taxonomical composition and functional analysis of biofilms sampled from a nuclear storage pool. *Frontiers in Microbiology* 14:1148976.

Oumarou Hama H, Chenal T, Pible O, Miotello G, Armengaud J, Drancourt M (2023) An ancient coronavirus from individuals in France, circa 16th century. *International Journal of Infectious Diseases* 131:7-12.

Armengaud J (2023) Metaproteomics to understand how microbiota function: the crystal ball predicts a promising future. *Env Microbiol* 25(1):115-125.

## Deep proteomic profiling reveals the dynamics of the host- microbiota interactions from cystic fibrosis sputum samples

Dr Lucia GRENGA<sup>a</sup>, Hamid HACHEMI<sup>b</sup>, Pauline HARDOUIN<sup>b</sup>, Olivier PIBLE<sup>c</sup>, Raphael CHIRON<sup>d</sup>, Helene MARCHANDIN<sup>e</sup>, Jean ARMENGAUD<sup>c</sup>

<sup>a</sup> University of Paris-Saclay, CEA, INRAE, Laboratoire Innovations technologiques pour la Détection et le Diagnostic (Li2D), <sup>b</sup> University of Montpellier, Laboratoire Innovations technologiques pour la Détection et le Diagnostic (Li2D), Bagnols-sur-Cèze, France, <sup>c</sup> University of Paris-Saclay, CEA, INRAE, Laboratoire Innovations technologiques pour la Détection et le Diagnostic (Li2D), Bagnols-sur-Cèze, France, <sup>d</sup> HydroSciences Montpellier, CNRS, IRD, Centre de Ressources et de Compétences de la Mucoviscidose, Université de Montpellier, CHU de Montpellier, Montpellier, France, <sup>e</sup> HydroSciences Montpellier, CNRS, IRD, Service de Microbiologie et Hygiène Hospitalière, Université de Montpellier, CHU de Nîmes, Nîmes, France

Cystic fibrosis (CF) is a genetic disorder characterized by persistent bacterial infection and inflammation in the respiratory tract. While modern therapies can sometimes help alleviate respiratory symptoms, by ameliorating the residual function of the protein responsible for the disorder, managing chronic respiratory infections remain a significant challenge. The lung microenvironment, including its microbiota, are potential driving causes for the worsening of patient symptoms and the variability in the treatment outcome. In this study, we applied cutting-edge proteomic profiling (PMID: 34207804, PMID: 36312921, PMID: 32133743, PMID: 34917757) to investigate the dynamic interactions between the host and the respiratory microbiota in CF patients chronically infected with *Mycobacterium abscessus*.

Our results show a dynamic variation of the microbiota composition and significant changes in the metabolic activity of known CF pathogens over time and as a function of the presence of *Mycobacterium*. Additionally, the profile of the associated host's response suggested that, after being internalized by macrophages, *Mycobacterium* enters a dormant state. To assess the correlation between the proteomics data and the clinical ones over time and shortlist potential biomarkers, an integrative machine learning method based on gradient boosting was developed.

The approach used allows a better understanding of the roles of the microbiota components, and their contribution to disease progression and host response. It also highlights the potential of proteomic profiling to provide new insights into the complex interactions among microbial species in the respiratory microbiota of CF patients. This could lead to the identification of new targets for potential antimicrobial therapy or microbiota-based intervention, ultimately improving the treatment and management of CF.

## Exploring prostate carcinogenesis progression through urine proteome profiling: an animal model study

Alexandra MOREIRA-PAIS<sup>a</sup>

<sup>a</sup> LAQV-REQUIMTE, Department of Chemistry, University of Aveiro, 3810-193, Aveiro, Portugal

**Authors:** Alexandra Moreira-Pais <sup>a,b,c</sup>, Rita Nogueira-Ferreira <sup>d</sup>, Susana Aveiro <sup>e</sup>, Tânia Melo <sup>a,f</sup>, José Alberto Duarte<sup>b,g</sup>, Fernanda Seixas <sup>h</sup>, Pedro Domingues <sup>a</sup>, Francisco Amado <sup>a</sup>, Paula Oliveira <sup>c</sup>, Rita Ferreira <sup>a</sup> and Rui Vitorino<sup>d,i</sup>

<sup>a</sup>LAQV-REQUIMTE, Department of Chemistry, University of Aveiro, 3810-193, Aveiro, Portugal

<sup>b</sup> Research Center in Physical Activity, Health and Leisure (CIAFEL), Faculty of Sports, University of Porto (FADEUP) and Laboratory for Integrative and Translational Research in Population Health (ITR), Porto, Portugal

<sup>c</sup> Centre for Research and Technology of Agro Environmental and Biological Sciences (CITAB), Inov4Agro, University of Trás-os-Montes and Alto Douro (UTAD), Quinta de Prados, 5000-801, Vila Real, Portugal

<sup>d</sup>UnIC@RISE, Department of Surgery and Physiology, Faculty of Medicine of the University of Porto, Alameda Professor Hernâni Monteiro, 4200-319, Porto, Portugal

<sup>e</sup>GreenCoLab-Green Ocean Association, University of Algarve, 8005-139, Faro, Portugal

<sup>f</sup>CESAM-Centre for Environmental and Marine Studies, Department of Chemistry, University of Aveiro, 3810-193, Aveiro, Portugal

<sup>g</sup>TOXRUN - Toxicology Research Unit, University Institute of Health Sciences, CESPU, CRL, Gandra, Portugal

<sup>h</sup> Animal and Veterinary Research Center (CECAV), University of Trás-os-Montes and Alto Douro (UTAD), 5000-801, Vila Real, Portugal

<sup>i</sup> Institute of Biomedicine - iBiMED, Department of Medical Sciences, University of Aveiro, 3810-193, Aveiro, Portugal

**Abstract:** Prostate cancer (PCa) is a deadly disease in men, requiring the development of new diagnostic tools. In this study, we analyzed the urine proteome using GeLC-MS/MS to gain insights into the progression of prostate carcinogenesis. We used an animal model of chemically and hormonally induced adenocarcinoma and considered two time points (27 and 54 weeks after cancer induction) to reflect the progression of prostate lesions from preneoplastic to neoplastic. Age-matched healthy animals were also used as controls, allowing us to investigate the influence of aging on prostate remodeling. At week 54, the incidence of preneoplastic and neoplastic lesions in animals with PCa was 100%. GeLC-MS/MS and subsequent bioinformatics analyses revealed several proteins involved in prostate carcinogenesis. Retinol-binding protein 4 levels increased, while cadherin-2 levels decreased, in early stages of the disease, whereas enolase-1 and T-kininogen 2 levels increased and isocitrate dehydrogenase 2 levels decreased in more advanced stages. With increasing age, urinary levels of clusterin and corticosteroid-binding globulin increased, and neprilysin levels decreased, all of which appear to play a role in prostate hyperplasia or carcinogenesis. This study provides a valuable starting point for future studies aimed at identifying human urine proteins that could be targeted for the early detection of aging-related maladaptive changes in the prostate, potentially leading to cancer.

# Sponsor Oral Presentations



## Scalable Sample Preparation Methods for LC-MS Assays

Nicolas Autret, PhD - Covaris

Translational or clinical laboratories require automated, hands-off solutions for protein sample preparation to provide reproducibility, increased efficiency, higher quality results, and faster turnaround time. This presentation will describe the robust protocols elaborated with leading mass spectrometry labs for FFPE or LCM samples, as well as more recent developments for ultra-low input processing, or sensitive extraction including for phosphoproteomics.

Those methods are perfectly suited for batches from 8 to 96 samples and are designed to allow robust preparation with every clean-up and digestion method. When working with bead-based protein aggregation capture, the workflow can now include an accelerated trypsin digestion step which yields similar quality compared to overnight digestion, in a more controlled way, and leads to 96 samples processed in 7 to 8 hours.

In summary, AFA-based processing:

- makes FFPE tissue proteomics compatible with the clinical setting, through a non-toxic deparaffinization and rehydration approach,
- facilitates workflow simplification by allowing homogenization through digestion in a single well with an automation compatible high-throughput (HTP) workflow,
- allows an increased throughput and decreased turn-around-time,
- makes one-day sample prep also fully applicable to laser capture microdissections (LCM) and fresh/frozen tissue preparation.

## Pushing Frontiers of High-throughput High-resolution Analysis: Orbitrap Technology Unites with a New Star

Daniel Hermanson, Alexander Makarov, Christian Hock, Hamish Stewart, Dmitry Grinfeld, Anastassios Giannakopoulos, Johannes Petzoldt, Toby Shanley, Eduard Denisov, Amelia Peterson, Matthew Garland, Eugen Damoc, Martin Zeller, Tabiwang Arrey, Anna Pashkova, Andreas Kühn, Matthias Biel, Arne Kreutzmann, Bernd Hagedorn, Immo Colonius, Adrian Schütz, Arne Stefes, Ankit Dwivedi, Daniel Mourad, Max Hoek, Philipp Cochems, Alexander Kholomeev, Robert Ostermann, Gregor Quiring, Maximilian Ochmann, Sascha Möhring, Alexander Wagner, André Petker, Sebastian Kanngiesser, Michael Wiedemeyer, and Wilko Balschun

*Thermo Fisher Scientific, Bremen, Germany*

Speaker: Shakir Shakir

*Thermo Fisher Scientific, Les Ulis, France*

### Introduction

Over the last decades translational medicine has advanced by studying larger and larger patient cohorts using next-generation genome sequencing, increasingly complemented by proteomic and metabolomic analysis.

To remain competitive against other technologies, mass spectrometry (MS) needs to rise to the challenge of unprecedented throughput of very deep analyses of very complex samples. With analysis time reduced to several minutes by fast liquid separations, confident identification and quantitation over many orders of dynamic range still requires a generational leap in the sequencing performance of MS.

This need is addressed in this work by a novel instrument that complements Orbitrap™ analyzer performance by combining some principles of electrostatic trapping with destructive detection.

### Methods

A new class of high-resolution accurate-mass analyzer has been deployed on a novel MS to break the sound limit of sequencing by LC/MS/MS.

Operation of the analyzer starts from trapping of the incoming ion beam in a newly developed dual-pressure linear trap, followed by ion ejection as short packets, exactly like from the C-trap in the Orbitrap analyzer. Ions are then directed to an open electrostatic trap of asymmetric planar geometry that ensures multiple ion reflections on overlapping paths. Absence of any grids or collimators on the way ensures lossless delivery of ions through the entire analyzer onto a novel secondary electron detector.

Performance of the entire instrument was characterized using Pierce™ Flexmix™ calibration solution, HeLa digest and intact proteins.



## Preliminary data

Integration of the novel analyzer with quadrupole and Orbitrap analyzers into a single instrument has been enabled by optimized synchronization of all ion-optical elements of the resulting system. As a result, a continuous ion beam from the electrospray ion source is converted into sequential ion packages, with multiple ion packages being simultaneously processed by different elements of ion optics. In operation, panoramic MS1 spectra are typically acquired on the Orbitrap analyzer in parallel to tens of MS/MS spectra acquired on the new analyzer.

Experiments have confirmed that the entire MS/MS cycle could be compressed into 5ms duration for most analyses due to the combination of lossless operation and high space charge capacity of the new analyzer with single-ion efficiency, narrow pulse width and high dynamic range of secondary electron detection. At the same time, elaborate design and very high precision of geometry and voltages of trap electrodes enables mass resolutions up to 100,000 and low-ppm mass accuracies over the entire mass range of interest.

This unique combination of analytical parameters was shown to enable unprecedented performance in several high-throughput applications. For example, analysis of tryptic peptides demonstrated a several-fold (3-4x) improvement of throughput for the same depth of analysis (identification and quantitation) compared to state-of-the-art Thermo Scientific™ Orbitrap Exploris™ 480 MS. Further, single-shot proteome analysis at the depth >8,000 proteins in HeLa digest could be now carried out at the rate above 180 proteomes per day, and with depth >10,000- at the rate of 60 proteomes per day.

# PREOMICS

## **Solution to move to proteomics sample prep standardization for routine analysis**

Abstract: PreOmics provides solutions for LC-MS based proteomics sample preparation to support scientists in their daily experiments. Presentation will be focused on biofluid and FFPE samples.





The 4D-Proteomics<sup>TM</sup> ecosystem : enabling large-scale personalized medicine research



## Enabling high-throughput and standardised proteomics using Evosep One and end-to-end, automated sample preparation

### Evosep

Translational or clinical laboratories require automated, hands-off solutions for protein sample preparation to provide reproducibility, increased efficiency, higher quality results, and faster turnaround time. This presentation will describe the robust protocols elaborated with leading mass spectrometry labs for FFPE or LCM samples, as well as more recent developments for ultra-low input processing, or sensitive extraction including for phosphoproteomics.

Those methods are perfectly suited for batches from 8 to 96 samples and are designed to allow robust preparation with every clean-up and digestion method. When working with bead-based protein aggregation capture, the workflow can now include an accelerated trypsin digestion step which yields similar quality compared to overnight digestion, in a more controlled way, and leads to 96 samples processed in 7 to 8 hours.

In summary, AFA-based processing:

- makes FFPE tissue proteomics compatible with the clinical setting, through a non-toxic deparaffinization and rehydration approach,
- facilitates workflow simplification by allowing homogenization through digestion in a single well with an automation compatible high-throughput (HTP) workflow,
- allows an increased throughput and decreased turn-around-time,
- makes one-day sample prep also fully applicable to laser capture microdissections (LCM) and fresh/frozen tissue preparation.

## Improved immunoaffinity enrichment methods for arginine methylation

Authors: **Maxime Jacquet**, B.M. Zee, H. J. Peckham, C. L. Farnsworth, A. J. Nelson, J. Min Ren, M. C. Palazzola, K. E. Abell, M. P. Stokes.

### Abstract:

Introduction: Immunoaffinity purification (IAP) coupled with LC-MS/MS is an established technique for the study of posttranslational modifications (PTMs). Recently, we introduced an antibody-based magnetic bead conjugate IAP method that enables identification of sites of PTMs including lysine ubiquitination (KGG), lysine acetylation (AcK) and tyrosine phosphorylation (pY) with improved sensitivity and specificity over preexisting agarose bead methods. Here, we present three novel magnetic bead IAP methods that enable identification of monomethylated (Me-R), symmetric dimethylated (SDMA), and asymmetric dimethylated (ADMA) arginine sites, with improvements over the respective agarose bead methods.

Arginine methylated proteins are involved in processes including transcriptional regulation and signal transduction. Arginine methylation is carried out by the arginine N-methyltransferase (PRMT) family of enzymes. Though all PRMT proteins catalyze monomethylation, Type I PRMTs such as PRMT1 add an additional methyl group to produce ADMA, while Type II PRMTs such as PRMT5 produce SDMA. Recent reports have highlighted the relevance of PRMT5, and SDMA, to cancer and stem cell self-renewal.

Methods: Tryptic mouse and human peptides were used as input material for Me-R, SDMA, and ADMA magnetic IAP beads. Peptides were incubated with beads using cold 1x IAP buffer for 2 hours at 4°C. Beads were washed with cold 1x IAP buffer, then water, and bound peptides were eluted with trifluoroacetic acid. Eluted peptides were desalted using C18 Stage-Tips and analyzed on ThermoFisher Q-Exactive or Fusion Lumos mass spectrometers using data-dependent acquisition (DDA).

Peptides were resolved using a 90 or 120min reversed phase gradient from 7.5 to 32% acetonitrile on a 25cm or 50cm C18 column. Peptides were identified by Comet with 1% of total identifications mapped to reverse sequences. Skyline software was used for MS1 peak review and quantification. For automation on Kingfisher Apex robot, input peptides were resuspended in bulk and aliquoted to deep well plates. After IAP on Kingfisher Apex robot, all enriched peptides were manually desalted as described above.

### CONCLUSIONS

The new Me-R, SDMA, and ADMA magnetic IAP beads provide 6-60% improvement in number of unique modified peptides identified and nearly 100% improvement in PTM capture specificity across a variety of different cell and tissue sample types. The ease of handling of magnetic beads facilitates manual benchwork and implementation on automation platforms.

### Innovative ready-to-use SPE kits for proteomic workflows

Dr Florine Hallez, *Product Manager Bioanalytics, AFFINISEP*

AFFINISEP was founded and has continued to operate for over a decade on the principle of manufacturing high quality and innovative sample preparation kits to simplify your procedures and help you achieve reliable results.

On this principle, AFFINISEP has recently developed the BioSPE® product range, composed of ready-to-use kits specially designed for various applications in bottom-up proteomic workflows, such as peptide desalting (BioSPE® PurePep), peptide fractionation (BioSPE® PepFrac), magnetic beads removal after SP3 procedure (BioSPE® BeadRem), or glycopeptide enrichment (BioSPE® GlycaClean). The BioSPE® kits can also be used for intact protein purification (BioSPE® PureProt), in top-down approaches.

These BioSPE® kits are based on our unique Solid Phase Extraction (SPE) technology, made of small sorbent particles densely packed in a soft, uniform and mechanically stable membrane with no dead volume, that has demonstrated excellent reproducibility for the efficient extraction of analytes with high recovery yields.

All our BioSPE® kits offer a real flexibility and complementarity of formats, as they are available as StageTips, Spin columns, Cartridges or 96 Well Plates for high throughput. Finally, all our microelution tools are available with different capacities to perfectly adapt to your sample and best meet your needs, and can thus be used for all sample types, ranging from for single-cell analysis to samples containing several mg of proteins.



SP8

**comprehensive peptide mapping of proteins using a complementary approach of electron activated dissociation (EAD) and Collision Induced Dissociation (CID) fragmentation techniques**

Heather Chassaing, Sciex, Villebon sur Yvette, France

[Heather.Chassaing@Sciex.com](mailto:Heather.Chassaing@Sciex.com)

Extensive protein characterization can be achieved by utilizing the complementary fragmentation obtained from both collision induced dissociation (CID) and electron activated dissociation (EAD). The performance of CID and EAD in terms of peptide backbone coverage for disulfides at different complexity level were compared. For simple disulfides with relatively short peptides, both CID and EAD works equally well in generating fragments for good peptide backbone sequence coverage. However, when the length of peptides increases, EAD provides unique benefits in generating wealth of fragments for every peptide involved in the disulfides. Particularly, for disulfide peptides containing more than 2 disulfide bonds, CID generated very limited peptide backbone fragments; whereas EAD generated abundant fragments for the solid identification of every peptide. In addition, as a soft fragmentation technique, EAD also was able to successfully preserve the labile O-glycan on the disulfides, providing additional benefit of accurately locating the O-glycosylation site.





**Mag-Net: deep plasma profiling using streamlined and single-shot membrane vesicle enrichment enabled by MagReSyn® hyper-porous magnetic microparticles**

Extracellular vesicles (EVs) are released by nearly all cell types and thought to be involved in intercellular communication. The rich protein, DNA and RNA cargo they carry can provide valuable information on disease onset and progression. We present a single-shot enrichment strategy (Mag-Net) based on the unique properties of MagReSyn® magnetic microparticles, to capture membrane-bound vesicles improving the dynamic range of proteins detectable in plasma. Mag-Net is sensitive (requires <100 µL plasma input) robust, streamlined, high-throughput and cost-effective. The workflow is automated on an open-source magnetic handling station, does not require any time-consuming filtration and/or centrifugation steps, and is thus ideally suited for processing large clinical cohorts. Coupled to data independent acquisition, we demonstrate that using Mag-Net >42,000 peptides from >4,500 plasma proteins could be detected with high precision.

We further feature the latest developments in the MagReSyn® product range, including mass spectrometry compatible protease-resistant Streptavidin beads, that result in multiple-fold reduction in streptavidin contaminating peptides, compacting the dynamic range thereby increasing the detection of target peptides with increased selectivity, likely due to the lower ionic charge of the Streptavidin MS.

**ReSyn Biosciences** utilises a proprietary microparticle technology platform, which is applied to improve the efficiency of Mass Spectrometry and Bioseparation workflows.



## cellenONE for multiplexed and label-free single cell proteomics sample preparation.

### Background

Single cell proteomics is a rapidly developing field with the potential to make important contributions to the understanding of cellular heterogeneity. Recent enhancements of fast and sensitive mass spectrometry paired with automated single cell sorting and sample preparation realized with the cellenONE platform allows for sensitive proteome analyses at the single cell level. Sample preparation for both multiplexed or label-free single cell proteomics can be undertaken using the cellenONE, a liquid- and cell-dispensing robot, associated with the proteoCHIP consumables product line. This combination allows processing of single cells in sub  $\mu\text{L}$  volumes, reducing adsorptive losses and chemical noise.

### Results

Compared to other single cell dispensers on the market, the cellenONE is standing out due to its high precision and versatility. The capability of performing both cell isolation and reagent dispensing on the same platform allows to run the sample preparation workflow in one single step, removing all error prone operator manipulations. The high accuracy and precision of spotting allows to miniaturize the sample preparation to a level never achieve before.

Unlike existing cell sorting/dispensing technologies, the gentle, piezo-driven acoustic dispense technology preserves the viability of dispensed cells and maintains protein expressions. Cells of any size between 0.5 and 80  $\mu\text{m}$  in diameter can be processed within the same sample. Compared to other cell sorting technologies, only low volumes of cell suspension (1-50  $\mu\text{L}$ ) are required at comparatively low concentration (100-200 cells/ $\mu\text{L}$ ) to perform isolation, which makes the cellenONE especially interesting for sorting rare cell species (liquid biopsy, CTCs, CSF, etc.)

Images of nozzle tip enables real-time observation of cells during the isolation process and are also saved for documentation. This allows combined analysis of morphological and proteomics data.

The ability to work with both aqueous and organic reagents, in combination with humidity and temperature control to prevent evaporation of low volumes, makes the cellenONE the perfect instrument to perform end-to-end sample preparation for single cell proteomics workflows.

### Conclusion

Both multiplexed and label-free analysis workflow using the cellenONE for sample preparation resulted in identifying > 1500 proteins per run.

## TEACHING AN OLD DOG NEW TRICKS: EXTENDING THE UTILITY OF YOUR MASS SPECTROMETER

Mass spectrometry is a ubiquitous tool in a multitude of areas of research, and has a huge industry to supply instruments to research institutions around the globe. There are a multitude of commercially available instruments that can do almost anything you can desire. However, one of the downsides of this is that although there are many mass spectrometers available, individual mass spectrometers are typically targeted for a specific application. Therefore, the mass spectrometer you have may not be able to do exactly what you want them to do. In this presentation, modification of mass spectrometers to be able to perform new experiments will be discussed.

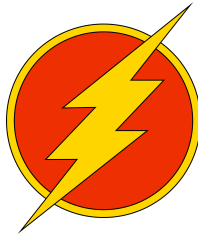
One such set of modification is the extension of mass spectrometers to be able to perform something that is currently unavailable on commercially available mass spectrometers. Two examples of this are IR/UV spectroscopy and soft landing. In both these cases, the capability of a Waters Synapt instrument to prepare mass selected ions and separate them by ion mobility is leveraged to prepare mass and shape selected ion beams, which can then either be trapped for spectroscopic measurements which can be used to elucidate structure, or to land them on a surface which can then be removed from the vacuum and studied by alternative structural techniques such as electron microscopy.

Alternatively, the mass spectrometry itself is fine for the application, but additional functionality can help with interpretation of better performance of experiments. One example of this would be the study of large proteins or protein assemblies containing PTMs or glycan modifications. This can lead to an extremely dense spectrum that is difficult to convolute. Here, addition of charge reduction can help to deconvolute the spectrum and aid assignment of peaks. Another example is the modify a mass spectrometer so it is more suitable for performing native mass spectrometer, by improving desolvation and extending the range of the quadrupole allowing for mass selection of higher  $m/z$  ions.

In conclusion, modification of existing mass spectrometers can significantly extend the utility and lifetime of mass spectrometers, allowing them to perform experiments not possible on current commercially available machines.

# Flash Presentations





## Immunoglobulins expression by non immune cells

Dr Alice CAPUZ<sup>a</sup>, Soulaimane ABLOUARD<sup>a</sup>, Marco TREROTOLA<sup>a</sup>, Dasa CIZKOVA<sup>b</sup>, Etienne COUYAUD<sup>a</sup>, Firas KOBEISSY<sup>c</sup>, Franck RODET<sup>a</sup>, Michel SALZET<sup>a</sup>

<sup>a</sup> Univ. Lille, Inserm, U-1192 - Laboratoire Protéomique, Réponse Inflammatoire et Spectrométrie de Masse-PRISM, F-59000 Lille, France, <sup>b</sup> Institute of Neuroimmunology, Slovak Academy of Sciences, Dúbravská cesta 9, 845 10 Bratislava, Slovakia, <sup>c</sup> Department of Biochemistry and Molecular Genetics, Faculty of Medicine, American University of Beirut, Lebanon.

The CNS includes the brain and spinal cord, and is composed of neurons, glial cells, and immune cells. Neurons carry nerve information while glial cells maintain CNS stability. Astrocytes are the most abundant glial cells and regulate nutrient supply, CNS homeostasis, nerve information transmission, and synaptic formation. The central nervous system has long been considered to be excluded from the peripheral immune system. However, contrary to the immunology dogma stipulating that only B lymphocytes are capable of producing antibodies, we have shown that astrocytes and neurons could be a new cellular source in the central nervous system. An antibody of two identical heavy and two light chains with a paratope that binds to specific antigens. This diversity is achieved through V(D)J recombination, where the V, J, and D segments are rearranged. However, the edification of antibodies issued of these neural cells is different from the ones found in B cells and resembles to aberrant immunoglobulins. Some are pseudogenes and belongs what we named, the Ghost proteome or alternative proteins. Alternative proteins originate from the 5'- and 3'- UTR mRNA regions, frameshifts of mRNA ORFs or from non-coding RNAs. Long-time considered as non-coding, recent in-silico translation prediction methods enriched the protein databases allowing the identification of new target structures that have not been identified previously.

Then, we focused our interest on one a kappa light chain variable region. We demonstrated that this protein named "Heimdall" was expressed by astrocytes and was secreted under inflammatory conditions with lipopolysaccharides (LPS). Immunoprecipitation experiments showed that the Heimdall interactome contained proteins related to astrocyte fate keepers such as "NOTCH1, EPHA3, IPO13" as well as membrane receptor protein including "CHRNA9; TGFBR, EPHB6, and TRAM". However, when Heimdall protein was neutralized utilizing a specific antibody or its gene knocked out by CRISPR-Cas9, sprouting elongations were observed in the corresponding astrocytes. Interestingly, depolarization assays and intracellular calcium measurements confirmed a phenotype switch from astrocytes to neuron-like phenotype. Proteomic analyses performed under LPS stimulation, revealed the expression of neuronal factors, stem cell proteins, and neurogenesis of astrocyte convertor factors such as EPHA4, NOTCH2, SLIT3, SEMA3F, suggesting a role of Heimdall to regulate astrocytic fate.

Besides Heimdall, we also identified an aberrant form of the IgG2b heavy chain constant region. In fact, transcriptomic analysis showed that the messenger encoding this IgG2b constant part didn't have a variable part coding sequence at its 5' end, but a Kozak sequence located directly upstream of the exon coding the CH1 domain. Moreover, invalidation of IgG2B by CRISPR-Cas9 in astrocytes show a higher abundance of proteins characteristic of hematopoietic cells (FRG1, PIR), neural stem cells (YAP1, TEAD3, SMAD2/9, MCC, KDM2A and DDHD2) and regulation of neuritogenesis (PLXNA1, ROCK1, MAGI3, TBL1). These results suggest that this inhibition of this heavy chain led to dedifferentiation of astrocytes into neural progenitor cells via activation of the BMP-YAP1-TEAD3 pathway (Capuz et al., 2023). Conversely, overexpression of IgG2B in astrocytes induced the CRT1-CREB-BDNF signaling pathway involved in gliogenesis and instead directs astrocytes toward a pro-inflammatory profile. Furthermore, proximal labeling (Bio-ID) experiments revealed that IgG2B is N-glycosylated by the OST complex, addressed to the membranes of ATPase complex-containing vesicles, and partially behaves as SLC3A2 (CD98hc) through its association with SLC7A5 (LAT1) and its ability to recruit ILKAP that potentially regulates downstream signaling pathways (Capuz et al., 2023).

Taken together, these results support that aberrant IgG2B and Heimdall chains are involved as a gatekeeper of astrocyte identity and that its inhibition leads to the conversion to a neural stem cell phenotype.



## Going against the grain: development of denaturing mass photometry for fine-tuning of protein-protein cross-linking reactions

Hugo GIZARDIN-FREDON<sup>a</sup>, Oscar HERNANDEZ-ALBA<sup>a</sup>, Sarah CIANFERANI<sup>a</sup>

<sup>a</sup> 1 Laboratoire de Spectrométrie de Masse BioOrganique, IPHC UMR 7178, Université de Strasbourg, CNRS, 67000 Strasbourg, France ; 2 Infrastructure Nationale de Protéomique ProFI - FR2048, 67087 Strasbourg, France

Structural MS is a generic term that encompasses a series of MS-based strategies adapted for the characterization of non-covalent assemblies. Among them, cross-linking followed by mass spectrometry (XL-MS) is a covalent labelling proteomics-based technique that has drastically progressed this past decade from *in vitro* to even *in situ/in cellulo* applications<sup>1</sup>. XL-MS approaches provide snapshots of non-covalent protein-protein interactions (PPI) through determination of spatial proximities and distance constraints. The first step of chemical XL reaction itself is of utmost importance in the XL-MS workflow. Indeed, before LC-MS/MS analysis, optimal XL conditions have to be determined (optimal reagent and XL conditions) without generating extensive non-specific XL aggregates. To date, denaturing 1D SDS-PAGE is recommended to monitor and optimize XL conditions<sup>2</sup>, allowing visual observation of high mass bands on the gel corresponding to cross-linked species with concomitant vanishing of bands corresponding to individual free protein partners. Even if denaturing SDS-PAGE analysis is easy, widely available and robust, it has some limitations: i) low mass-resolution, ii) time consuming (24 hour time-scale from gel casting to protein staining), iii) not suited for high-masses or highly heterogeneous complexes which do not enter the gel.

To circumvent these limitations, we aimed at evaluating mass photometry (MP), a recently developed “single-molecule” biophysical technique<sup>3</sup> that gains interest in structural biology, for XL reaction optimization. MP is a versatile, fast and low sample-consuming technique that complements native MS for protein mass distributions measurements in native conditions. Despite its routine use to study native protein-protein interactions or multiprotein complexes oligomerization<sup>4</sup>, MP workflows are not adapted for the characterization of covalent assemblies.

We report here on the development of a suitable and efficient MP methodology in denaturing conditions (called dMP) for the straightforward characterization of covalent assemblies, with a special emphasis on its benefits for rapid XL reaction optimization. The developed workflow consists of a robust and efficient 2-step protocol that ensures 95% of irreversible denaturation within 5min in urea, just before MP measurement. After evaluating its efficiency on reference multimeric proteins (BSA, ADH, GLDH and 20S proteasome), we benchmarked dMP against the reference 1D SDS-PAGE gel migration to fine-tune protein XL reactions (impact of XL reagents, concentrations, etc.). dMP allowed a rapid screening of 20 different XL conditions including different commonly used reagents (PhoX, DSAU, DSSO, DSBU) within ~1.5 hour (with measurements in triplicates), highlighting that size and flexibility of the spacer arm are of utmost importance on XL efficiency. Finally, we successfully applied the dMP-based XL optimization within a complete XL-MS workflow that aimed to study the PPI of R2SP, a ~540 kDa complex involved in major cell functions as DNA repair, remodeling of chromatin or transcriptional regulation<sup>5</sup>.

By providing direct visualization along with relative quantification of all sub-complexes and aggregates, dMP can be uniquely positioned as the go-to method for more precise and accurate, mass-based optimization of best XL conditions. Consequently, optimal XL conditions (cross-linker, reaction time, temperature, cross-linker excess, pH value, etc.) can be rationally screened in a straightforward manner and with more accurate readouts (visualization of sub-complexes vs over-XL aggregates, relative quantification of the species, etc.). We envision dMP as a more precise and accurate alternative technique than 1D SDS-PAGE analysis for XL reaction optimization, which might be fully automatized in a near future.

(1) Piersimoni, L.; and al. *Chem. Rev.* 2021

(2) Iacobucci, C.; and al. *Anal. Chem.* 2019

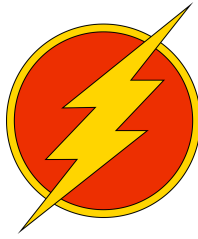
(3) Asor, R.; and al. *Curr. Opin. Chem. Biol.* 2022

(4) Soltermann, F.; and al. *Angew. Chem. Int. Ed.* 2020

(5) Maurizy, C.; and al. *Nat. Commun.* 2018







## Deciphering the signaling mechanisms of the C-type lectin receptor Dectin-1 in myeloid cells using high throughput proteomics

Dr Anne GONZALEZ DE PEREDO<sup>a</sup>, Rim ABDERRAHIM<sup>a</sup>, Alexandre STELLA<sup>a</sup>, Karima CHAOUI<sup>a</sup>, Marie LOCARD-PAULET<sup>a</sup>, Benjamin RAYMOND<sup>a</sup>, Thibaut SANCHEZ<sup>a</sup>, Véronique LE CABEC<sup>a</sup>, Frédéric LAGARRIGUE<sup>a</sup>, Odile SCHILTZ<sup>a</sup>, Olivier NEYROLLES<sup>a</sup>, Yoann ROMBOUTS<sup>a</sup>

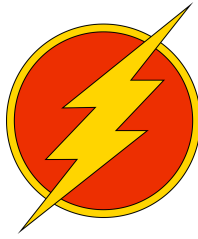
<sup>a</sup> CNRS, Institut de Pharmacologie et de Biologie Structurale

Dectin-1 is a Pathogen Recognition Receptor expressed on myeloid immune cells, known to play a major role in the antimicrobial immune response. It belongs to the C-type lectin Receptors (CLR) family and can bind to carbohydrate structures, such as  $\beta$ -glucans present on pathogens. This recognition triggers the phosphorylation of key tyrosine residues in the ITAM domain of Dectin-1, leading to the activation of SYK-dependent and independent signaling pathways, and finally to the secretion of pro-inflammatory cytokines and chemokines as well as to the production of reactive oxygen species and nitric oxide. Even though Dectin-1 is an important CLR known to be involved in many diseases such as fungal infections, inflammation and autoimmune diseases, its signaling mechanisms remain unclear. In this project we applied unbiased global proteomic approaches to explore Dectin-1 signaling and identify some potential medical targets. We used global phosphoproteomics to obtain a time-resolved characterization of the early signaling events taking place in the first minutes following Dectin-1 engagement, and identify key intracellular signaling proteins involved in that pathway. In addition, we established a high-throughput approach for secretome analysis, to monitor in a fast and global manner the inflammatory mediators secreted by immune cells 24 h after Dectin-1 stimulation.

To that aim, we generated ER-Hoxb8 immortalized murine cell lines from either WT or Dectin-1 KO progenitors, that were in-vitro differentiated into dendritic cells (DC). Cells were stimulated with depleted Zymosan, a canonical ligand for Dectin-1, during different time length. Phosphorylated peptides were enriched using a two-step protocol based on TiO<sub>2</sub> and phospho-Tyrosine antibody affinity purification, and resulting phosphopeptides were analyzed on an Orbitrap-Exploris instrument equipped with a FAIMS interface. To monitor the secretory phenotype of DCs, a high-throughput approach for secretome analysis was developed using the Evosep chromatographic system and DIA-PASEF analysis on a TIMS-TOF SCP mass spectrometer.

Using this analytical pipeline, we could globally detect more than 10,000 phosphosites with localization confidence  $>0.75$ , and identify after statistical analysis more than 300 phosphosites significantly regulated upon Dectin-1 stimulation, showing different kinds of kinetic profiles. These included sites from known proteins of the Dectin-1 signaling pathway, as well as novel targets that were selected for further functional validation by generating the corresponding KO cells. We could also provide a detailed characterization of the DC phenotypic response following Dectin-1 activation, with more than 500 proteins significantly over-represented in the secretome of activated cells, including many canonical inflammatory mediators such as TNF or IL6. The high-throughput secretomic analytical workflow based on the Evosep system and fast DIA-MS runs provides a quick and comprehensive read-out to monitor the phenotypic response of immune cells in many different condition. It thus allows the systematic screening of different KO cells to assess the importance of newly identified signaling proteins in on-going validation studies.



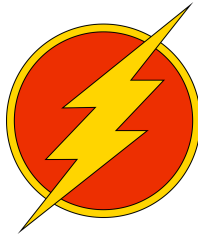


## Characterization of abundant periprostatic adipose tissue (PPAT) in prostate cancer cells: from inflammation-free expansion to over-secretion of endotrophin.

Dr Manuelle DUCOUX-PETIT<sup>a</sup>, David ESTÈVE<sup>a</sup>, Aurélie TOULET<sup>a</sup>, Mathieu ROUMIGUIÉ<sup>a</sup>, Dawei BU<sup>b</sup>, Sarah PERICART<sup>c</sup>, Chloé BELLES<sup>d</sup>, Cécile MANCEAU<sup>d</sup>, Cynthia HOUËL<sup>d</sup>, Nathalie VAN ACKER<sup>c</sup>, Stéphanie DAUVILLIER<sup>a</sup>, Yiyue JIA<sup>a</sup>, Camille FRANCHET<sup>e</sup>, Nicolas DOUMERC<sup>f</sup>, Mathieu THOULOZAN<sup>f</sup>, Sophie LE GONIDEC<sup>g</sup>, Philippe VALET<sup>g</sup>, Bernard MALAUAUD<sup>h</sup>, Odile BURLET-SCHILTZ<sup>a</sup>, Anne BOULOUMIÉ<sup>d</sup>, Philipp SCHERER<sup>i</sup>, Delphine MILHAS<sup>a</sup>, Catherine MULLER<sup>a</sup>

<sup>a</sup> IPBS, Institute of Pharmacology and Structural Biology, <sup>b</sup> Touchstone Diabetes Center, Department of Internal Medicine and Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA, <sup>c</sup> Department of Pathology, Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France, <sup>d</sup> Institut des Maladies Métaboliques et Cardiovasculaires, Université de Toulouse, INSERM, UPS, Toulouse, France, <sup>e</sup> Department of Pathology, Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France, <sup>f</sup> Institut des Maladies Métaboliques et Cardiovasculaires, Université de Toulouse, INSERM, UP, <sup>g</sup> Department of Urology and Renal Transplantation, Toulouse University Hospital ; Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France., <sup>h</sup> Institut RESTORE, Université de Toulouse, CNRS U-5070, EFS, ENVT, INSERM U1301, UPS, Toulouse, France., <sup>i</sup> Department of Urology and Renal Transplantation, Toulouse University Hospital ; Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France, <sup>i</sup> Touchstone Diabetes Center, Department of Internal Medicine and Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA.

Periprostatic adipose tissue (PPAT) abundance correlates with prostate cancer progression, but the mechanism remains unexplained. Here, we used a statistical approach to define abundant PPAT by normalizing PPAT volume to prostate volume in a cohort of 351 patients with a linear regression model. Applying this definition, we find that tumors specifically from patients with abundant PPAT exhibit several hallmarks of aggressiveness, suggesting that PPAT abundance might be used to improve risk stratification. We show that abundant PPAT expands by adipocyte hypertrophy but this does not result in inflammation. Extensive extracellular matrix remodeling, notably of the collagen network, and decreased expression of mechano-sensing proteins in adipocytes explains this inflammation-free expansion by decreasing the mechanical constraints on the adipocytes. Moreover, collagen VI degradation in abundant PPAT is associated with production of endotrophin, a matrikine that promotes cancer progression. We find high levels of endotrophin specifically in the urine of patients with abundant PPAT, indicating the clinical relevance of our findings.



FL5

## Flash MS/MS identification of microbial isolates

Madisson CHABAS<sup>a</sup>

<sup>a</sup> CEA

### Flash MS/MS identification of microbial isolates

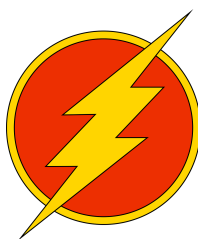
Madisson Chabas<sup>1,2</sup>, Jean-Charles Gaillard<sup>1</sup>, Olivier Pible<sup>1</sup>, Béatrice Alpha-Bazin<sup>1</sup>, Jean Armengaud<sup>1</sup>

<sup>1</sup> CEA, ProGénoMix platform; <sup>2</sup> Montpellier University

Rapid identification of microorganisms is vital for medical diagnostics, food product control, and screening for new enzymatic potentials for biotechnological applications. High throughput analytical platforms have been established in analytical laboratories to process thousands of samples per day. Furthermore, culturomics is very much in vogue to isolate and detail the components of the microbiota. MALDI-TOF MS technology has become the reference method for the rapid identification of microorganisms in microbiology for clinical diagnostic purposes, as it is a fast and inexpensive method. However, this method has some severe limitations because it requires that the information of a reference strain close to the sample is pre-registered in the database, and thus cannot be applied for most environmental isolates. Moreover, this technology does not allow distinguishing between closely related strains or is refractory to mixtures of microorganisms.

Tandem mass spectrometry proteotyping is more efficient because the method is able to identify the taxonomic position in the tree of life whatever the sample. By deciphering the protein sequences extracted from the samples, accurate taxonomic data can be obtained. The method we have developed, "Phylopeptidomics" (1), allows to quantify the relative biomass of the identified organisms (2-3) and is highly sensitive (4). We propose here a methodology for ultra-fast identification of microorganisms, which is amenable to very high throughput, based on direct sample infusion and adapted computer processing. This pipeline will be illustrated with results obtained on numerous isolates from skin microbiota culturomics. The performance of this methodology will be detailed in terms of analysis time required per isolate identified. These results open new perspectives for the application of MS/MS proteotyping of microorganisms.

1. Pible O, Allain F, Jouffret V, Culotta K, Miotello G, Armengaud J (2020) Estimating relative biomasses of organisms in microbiota using "phylopeptidomics". *Microbiome* 8:30.
2. Hayoun K, Pible O, Petit P, Allain F, Jouffret V, Culotta K, et al. (2020) Proteotyping Environmental Microorganisms by Phylopeptidomics: Case Study Screening Water from a Radioactive Material Storage Pool. *Microorganisms* 8:1525.
3. Lozano C, Kielbasa M, Gaillard J-C, Miotello G, Pible O, Armengaud J (2022) Identification and Characterization of Marine Microorganisms by Tandem Mass Spectrometry Proteotyping. *Microorganisms* 10:719.
4. Mappa C, Alpha-Bazin B, Pible O, Armengaud J (2023) Evaluation of the limit of detection of bacteria by tandem mass spectrometry proteotyping and phylopeptidomics. *Microorganisms*, in press.



## Pathway Grabber: Exploring proteomics data with KEGG pathways

Alexandre BUREL<sup>a</sup>, Patrick GUTERL<sup>a</sup>, Sarah CIANFERANI<sup>a</sup>, Christine CARAPITO<sup>a</sup>, Fabrice BERTILE<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO), Institut Pluridisciplinaire Hubert Curien (IPHC), UMR7178, Strasbourg, France ; Infrastructure Nationale de Protéomique ProFI - FR2048, 67087 Strasbourg, France

### Introduction

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource for understanding high-level functions of biological systems from large-scale molecular datasets. It has been developed together with a collection of tools for mapping molecular objects to KEGG pathway maps by the Kanehisa Laboratories ([www.kegg.jp](http://www.kegg.jp))

[www.kegg.jp](http://www.kegg.jp)

### Methods

Pathway Grabber was developed with Julia 1.8.5 and some Javascript. From an Excel file containing UniProt or KEGG identifiers and associated statistical scoring values, including as well p-values from single (e.g. t-tests) or multiple comparisons (e.g. Anova and post-hoc tests) and fold changes, and the definition by the user of thresholds for each parameter, proteins are categorized as “non-significant” or “significant” and the information “upregulated” or “downregulated” is kept. In parallel, KEGG annotations are downloaded for each protein, which allows proteins from the dataset to be distributed among Pathway maps. A cache mechanism has been added to reduce the amount of data transfer, hence making the tool faster and decreasing the impact on the bandwidth. Extracted data is provided in the form of an Excel file and a list of HTML files.

### Results

Output HTML files correspond to all Kegg Pathway maps that contain at least one protein from the dataset. On these maps, the information that has been made available is highlighted using a colour code relative to the ‘statistical category’ (for any item of interest, whether it is, e.g., a module, a protein/gene, a compound, a relation). Additional details about statistical scoring is also given as tooltip text items. All the items in the HTML files are clickable and reroute the user to the corresponding KEGG entries, but these files can also be used completely offline, once they are generated they do not require an Internet connection anymore.

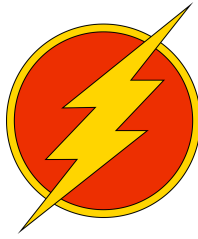
The output Excel file summarizes the whole information that is highlighted on maps, one sheet listing all the pathways associated with each with protein, another sheet listing all proteins associated to each revealed pathway.

### Conclusions

Pathway Grabber is a user-friendly tool for biologists and proteomists, helping them to get a comprehensive view of the molecular regulations from large lists of proteins, including the statistical metrics of differential analysis. Therefore, it allows to greatly accelerate omics data mining and it helps to nicely draw graphical representations of omics results (1, 2). Today, Pathway Grabber constitutes a solid basis for many future improvements (e.g. network analysis, extension to other resources like Reactome).

### References

1. Chazarin, B.; Benhaim-Delarbre, M.; Brun, C.; Anzeraey, A.; Bertile, F.; Terrien, J. Molecular Liver Fingerprint Reflects the Seasonal Physiology of the Grey Mouse Lemur (*Microcebus murinus*) during Winter. *Int. J. Mol. Sci.* **2022**, *23*, 4254. <https://doi.org/10.3390/ijms23084254>
2. Tascher, G.; Burban, A.; Camus, S.; Plumel, M.; Chanon, S.; Le Guevel, R.; Shevchenko, V.; Van Dorsselaer, A.; Lefai, E.; Guguen-Guillouzo, C.; Bertile, F. In-Depth Proteome Analysis Highlights HepaRG Cells as a Versatile Cell System Surrogate for Primary Human Hepatocytes. *Cells* **2019**, *8*, 192. <https://doi.org/10.3390/cells8020192>



FL7

## Comparison of tissue homogenization protocols for optimization of quantitative proteomics and phosphoproteomic analysis

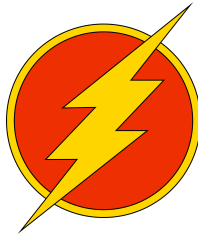
Marta ROJAS-TORRES<sup>a</sup>, Ana MARTINEZ-VAL<sup>b</sup>, Lucia BELTRAN-CAMACHO<sup>c</sup>, Sara ESLAVA-ALCON<sup>a</sup>, Jesper V. OLSEN<sup>b</sup>, M<sup>ª</sup> Carmen DURAN-RUIZ<sup>a</sup>

<sup>a</sup> Biomedicine, Biotechnology and Public Health, University of Cadiz / Inhibica, <sup>b</sup> Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, <sup>c</sup> Cell biology, Physiology and Immunology, Imibic / University of Cordoba

Tissue homogenization and cell lysis are the first steps in virtually all molecular biology and diagnostic techniques. The reproducible and efficient extraction of proteins from tissue biopsies for quantitative analysis is a critical step in biomarker and translational research, as it allows maximal recovery from limited quantity specimens and minimize the effects of intratissue heterogeneity. Tissues are specially challenging to phosphoproteomics analysis, since it is crucial to inactivate phosphatases that might be active during the homogeneization process. Nowadays, there are several methods and specific devices focused on the optimization of the abovementioned. In this regard, some of the most used are automatic dry pulverizer (Covaris <sup>®</sup>) and Precellys <sup>®</sup> 24 (Bertin technologies), alone or in combination with proteome stabilizer (Denator AB).

In the current study we have evaluated, by application of a label-free quantitative (LFQ) proteomic approach, the MS profiles obtained from different tissue samples (liver, kidneys, spleen, lungs, heart and brain) after homogenization using Precellys, Denator with Precellys and Covaris, in order to identify the optimal approach for each tissue when performing a total proteome and phosphoproteome quantitative and qualitative analyses. After the homogenization and protein extraction step, proteins were digested using Protein Aggregation Capture in the KingFisher platform. Moreover, phospho-peptide enrichment was performed in the KingFisher platform using Ti-IMAC HP beads. Samples were analysed on a Orbitrap Exploris (Thermo Scientific) coupled to an EvosepOne LC platform, operating in Data Independent acquisition (DIA). Spectronaut<sup>TM</sup> and Perseus software were used for data processing and statistical analysis.

We evaluated the performance on each homogeneization protocol in terms of overall protein and phospho-peptide coverage, as well as enrichment efficiency and applicability of the method for large scale studies. All methods were comparable in terms of protein coverage, as well as on phospho-peptide enrichment efficiency. On average, we quantified 5,000 proteins per tissue. The phospho-peptide coverage, as well as other PTMs such as oxidation, revealed higher tissue specificity reflecting the different physiology and signaling of each tissue. However, we did not find any consistent bias on any of the pipelines studied here. Several differences were found depending on the tissue and homogenization approach, especially from the processing time required in each pipeline. Quantification results regarding total proteome and phosphoproteome will be further discussed.



FL8

## A novel DIA label-free proteomics workflow to study protein glycation in plasma samples

Sarahi JARAMILLO ORTIZ<sup>a</sup>, Charlotte BRUN<sup>b</sup>, Fabrice BERTILE<sup>b</sup>, Francois CRISCUOLO<sup>c</sup>, Christine SCHAEFFER<sup>b</sup>

<sup>a</sup> IPHC-LSMBO - Université de Strasbourg/CNRS, 25 rue Becquerel 67087, Strasbourg, France/Infrastructure Nationale de Protéomique ProFI - FR2048, 67087, <sup>b</sup> IPHC-LSMBO - Université de Strasbourg/CNRS, 25 rue Becquerel 67087, Strasbourg, France/ Infrastructure Nationale de Protéomique ProFI - FR2048, 67087 Strasbourg, France, <sup>c</sup> IPHC-DEPE - Université de Strasbourg/CNRS, 23 rue Becquerel 67087, Strasbourg, France

Glycation is a post-translational modification (PTM) induced by the non-enzymatic reaction between reducing sugars and primary amino groups of proteins, resulting in the formation of Amadori compounds and advanced glycation end products (AGEs). Protein glycation has been associated with diabetes and aging-related pathologies<sup>1</sup>. Although very few proteomics studies have been conducted on glycation research, the preferred recurrent analytical platform to study this PTM is mass spectrometry with enzymatic digestion (bottom-up approach)<sup>2</sup>. Glycated proteins in plasma are relatively low abundant and comprehensive glycation studies require the enrichment of modified peptides coupled with advanced mass spectrometry-based methods. To date, no glycated proteome studies have been published in which data-independent acquisition (DIA) approach has been applied.

In this study, a label-free quantitative approach was developed to study glycations in undepleted plasma. Two sample preparation workflows were optimized. The first approach involved isopropanol gradient peptide fractionation (CIF) with paramagnetic beads to decomplexify the proteome prior to LC-MS/MS analysis. The second one sequentially combined CIF and automated boronate affinity chromatography (BAC) enrichment on an AssayMAP Bravo platform. Herein, we evaluated dia-PASEF methods on a timsTOF Pro platform.

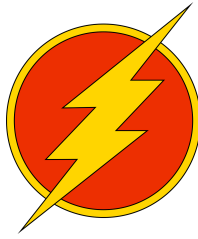
Commercial human plasma proteins (450 µg) were digested overnight using trypsin/Lys-C. In the first approach, CIF fractionation was performed on tryptic peptides and then analyzed by LC-MS. In parallel, for the second approach, digested peptides were sequentially fractionated and glycated peptides were enriched prior to LC-MS/MS analysis. CIF was performed using paramagnetic beads grafted with carboxylate groups; peptides were eluted into five fractions (F<sub>1</sub> ACN; F<sub>2</sub> 80, F<sub>3</sub> 60, F<sub>4</sub> 30, and F<sub>5</sub> 0% isopropanol). Custom BAC cartridges high capacity (BeneLab, Zagreb, Croatia) were used to perform automated enrichment on the AssayMap Bravo platform (Agilent). All samples were prepared in triplicate and analyzed by nanoLC-MS/MS using a TimsTOF Pro (Bruker) in DIA mode. Data were processed with Spectronaut 17.4 software (Biognosys) and DIA-NN 1.8.1<sup>3</sup>.

Our results showed that, in terms of total unmodified proteins and peptides, using DIA mode increases the number of identifications by a factor of two compared to previous experiments performed in DDA mode, albeit not when using sample fractionation.

Regarding modified peptides and proteins in control samples (non-fractionated and non-enriched); 7 times more peptides were identified in DIA compared to DDA acquisition mode and 78% more proteins. The DIA acquisition mode analysis of fractionated CIF samples enabled a 45% increase in the number of modified proteins and a 47% increase in the number of modified peptides and glycation sites recognized. The combined implementation of CIF fractionation and BAC enrichment of glycated peptides in DIA mode acquisition achieved to increase in the number of glycated peptides and glycation sites (twofold) as well as proteins (three-fold) compared to non-fractionated and non-enriched samples.

The bottom-up proteomic method developed has successfully increased the identification and quantification of numerous peptides and proteins modified by Amadori products in human undepleted plasma samples. To our knowledge, this is the first glycated proteome using a combination of fractionation/boronate-based enrichment strategies coupled with comprehensive DIA-MS. This approach will now be applied to samples from different avian species in order to better understand how birds manage their usually high levels of blood glucose.

1. A. Shin *et al.*, *Adv Clin Chem* 113:101-156 (2023).
2. C. Seonghyeon *et al.*, *Reviews in Analytical Chemistry* 41: 83-97 (2022).
3. V. Demichev *et al.*, *et al. Nat Methods* 17, 41-44 (2020).



## Identification of potential biomarkers related to A $\beta$ deposition through proteomic profiling in patients with Mild Cognitive Impairment

Vera M. MENDES<sup>a</sup>, Daniela MOUTINHO<sup>b</sup>, Sara C. MADEIRA<sup>c</sup>, Inês BALDEIRAS<sup>d</sup>, Manuela GUERREIRO<sup>b</sup>, Sandra CARDOSO<sup>b</sup>, Helena AIDOS<sup>c</sup>, Isabel SANTANA<sup>e</sup>, Alexandre MENDONÇA<sup>b</sup>, Bruno MANADAS<sup>a</sup>

<sup>a</sup> CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal ; Centre for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Coimbra, Portugal, <sup>b</sup> Faculty of Medicine, University of Lisbon, Lisboa, Portugal, <sup>c</sup> LASIGE, Faculty of Sciences, University of Lisbon, Lisboa, Portugal, <sup>d</sup> CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal ; Centre for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Coimbra, Portugal ; Faculty of Medicine, University of Coimbra, Coimbra, Portugal, <sup>e</sup> Faculty of Medicine, University of Coimbra; Department of Neurology, Hospital and University Centre of Coimbra; Center for Neuroscience and Cell Biology ; Centre for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Coimbra, Portugal

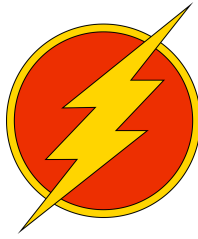
Alzheimer's disease is the leading type of dementia, characterized by a gradual decline in cognitive and behavioral function. It usually starts with mild memory loss, progressing over time to difficulties with communication and response to the environment. Gaining insight into the underlying causes of this disorder is critical for enabling early and precise diagnosis, as well as developing effective treatments. Mild cognitive impairment (MCI) is a condition that often precedes AD, and studying MCI can help identify early signs and risk factors for AD, allowing for earlier intervention and potential prevention of the disease. A comparative proteomic study was carried out between an MCI group of patients without any biomarkers of A $\beta$  deposition or neuronal injury (MCI<sub>A $\beta$ -</sub>) and an MCI group of patients fulfilling criteria of high likelihood for AD (MCI<sub>A $\beta$ +</sub>). In order to enhance the reliability of the results and avoid limitations to a specific patient group, this study employed analysis and comparison of samples obtained from two distinct cohorts. The protein content from 60 CSF samples from MCI<sub>A $\beta$ -</sub> patients and 68 from MCI<sub>A $\beta$ +</sub> patients was precipitated with methanol and analyzed using the Short-GeLC-SWATH approach. Pooled samples were prepared to generate the peptide ion libraries by Data Dependent acquisition (DDA) and each sample was analyzed by SWATH/DIA for relative protein quantification.

Results show that, out of the 617 proteins in the library, 517 were quantified and subjected to statistical analysis to identify the most significant proteins that distinguished between the two MCI groups. This analysis was conducted separately for each cohort and later for all samples. Differences between the two cohorts were identified through PCA analysis, and to mitigate any potential technical variations, batch effect correction was conducted. This facilitated a reassessment of the observed disparities between the two cohorts.

The most interesting proteins related to A $\beta$  deposition (MCI<sub>A $\beta$ +</sub>) were selected based on a VIP score above 1, extracted from the PLS-DA analysis. Although the number of selected proteins was similar between both cohorts (around 160 proteins), the proteins themselves differed significantly. Gene Ontology enrichment analysis was conducted on the identified sets of interesting proteins. While some biological processes, such as cell adhesion and post-translational protein modification, were common to both cohorts, a large number of biological processes were found to be enriched in only one of the cohorts.

Despite the differences among the selected proteins from each cohort, the combination of all the analyses resulted in a group of proteins that distinguishes MCI<sub>A $\beta$ -</sub> from MCI<sub>A $\beta$ +</sub> and that can be considered potential biomarkers for Alzheimer's disease.





## Optimization of sample preparation and nanoLC-MS/MS methods for single cell proteomic analysis

Pauline PERDU-ALLOY<sup>a</sup>, Laurence SABATIER<sup>a</sup>, Christine CARAPITO<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique, IPHC (UMR 7178), CNRS, 67000 Strasbourg, France

Understanding the various mechanisms in which proteins are involved is crucial for comprehending biochemical phenomena, including the onset and development of diseases<sup>1</sup>. Conducting qualitative and quantitative analysis of the entire proteome can provide insights into the response of an organ or a population of cells to treatments or diseases. However, analysis of proteins from an average cell population often neglects cellular heterogeneity and does not allow to distinguish healthy cells from infected cells, which cohabit in the same tissue<sup>2</sup>. The rise of single-cell proteomics is therefore meaningful because it allows the identification of the proteome expressed in individual cells with increasing depth at a given time<sup>3</sup>, with applications ranging from disease diagnosis to the response of a chemical treatment. However, despite its importance, the proteome analysis at the single-cell level remains a major analytical challenge due to the very low abundance of protein material available within a cell (<ng). Thus, every step of the workflow needs to be revisited to minimize losses during sample preparation and achieve highest sensitivity in liquid chromatography tandem mass spectrometry analysis (LC-MS/MS).

In recent years, the development of trapped ion mobility spectrometry (TIMS) coupled to a high sensitivity mass spectrometer (TimsQTOF) as well as the development of automated methods for sample preparation and sophisticated data software constitute an anchor for the analysis of protein material at the single cell scale<sup>4</sup>.

On the one hand, Hela cell digests (Pierce, Ref. 88329) were used to prepare a dilution series from 0.1 to 10 ng of injected material. On the other hand, HeLa cells were sorted, lysed and digested using the CellenONE instrument at Cellenion (Lyon). Varying cell numbers, from 1 to 20 cells were sorted, lysed, digested prior to nanoLC-MS/MS analysis.

Peptide mixtures were separated on an Aurora C18-RP 120Å (75µm x 250mm, 1.9µm) (IonOpticks) column using a nanoElute (Bruker Daltonics) coupled to a timsTOFpro2 (Bruker Daltonics) mass spectrometer. Separation was performed on a gradient from 2 to 30% ACN in 0.1% formic acid over 30min at 0.2µL/min. Data Dependent Acquisition - Parallel Accumulation and Serial Fragmentation (DDA-PASEF) and Data Independent Acquisition - PASEF (DIA-PASEF) methods were used. A comparative analysis was performed to assess the performance of two data analysis tools, MaxQuant and Proline for DDA data analysis, and Spectronaut and DIA-NN for DIA data interpretation.

In DDA-PASEF mode, around 300 proteins were identified in the 0.1 to 0.3ng region, while more than 1000 proteins could be identified in the same region in DIA-PASEF. These first results were encouraging to further explore the instrument's performances for very low amounts as these protein amounts are close to single cell protein contents (approx. 0.3ng). In isolated cells, an average of 540 proteins were identified in single cells in DDA-PASEF mode, while more than 2400 proteins were identified using a DIA-PASEF method. This remarkable outcome highlights the great capacities of DIA-PASEF for single-cell analysis.

Investigating the data processing aspect of DIA, comparing the two data processing software (Spectronaut and DIA-NN) revealed close average numbers of quantified proteins for single cells and up to 5 isolated cells.

While those initial results are promising, further optimizations are still required from sample preparation to nanoLC-MS/MS acquisition methods optimization and finally data processing to be able to unlock deeper insights into cellular heterogeneity and disease mechanisms.

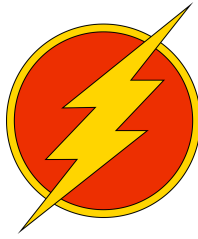
<sup>1</sup> Minakshi, *Single-Cell Proteomics: Technology & Applications*. (2019).

<sup>2</sup> Marino, Zito, Accardo, Ronchi, Cozzolino, Morgillo, Rossi, Franco, *Molecular heterogeneity in lung cancer*. (2019).

<sup>3</sup> Templer, Ces, *New frontiers in single-cell analysis*. (2008).

<sup>4</sup> Meier, Park, Mann, *Trapped Ion Mobility Spectrometry and Parallel Accumulation-Serial Fragmentation in Proteomics*. (2021).





FL11

## Promising Salivary Protein Markers for Head and Neck Squamous Cell Carcinoma Prognosis

Dr Rui VITORINO<sup>a</sup>

<sup>a</sup> UNIVERSITY OF AVEIRO

Oriana Barros<sup>1,2</sup>, Vito G. D'Agostino<sup>3</sup>, Lucio Santos<sup>2</sup>, Rita Ferreira<sup>4</sup>, Rui Vitorino<sup>1,2,5\*</sup>

<sup>1</sup> Department of Medical Sciences, Institute of Biomedicine iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal

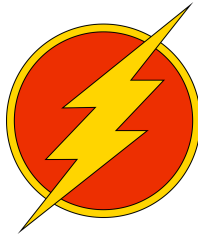
<sup>2</sup> Experimental Pathology and Therapeutics Group, Research Center of IPO Porto (CI-IPOP) / RISE@CI-IPOP (Health Research Network) and Surgical Department of Portuguese Oncology Institute of Porto (IPO Porto) / Porto Comprehensive Cancer Center (Porto.CCC)

<sup>3</sup> Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Via Sommarive 9, 38123, Trento, Italy

<sup>4</sup> LAQV-REQUIMTE, Department of Chemistry, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.

<sup>5</sup>UnIC, Department of Surgery and Physiology, Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal

*Head and neck squamous cell carcinoma (HNSCC) is one of the most common malignant cancers in the world. In advanced stages, the mortality rate is high. Stratification of HNSCC patients by molecular characteristics allows optimization of clinical management of these patients. Here, we explored the GEO and PRIDE databases to identify putative prognostic biomarkers for HNSCC. Thus, an integrated transcriptome analysis was performed to identify the major differentially expressed genes in HNSCC tissues (using GSE12452, GSE13597, GSE31056, GSE6631 and GSE3524 GEO datasets) in combination with a proteomic analysis to identify the overrepresented proteins in saliva samples from HNSCC patients (using PRIDE PXD012436 dataset). The panel of identified biomarkers was characterized using web tools, namely UALCAN, ToPP and PINAv3.0. From the combinatorial analysis of these Omics data, the salivary biomarkers with the greatest potential for clinical application for prognostic stratification were identified: ADH7, MMP9, and S100A14. Overall, this Multi-Omics approach provides comprehensive information on the potential value of combining ADH7, MMP9, and S100A14 as prognostic biomarker panel for risk stratification. Future studies should validate this panel of biomarkers in saliva samples from a large cohort of patients using targeted approaches, envisioning its translation into the clinical setting.*



## LC-MS/MS analysis of penguin tissues and customization of protein database to identify novel antimicrobial peptides

Bastien ARNAUD<sup>a</sup>, Adrien BROWN<sup>b</sup>, Philippe BULET<sup>c</sup>, Fabrice BERTILE<sup>b</sup>

<sup>a</sup> Inovarian, Paris, France, <sup>b</sup> BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC UMR 7178, University of Strasbourg, CNRS, Strasbourg, France, <sup>c</sup> Institute for Advanced BioSciences, IAB UMR 5309, CNRS, INSERM, Université Grenoble Alpes, La Tronche, France

### Introduction

According to the WHO, the clinical pipeline and recently approved antibiotics are insufficient to tackle the challenge of increasing emergence and spread of antimicrobial resistance. The identification and improvement/development of natural antimicrobial peptides (AMPs) is a promising alternative to conventional antibiotics, due to their low probability of inducing resistance. In the early 2000s, we have isolated and fully characterized two isoforms (AvBD103a, AvBD103b) of an AMP in the King penguin (*A. patagonicus*), namely the avian  $\beta$ -defensin spheniscin<sup>1</sup>. We showed that AvBD103b is active against micro-organisms (G+, G- bacteria, yeast and filamentous fungi) even at high salt concentration<sup>2</sup>. This is a considerable advantage to fight the pathogens that develop in salt-rich environments (e.g. during cystic fibrosis or ocular infections). Therefore, our ultimate goal is to decipher the mechanisms of action of spheniscin. Here we seek to analyze the structural diversity of AMP members of the spheniscin family by refining the annotation of the King penguin genome analyzing and several penguin tissue samples by mass spectrometry.

### Methods

Penguin samples selected based on their septic and saline environment (bone, lung, kidney, mouth floor, tongue, stomach, esophagus, duodenum and jejunum) were cryogrinded using a ball mill, and proteins were extracted in urea/thiourea/CHAPS or Laemmli buffer. After protein assay, 50  $\mu$ g of proteins were loaded onto glycine SDS-PAGE and tricine SDS-PAGE gels and electrophoresed. Eight to ten protein bands of 2mm width were excised in the 1-10kDa range, and trypsin *in-gel* digestion was performed after reduction and alkylation. Samples were analyzed using a nanoElute chromatography system coupled to a timsTof Pro mass spectrometer (Bruker). MS raw data were processed using Mascot (v2.8.1) and ProlineStudio (v2.1.2) with an FDR of 1% at PSM and protein levels. Spectra were searched against a home-made protein database containing protein sequences available from three different annotations of the *A. patagonicus* genome. We also added to the database protein sequences derived from our own annotation of the *A. patagonicus* genome and comparison (blast) with sequences of all known beta-defensins in specialized resources (APD and DRAMP). In addition, a de novo sequencing strategy was considered using Novor (v1.1; min peptide score 50) and blastp v2.12.0.

### Results

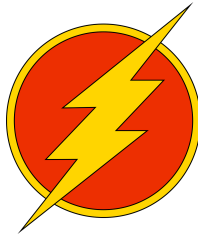
We successfully built a protein database enriched with 29 putative penguin beta-defensins in addition to the 8 already annotated. From there, sequence clustering and alignment allowed us to eliminate redundancies, and we finally retained a database version with 25 penguin defensin sequences. Strikingly, retained sequences were rich in well-aligned cysteine motifs, which supports their actual defending status. Mascot searches allowed 10 penguin defensins to be identified, and the de novo sequencing strategy allowed detection of an additional one. Defensin sequence coverage was determined in the 9-47% range, which is unreasonably low and will be corrected, due to the fact that the sequences we predicted were artificially elongated to avoid missing the true 'start' and 'stop' of penguin defensins. Tissue specificity could be noticed, the bone, lung and tongue already bringing the identification 10 of the 11 defensins.

### Conclusion/perspectives

The combination of bioinformatics manipulation of protein sequence databases, mass spectrometry, classical Mascot analysis, and de novo sequencing proved relevant and successful. Further analyses should increase the coverage of penguin defensin sequences. These results will be used to better understand the structure/function relationship of penguin defensins. In parallel, we will conduct a bio-guided strategy to evaluate which of the 11 defensins we identified are actually active against microorganisms under high salinity conditions.

<sup>1</sup>Thouzeau et al., J Biol Chem 2003, DOI:10.1074/jbc.M306839200

<sup>2</sup>Landon et al., J Biol Chem 2004, DOI:10.1074/jbc.M401338200



## Short modified peptides or dimeric peptide isolated from animal venoms: Are LC-MS-MS, Edman sequencing enough to determine their sequences?

Yasmine BOUGHANMI<sup>a</sup>, Pascal MANSUELLE<sup>b</sup>, Regine LEBRUN<sup>b</sup>, Patrick FOURQUET<sup>c</sup>, Marielle BALZANO<sup>d</sup>, Caroline BERENQUER<sup>e</sup>, Maxime MOULARD<sup>d</sup>, Harold DE POMYERS<sup>f</sup>, Didier GIGMES<sup>g</sup>, L'houcine OUAFIK<sup>e</sup>, Kamel MABROUK<sup>g</sup>

<sup>a</sup> Aix-Marseille Univ CNRS ICR UMR 7273, Équipe CROPS, <sup>b</sup> Aix Marseille Univ, IMM Institut de Microbiologie de la Méditerranée (Marseille Protéomique), FR 3479, Marseille, France, <sup>c</sup> Aix-Marseille Univ, INSERM, CNRS, Institut Paoli-Calmettes, CRCM, 13009 Marseille Protéomique, Marseille, France, <sup>d</sup> BioCytex, Marseille, 13010, France, <sup>e</sup> Aix Marseille Univ, CNRS, Institut de Neurophysiopathologie (INP), Marseille, France, <sup>f</sup> LATOXAN SAS, 845 Avenue Pierre Brossolette, 26800 Portes-lès-Valence, France, <sup>g</sup> Aix-Marseille Univ CNRS ICR UMR 7273, Équipe CROPS, 13397 Marseille, France

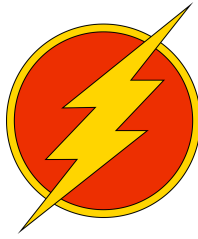
The research for new therapeutic molecules is a major public health issue, more particularly in the treatment of certain diseases. In our group, venoms have been screened in several areas in order to identify new painkillers and mainly anti-tumors or antibacterial agents.

So far, the success of cancer treatments has not been satisfactory. Indeed, synthetic chemicals used in chemotherapy and radiotherapy have low solubility and sometimes extremely high cytotoxicity that can cause significant collateral damage to normal tissues. In addition, they can induce chemoresistance or radioresistance after a period of treatment. It is therefore essential to develop new molecules capable of blocking the division of cancer cells and having less toxic effects on healthy tissues. Many anti-cancer peptides (ACPs) have the ability to specifically and selectively target cancer cells.

Moreover, natural resistance of bacteria consecutive to overuse of antibiotics in human and animal therapies as well as in various industrial fields have led to the selection of pathogenic multi- drug resistant bacteria. Over the last few decades, resistance to conventional antibiotics has limited treatment options, resulting in a significant increase in mortality and morbidity in hospitals. Venoms are known to be a rich source of unique and innovative pharmacologically active substances, including ACPs or AMPs.

In our study, the antibacterial or antiproliferative properties of 200 venoms of snakes, scorpions, amphibians and insects were studied. The identification of these AMPs or ACPs was carried out by: i) tests for the inhibition activity ii) HPLC for the purification of compounds from fractions found active and iii) mass spectrometry conjugated with amino acid sequencing for sequences/structures determination.

The difficulties encountered when determining 2 sequences of bioactive peptides, one short and is modified and the other long and is dimeric will be discussed.



## Comparison of *Bacillus thuringiensis* virulence potential on human Caco-2 intestinal cells using multi-omics investigation

Sarah ALILAT<sup>a</sup>, Rachele LANCELEUR<sup>b</sup>, Mathilde BONIS<sup>c</sup>, Olivier FIRMESSE<sup>c</sup>, Gaëlle VIAL<sup>d</sup>, Xavier SARDA<sup>d</sup>, Valérie FESSARD<sup>b</sup>, Sophie LIUU<sup>c</sup>, Thibaut LÉGER<sup>b</sup>

<sup>a</sup> Toxicology of Contaminants Unit, National Agency for Food, Environmental and Occupational Health & Safety, <sup>b</sup> Toxicology of Contaminants Unit, National Agency for Food, Environmental and Occupational Health & Safety (ANSES, Fougères), <sup>c</sup> Staphylococcus, Bacillus and Clostridium Unit, Food Safety Laboratory, National Agency for Food, Environmental and Occupational Health & Safety (ANSES, Maisons-Alfort), <sup>d</sup> Direction of Regulated Product Evaluation (DEPR), Residue and Food Safety Unit (URSA), National Agency for Food, Environmental and Occupational Health & Safety (ANSES, Maisons-Alfort)

The soil bacterium *Bacillus thuringiensis* (Bt) can produce spores and crystals composed of toxins with insecticidal activity. Due to these properties, some Bt strains have been developed as biopesticides for pest control, with a high value to restrict the use of chemical insecticides. However, Bt strains belong to the group of *Bacillus cereus* (Bc) and share some virulence genes encoding diarrheal enterotoxins such as hemolysin BL (Hbl), as well as non-hemolytic enterotoxin (Nhe) and cytotoxin K (CytK2), known to be involved in foodborne outbreaks (FBOs) following Bc contamination. For this reason, it has been suggested that Bt may also provoke FBOs and it is therefore crucial to evaluate their virulence potential and to characterize their effects on humans.

For this purpose, differentiated human intestinal Caco-2 cells have been treated with serial dilutions of the culture supernatant of five different Bt strains used as biopesticides. The cytotoxicity (MTT assay) and inflammatory response (interleukin-8 release) were first measured after a 24 h treatment. No cytotoxicity was induced with all the Bt supernatants up to a 25% dilution in cell medium, but the secretion of interleukin 8 was significantly induced.

To investigate the molecular effects of Bt supernatants on Caco-2 cells, untargeted proteomics and metabolomics were performed on both cell media and lysates after a 24h treatment with 3 selected non-cytotoxic concentrations. Peptides and metabolites were detected using a Q Exactive Plus mass spectrometer coupled to a Vanquish micro-UHPLC system (Thermo Fisher Scientific) and data were analyzed using Maxquant and Compound Discoverer tools. The pathways affected will be compared for all the Bt strains.

In conclusion, this project will better characterize the toxic effects induced by a panel of Bt strains on the human intestinal barrier. The relationship between the production of toxins and the pathways affected will be further investigated.

# Poster Presentations



## Characterization of brain $\alpha$ -synuclein in neurodegenerative diseases by mass spectrometry

Kim-Thanh VAN<sup>a</sup>, Susana BOLUDA CASAS<sup>b</sup>, Nicolas VILLAIN<sup>b</sup>, Foudil LAMARI<sup>b</sup>, François FENAILLE<sup>a</sup>, François BECHER<sup>a</sup>

<sup>a</sup> CEA Paris-Saclay, <sup>b</sup> Hôpitaux Universitaires Pitié Salpêtrière

### Introduction

$\alpha$ -synuclein is the main protein involved in the pathology of a number of neurodegenerative diseases, including Parkinson's disease, dementia with Lewy bodies (LBD), multiple system atrophy, and about half of Alzheimer's disease cases [1]. A primary hallmark of these synucleinopathies is the presence of Lewy bodies - insoluble particles made of proteinaceous materials in which  $\alpha$ -synuclein is the main component [2]. The literature has reported numerous post-translational modifications (PTMs) along the sequence of  $\alpha$ -synuclein, including phosphorylation, truncation, nitration, ubiquitination, acetylation, nitration, O-GlcNAcylation, and SUMOylation [3-5]. Most of these PTMs, however, were detected using site-specific antibodies *in vitro* or on animal models. In this study, we developed a novel immunoprecipitation method combined with nanoLC-HRMS/MS to identify and characterize the proteoforms of  $\alpha$ -synuclein from post-mortem brains of patients suffering from LBD.

### Methods

We used frozen brain samples (temporal cortex) from two LBD patients and two control cases. Brain tissue (~100 mg) was homogenized in TBS buffer containing protease and phosphatase inhibitors, using a Dounce homogenizer. We then used ultracentrifugation (100,000 g for 1 h) to fractionate brain proteins into three fractions using buffers with increasing stringency: soluble fraction (TBS), detergent-soluble fraction (2% Triton-X-100), and detergent-insoluble fraction (8 M urea + 2% SDS). Then, we performed immunoprecipitation with a panel of commercial  $\alpha$ -synuclein antibodies and analyzed in parallel by nanoLC-HRMS/MS and Western blot. Data analysis was done with Proteome Discoverer 2.5.

### Results

Western blots using pan  $\alpha$ -synuclein antibody showed that there were differences in the pattern of  $\alpha$ -synuclein signals between the three fractions and between disease and control brains. While truncations seemed to appear ubiquitously, high molecular weight bands were uniquely visible in the detergent-insoluble fraction of disease brains. We also observed that different antibodies recognized different forms of  $\alpha$ -synuclein. Therefore, we first optimized the immunoprecipitation protocol with different types of magnetic beads, capturing antibodies with different epitopes, elution buffers, and digestion enzymes. Evaluation was based on Western blots and bottom-up mass spectrometry.

On going bottom-up MS analysis results already showed phosphorylation at Serine 129 and truncations at residues 65, 66, and 101 in both the insoluble and insoluble fractions of brain extract, highlighting the potential pathophysiology role of these modified  $\alpha$ -synuclein species. Our future plan is to identify and quantify these PTMs in cerebrospinal fluid of LBD patients and controls in large cohort, using targeted MS method (parallel reaction monitoring). We expect that these PTMs could become robust biomarkers for the diagnosis of LBD from the early year and support the classification of LBD amongst other synucleinopathies.

### References

1. Serratos IN, Elizabeth HP, Carolina C, et al.; Mol Neurobiol **59** (1): 620-42 (2022).
2. Manzanza N de O, Sedlackova L, Kalaria RN; Front Aging Neurosci **13** (2021).
3. Schmid AW, Fauvet B, Moniatte M, et al.; Mol Cell Proteomics **12**: 3543-3558 (2013).
4. Oliveira LMA, Gasser T, Edwards R, et al.; Npj Park Dis **7**: 1-23 (2021).
5. Chen R, Gu X, Wang X; Clin Chim Acta **529**: 76-86 (2022).

## Reproducible and scalable one-day FFPE sample preparation method for LC-MS assays with Adaptive Focused Acoustics®

Dr Nicolas AUTRET<sup>a</sup>

<sup>a</sup> Covaris

Translational or clinical laboratories require automated, hands-off solutions for protein sample preparation to provide reproducibility, increased efficiency, higher quality results, and faster turnaround time. This poster highlights simultaneous, one-day, multi sample processing from formalin-fixed and paraffin-embedded (FFPE) tissue for liquid chromatography-mass spectrometry-based (LC-MS) analysis, using focused ultrasonication or AFA.

Protocols are perfectly suited for batches from 8 [1] to 96 samples [2] and were designed to allow robust single pot sample preparation, as previously described : those highly reproducible methods work with other sample types like cells and fresh/frozen tissues [3], with long or short [4] LC-MS gradients, and with every clean-up and digestion method [5]. When working with bead-based protein aggregation capture [2], the workflow can now include an accelerated trypsin digestion step which yields similar quality compared to overnight digestion, in a more controlled way, and leads to 96 samples processed in 7 to 8 hours [6].

In summary, AFA-based processing:

- makes FFPE tissue proteomics compatible with the clinical setting, through a non-toxic deparaffinization and rehydration approach,
- facilitates workflow simplification by allowing homogenization through digestion in a single well with an automation compatible high-throughput (HTP) workflow,
- allows an increased throughput and decreased turn-around-time,
- makes one-day sample prep also fully applicable to laser capture microdissections (LCM) [5,7-8] and fresh/frozen tissue preparation [3,4].

1. Automated Protein Extraction from Clinical Samples with the Covaris ML230 Focused-ultrasonicator. Staeber et al., Covaris Application note, [https://www.covaris.com/wp/wp-content/uploads/resources\\_pdf/m020155.pdf](https://www.covaris.com/wp/wp-content/uploads/resources_pdf/m020155.pdf)
2. AFA-sonication Followed by Modified Protein Aggregation Capture (APAC) Enables Direct, Reproducible and Non-toxic Sample Preparation of FFPE Tissue for Mass Spectrometry based Proteomics. Schweitzer et al., Covaris Application note, <https://www.covaris.com/wp/wp-content/uploads/2020/06/M020141.pdf>
3. Automated sample preparation with SP3 for low-input clinical proteomics. Mueller et al., Mol Syst Biol (2020) DOI: 10.15252/msb.20199111
4. Dietary-challenged mice with Alzheimer-like pathology show increased energy expenditure and reduced adipocyte hypertrophy and steatosis. Schreyer et al., Aging 2021, Vol. 13, N° .8, P10891-10919. DOI: 10.18632/aging.202978
5. Scalable FFPE Sample Preparation Method for Clinical Proteomics LC-MS Assays with Laser capture microdissection coupled mass spectrometry (LCM-MS) for spatially resolved analysis of formalin-fixed and stained human lung tissues. Herrera et al., Clin Proteom (2020) 17:24
6. Adaptive Focused Acoustics® Dr Jessica Chapman, MSKCC - Covaris Seminar at ASMS 2022 <https://www.youtube.com/watch?v=pFAUnk3aXBo>
7. Proteome analysis of formalin-fixed paraffin-embedded colorectal adenomas reveals the heterogeneous nature of traditional serrated adenomas compared to other colorectal adenomas. Sohler et al., Journal of Pathology, 2020, 250, 3, 251-261
8. Novel Laser Capture Microdissection-Proteomic Analysis Identifies Spatially Distinct Extracellular Matrix Signatures in the Core and Infiltrating Edge of Human Glioblastoma. Pedley et al., Biorxiv 10.1101/2022.09.01.506199



## Innovative high-throughput plasma proteome profiling facilitates in-depth, rapid and robust analysis

Dr Chloé MORITZ<sup>a</sup>, Claudia MARTELLI<sup>b</sup>, Katharina LIMM<sup>c</sup>, Andreas SCHMIDT<sup>d</sup>, Fabian WENDT<sup>e</sup>, Zehan HU<sup>c</sup>, Xaver WURZENBERGER<sup>c</sup>, Katrin HARTINGER<sup>c</sup>, Gary KRUPPA<sup>d</sup>, Sebastian MUELLER<sup>f</sup>, Manuel BAUER<sup>e</sup>, Nils KULAK<sup>c</sup>

<sup>a</sup> PreOmics GmbH, <sup>b</sup> Bruker Switzerland AG, Fällanden, SWITZERLAND, <sup>c</sup> PreOmics GmbH, Planegg, GERMANY, <sup>d</sup> Bruker Daltonics GmbH & Co. KG, Bremen, GERMANY, <sup>e</sup> Tecan Group Ltd, Männedorf, SWITZERLAND, <sup>f</sup> Biognosys AG, Zürich, SWITZERLAND

### Introduction

Blood plasma is one of the least invasive biopsies and a precious sample for clinical research and patient health monitoring. Almost all tissues are supplied by a constant blood flow in which proteins are constantly secreted or leaked. This allows the plasma to provide comprehensive information on the state of health or disease. Today the discovery of protein biomarkers in plasma represents an exceptional challenge since the complexity and huge dynamic range of about ten orders of magnitude complicates in-depth profiling at the throughput required for the study of large cohorts. To tackle this challenge, we have developed a fast, robust and fully automatable LC-MS-based plasma workflow for high-throughput proteomics studies that enriches for low-abundant proteins and provides deeper coverage of the plasma proteome.

### Methods

The plasma sample preparation workflow showcased is based on the new ENRICH technology which allows efficient dynamic range compression by enriching low abundant proteins on paramagnetic beads. Starting with 20 µl of plasma, further protein preparation consists of denaturation, reduction and alkylation on beads, tryptic digestion and peptide clean-up using the biofluid optimized iST-BCT protocol (PreOmics). The kit can be used manually or in combination with automated liquid handling platforms. Results generated on the Tecan Fluent® liquid handling system with the Resolvex® A200 positive pressure module will be presented. Peptides (300ng) were analyzed by nLC-MS/MS using mass spectrometers of the TimsTOF family using the diaPASEF® acquisition mode (Bruker). Data processing was performed with either DIA-NN or Spectronaut (Biognosys).

### Results

Parallel processing of 96 samples using the new ENRICH technology in combination with the iST-BCT sample preparation takes ~5 hours. It was performed manually and automatically leading to a 2-fold increase in the number of proteins identified compared to neat plasma. Excellent intra-plate repeatability could be demonstrated with both manual and automated protocols with a median CV <20% within technical replicates.

To assess the effect of donor diversity and pre-analytical parameters, we analyzed blood plasma from four donors collected with four different anticoagulants for each donor (EDTA, CTAD, Citrate, Heparin). A strong influence of the anticoagulants on the depth of the proteomic analysis could be observed. Plasma from the same donor collected with sodium citrate or EDTA resulted in a delta of ~400 additional proteins identified in favor of EDTA. These results highlight the importance of experimental design, from planning to optimized and standardized biofluid collection. Initially developed for the analysis of human plasma, we also explored the possibility of applying the ENRICH technology to serum samples as well as to plasma from model animals such as mice and rats.

Finally, the ENRICH-iST kit was compared to the iST-BCT kit to prepare a clinical cohort of human plasma from lung cancer patients (n=10) and age-matched healthy donors (n=10). The samples were spiked with the PQ500™ reference panel (Biognosys). Without ENRICH, 350 protein groups per sample were recovered, covering 405 protein groups represented in the PQ500 panel. From the ENRICH-treated samples, we obtained an average of 1,000 protein groups per sample. Statistical analysis of the quantified proteins showed a clear stratification between healthy donors and patients. Strongly upregulated factors in the lung cancer samples included S110A8 and S100A9 which have previously been reported to be upregulated in lung cancer and other types of cancer (<https://doi.org/10.1016/j.bcp.2006.05.017>).

### Conclusion

The ENRICH-iST kit is a cutting-edge, all-in-one workflow for low abundance protein enrichment and bottom-up





proteomic sample preparation. It allows the study of clinically relevant biofluids from various species at high throughput, in a robust, fast and easy-to-use manner.



## Development of a proteomic approach for samples preserved in Formalin-Fixed Paraffin-Embedded and collected by laser microdissection for the study of the evolution of the tumorigenesis sequence of malignant peripheral nerve sheath tumors.

Cedric BROUSSARD<sup>a</sup>, Mohamed-Amin MATHLOUTHI<sup>a</sup>, Morgane LE GALL<sup>a</sup>, Eric PASMANT<sup>a</sup>, Maryline FAVIER<sup>a</sup>, Emilie-Fleur GAUTIER<sup>a</sup>, Pierre SOHIER<sup>a</sup>

<sup>a</sup> INSTITUT COCHIN

Proteomic analysis by mass spectrometry of biopsies in Formalin-Fixed Paraffin-Embedded (FFPE) blocks is possible and is closed to the technique on frozen tissue. However, biological questions may require to study separately several areas of interest within the same sample or a restricted area. This is the case with the progression of tumorigenesis sequence of Malignant Peripheral Nerve Sheath Tumors (MPNST) where several distinct areas can be observed on a single section.

From published data, we have implemented a protocol for the proteomic analysis of FFPE tissues with selection of areas by laser microdissection adapted to the Arcturus XT laser microdissector. The analysis is carried out on sampled areas with an average surface area of 10 mm<sup>2</sup>. We carried out various developments including the preparation of sections with a microtome, the sampling with the laser microdissector and the preparation of samples for nLC-MS/MS proteomic analysis as well as mass spectrometry data acquisition methods (DDA / DIA).

After these developments, we selected the following protocol: after microdissection of 10 mm<sup>2</sup> sections with an 8 µm thick, dried for 24 hours at +4 °C, a break of the covalent bonds caused by the attachment of formalin is obtained by heating and sonication steps. Then, a digestion of proteins is carried out using S-trap<sup>TM</sup> sample processing technology and the analysis by mass spectrometry is carried out using a DIAPasef acquisition method with 1 hour of chromatography gradient on a timsTOF Pro mass spectrometer.

We applied the selected protocol to the comparative analysis of different areas of biopsies from patients with MPNST preserved in FFPE blocks, cancerous and non-cancerous. The results of the analysis were compared with the analysis of the same samples using macrodissection. Among the 13 samples analyzed, a median of 3.475 proteins was quantified. A principal component analysis shows that the quantifications obtained make it possible to separate the non-cancerous samples from the cancerous samples whereas it was not the case when samples are analyzed by macrodissection. Even if the number of quantified proteins is the half of the number of proteins obtained with macrodissection, the number of statistically differentially expressed proteins is higher. Indeed, a *t-test* carried out on these two types of samples makes it possible to highlight 196 relevant proteins differentially expressed (p-value <0.01) between cancerous and non-cancerous samples although only 19 proteins were statistically differential with macrodissection on the same samples. Functional analyses on differentially expressed proteins performed using Ingenuity Pathways Analysis also show more than 20 proteins already found deregulated in MPNST.

In conclusion, we set up a protocol of proteomic analyses of FFPE samples using tissue selection with microdissection that shows best performance in the highlighting of differentially expressed proteins between cancerous and non-cancerous tissues in the same bloc. Further investigations have to be realized to validate the obtained results. The development of protocols to enhance this resource is therefore an important issue. In particular, we want to go further in the development of the protocol by reducing the sampling area (i.e. < 10 mm<sup>2</sup>). Then, we would then like to use laser microdissection to select not areas of interest but cell populations taken individually from within the tissue. This is particularly relevant in the study of intratumoral heterogeneity, immunology and microbiology.

## Benefits of circulating human metabolites from fish cartilage hydrolysate on primary human dermal fibroblasts, an ex vivo clinical investigation coupled to proteomic study for skin health applications

Dr Julie LE FAUDER<sup>a</sup>, Aurélie GUEHO<sup>b</sup>, Régis LAVIGNE<sup>b</sup>, Fabien WAUQUIER<sup>c</sup>, Line BOUTIN-WITTRANT<sup>c</sup>, Emmanuelle COM<sup>b</sup>, Elodie BOUVRET<sup>a</sup>, Yohann WITTRANT<sup>d</sup>, Charles PINEAU<sup>b</sup>

<sup>a</sup> Abyss Ingredients, <sup>b</sup> UNIVERSITÉ DE RENNES, <sup>c</sup> Clinic'n'Cell SAS, <sup>d</sup> INRAE, Clermont Auvergne University

**CONTEXT:** Due to its significant exposure to stressful environmental factors, the skin undergoes a high remodeling rate over time, which alters not only its appearance but also its functionality. This alteration of the skin, namely photoaging, is characterized by dryness and a loss of elasticity that mainly originates from the dysregulation of dermal fibroblast activities. In order to overcome such tissue outcome, cosmetic products have evolved toward nutricosmetics, thus promoting beauty from within. Among bio-actives of interest, peptides and polysaccharides deriving from plant or animal sources may exert various biological activities beyond their nutritional value. Recently, using an original ex vivo clinical approach, which considers the clinical digestive course of nutrients, we demonstrated that fish cartilage hydrolysate (FCH) supported hydration, elasticity and limiting the expression of catabolic factors involved in photoaging onset [1].

**OBJECTIVE AND METHOD:** In this study, the action of FCH, a natural active food ingredient obtained from enzymatic hydrolysis, which combines collagen peptides and chondroitin sulfate, has been evaluated on human dermal fibroblasts (HDFs). For that purpose, an ex vivo clinical approach coupled to a Data Independent Acquisition (DIA)-MS proteomic analysis (timsTOF Pro, Bruker Daltonics) has been performed in order to better elucidate the underlying mechanisms.

**RESULTS:** The DIA mode allowed the robust quantification of 6,842 proteins from HDFs. Mass spectrometry data show that the circulating metabolites produced in humans following FCH intake stimulate protein expression related to healing in HDFs. To validate our proteomics findings, scratch wound assays have been carried out, demonstrating a significative acceleration of tissue repair by FCH after 6 and 12h.

**CONCLUSION:** Altogether, these data provide additional clues on the mechanisms likely contributing to the beneficial impact of FCH on human skin functionality by supporting healing.

[1] Wauquier F., Boutin-Wittrant L., Bouvret E., Le Faouder J., Roux V., Macian N., Pickering G., Wittrant Y. *Nutrients*. 2022 Nov 25;14(23):5027.

## Diabetes mellitus and aortic stenosis head to head: toward personalized medicine in patients with both pathologies.

Dr Nerea CORBACHO ALONSO<sup>a</sup>, Tamara SASTRE OLIVA<sup>a</sup>, Ines PERALES SANCHEZ<sup>a</sup>, Cristina JUAREZ ALIA<sup>a</sup>, Montserrat CARRASCAL<sup>b</sup>, Luis F LOPEZ ALMODOVAR<sup>c</sup>, Jorge SOLIS<sup>d</sup>, Luis R PADIAL<sup>c</sup>, Teresa TEJERINA<sup>e</sup>, Laura MOURIÑO ALVAREZ<sup>a</sup>, Maria G BARDERAS<sup>a</sup>

<sup>a</sup> Hospital Nacional de Paraplégicos and Instituto de Investigaciones Sanitarias de Castilla-La Mancha (IDISCAM), <sup>b</sup> Institut d'Investigacions Biomèdiques de Barcelona, Consejo Superior de Investigaciones Científicas (IIBB-CSIC), <sup>c</sup> Hospital Universitario de Toledo, <sup>d</sup> Hospital Universitario 12 de Octubre and Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), <sup>e</sup> Universidad Complutense de Madrid

Diabetes mellitus (DM) and calcific aortic stenosis (CAS) are common morbidities in the elderly, which are both chronic, progressive and often concomitant diseases. Several studies revealed that DM increases the risk of developing severe CAS, yet clear information about the relationship between both these diseases and the influence of DM on the progression of CAS is currently lacking. To evaluate the effect of DM on aortic valves and on the process of calcification, and to achieve better patient management in daily clinical practice, we analyzed calcified and non-calcified valve tissue from patients with severe CAS, with or without DM. A proteomic strategy using isobaric tags was adopted and the plasma concentrations of nine proteins were studied using three orthogonal methods (Parallel Reaction Monitoring (PRM), Enzyme-Linked Immunosorbent Assay (ELISA) or Western blotting) and in a separate cell model. The differentially expressed proteins identified are implicated in biological processes like endopeptidase activity, lipid metabolism, coagulation and fibrinolysis. ROC curves were performed for these proteins in order to establish its sensibility and specificity. The diagnostic power of all proteins as a panel was much better than the proteins alone. The results obtained provide evidence that DM provokes changes in the proteome of aortic valves, affecting valve calcification. This finding may help enhance our understanding of the pathogenesis of CAS and how DM affects the evolution of this condition, an important step in identifying targets to personalize the treatment of these patients.

## BeatBox and iST for streamlined FFPE tissue processing: A xylene-free, robust, and high-throughput sample preparation for proteomic analysis.

Dr Quentin ENJALBERT<sup>a</sup>

<sup>a</sup> PreOmics

### Introduction:

Formalin-fixed paraffin-embedded (FFPE) tissues are an invaluable resource for retrospective clinical studies to investigate molecular mechanisms or to discover new biomarkers. The great interest of FFPE tissues lies in the valuable clinical metadata associated with them and the large size of FFPE biobanks worldwide, collected by hospital pathology departments with the primary objective of diagnosis.

Indeed, the FFPE tissue preservation has not been designed for proteomic analysis, which poses many problems for proteomics sample preparation. Harsh conditions have to be applied to reverse the cross-linking and to obtain an adequate extraction of proteins. In addition, paraffin interferes with Liquid Chromatography-Mass Spectrometry (LC-MS) analysis, reason why most protocols include an initial xylene-based deparaffinisation step that is time-consuming, poses a health risk due to high levels of toxicity, and results in a risk of sample loss.

To address these challenges, an optimised workflow combining the BeatBox tissue homogeniser and the iST kit for proteomic sample preparation has been developed. It eliminates the need for xylene-based deparaffinisation and allows efficient and robust processing of 96 samples in parallel from FFPE tissue to clean peptides in one working day.

### Method:

This innovative FFPE sample preparation process begins with the homogenisation of 10µm FFPE curls for 10 minutes, high settings in the BeatBox. Homogenisation is followed by a one hour incubation at 80-95°C, 1000 rpm to de-crosslink, extract, reduce and alkylate the proteins in one step. After spinning and cooling, the samples are transferred to clean containers with the solidified paraffin remaining in the primary tubes. Finally, the samples are subjected to the next steps of the iST sample preparation with tryptic digestion followed by optimised peptide clean-up with an additional wash step designed to effectively remove the last traces of paraffin. This workflow allows the FFPE tissue to be processed into clean peptides ready for LC-MS injection in less than 6 hours.

### Results:

A step-by-step benchmark of this innovative workflow against a traditional xylene-based deparaffinisation and sonication workflow was performed using mouse heart muscle, kidney and liver. The same mouse tissues were preserved in fresh frozen or paraffin-embedded state to assess the suitability of the two tissue preservation techniques for proteomic analysis.

For FFPE tissues, BeatBox outperformed sonication for both xylene-based and xylene-free methods, revealing a >10% increase in protein IDs in all tissue types, with the largest increase of 43% in mouse heart muscle FFPE tissue. Protein IDs and dynamic range of fresh frozen and FFPE tissues, treated with the BeatBox + iST workflow without xylene, showed a high overlap of up to 87% shared proteins and a similar dynamic range of up to 4 orders of magnitude.

To determine technical variability in the use of BeatBox for FFPE processing, coefficients of variation (CVs) were calculated from replicates (n=4) of the xylene-based and xylene-free methods, resulting in median CVs <7%. Bearing in mind the treatment of large cohorts, inter-day reproducibility was assessed over 4 days, resulting in median CVs <8%.

### Conclusion:

To conclude, the PreOmics BeatBox-iST workflow for FFPE tissues allows for thorough proteomic analyses, even when compared to fresh frozen tissues, whose storage is slightly more efficient but more energy consuming. This protocol saves considerable time compared to traditional xylene-based sonication methods, while using an easy-to-use paraffin removal step that is seamlessly integrated into the workflow. This new workflow aims to

provide an answer to large-scale retrospective studies of FFPE tissue enriched with clinical metadata.



## Effect of coronary artery disease on aortic valve calcification: an insight into oxidative stress

Tamara SASTRE OLIVA<sup>a</sup>, Nerea CORBACHO ALONSO<sup>b</sup>, Inés PERALES SÁNCHEZ<sup>b</sup>, Cristina JUAREZ ALIA<sup>b</sup>, Juan Antonio LÓPEZ<sup>c</sup>, Luis Fernando LÓPEZ ALMODÓVAR<sup>d</sup>, Jesús VÁZQUEZ<sup>c</sup>, Luis R PADIAL<sup>d</sup>, Teresa TEJERINA<sup>e</sup>, Laura MOURINO ALVAREZ<sup>b</sup>, M G BARDERAS<sup>b</sup>

<sup>a</sup> Hospital Nacional de Paraplégicos, SESCAM, <sup>b</sup> Hospital Nacional de Paraplégicos, SESCAM, Toledo, SPAIN, <sup>c</sup> Centro Nacional de Investigaciones Cardiovasculares, CNIC, Madrid, SPAIN, <sup>d</sup> Complejo Hospitalario Universitario de Toledo, SESCAM, Toledo, SPAIN, <sup>e</sup> Faculty of Medicine, Complutense University Madrid, SPAIN

### Objective

Calcific aortic valve disease (CAVD) and coronary artery disease (CAD) are related cardiovascular pathologies in which common processes lead to the calcification of the corresponding affected tissue. Among the mechanisms involved in calcification, the oxidative stress that drives the oxidation of sulfur containing amino acids, such as cysteines (Cys), is of particular interest.

### Methods

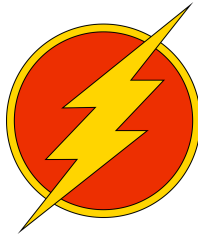
To evaluate the effect that CAD has on aortic valves, we have performed a proteomic analysis using the FASILOX technique on valve tissue from patients with severe CAVD with and without CAD. This allowed us to quantify proteins and reversible Cys modifications simultaneously in the same experiment. After that, we performed a functional analysis using DAVID to select the proteins of interest. These proteins were verified by Parallel Reaction Monitoring (PRM) and Western Blotting in plasma samples from an independent cohort of patients. Besides, we used the PEG-PCMal labelling technique to visualize changes in the redox state.

### Results

We identified 16 proteins with different levels of expression between the two conditions studied. After the verification phase, we defined a panel of 4 proteins related to inflammation, acute phase response and complement system. Besides, we found 7 peptides with differences in the redox state of their Cys, two of them belonging to albumin. Importantly, these sites of Cys oxidation in albumin have not been described previously. The results showed significant differences between the more oxidized states of the albumin, which may be relevant to define new therapeutic options.

### Conclusions

These results provide evidence that CAD affects valve calcification, modifying the molecular profile of aortic valve tissue. In addition, the study of the redox proteome has allowed us to define two new sites of Cys oxidation in albumin, which represents a starting point for future functional and pharmacological studies, with the final purpose of improving patients management.



## Characterization of abundant periprostatic adipose tissue (PPAT) in prostate cancer cells: from inflammation-free expansion to over-secretion of endotrophin.

Dr Manuelle DUCOUX-PETIT<sup>a</sup>, David ESTÈVE<sup>a</sup>, Aurélie TOULET<sup>a</sup>, Mathieu ROUMIGUIÉ<sup>a</sup>, Dawei BU<sup>b</sup>, Sarah PERICART<sup>c</sup>, Chloé BELLES<sup>d</sup>, Cécile MANCEAU<sup>d</sup>, Cynthia HOUËL<sup>d</sup>, Nathalie VAN ACKER<sup>c</sup>, Stéphanie DAUVILLIER<sup>a</sup>, Yiyue JIA<sup>a</sup>, Camille FRANCHET<sup>e</sup>, Nicolas DOUMERC<sup>f</sup>, Mathieu THOULOZAN<sup>f</sup>, Sophie LE GONIDEC<sup>g</sup>, Philippe VALET<sup>g</sup>, Bernard MALAUDAUD<sup>h</sup>, Odile BURLET-SCHILTZ<sup>a</sup>, Anne BOULOUMIÉ<sup>d</sup>, Philipp SCHERER<sup>i</sup>, Delphine MILHAS<sup>a</sup>, Catherine MULLER<sup>a</sup>

<sup>a</sup> IPBS, Institute of Pharmacology and Structural Biology, <sup>b</sup> Touchstone Diabetes Center, Department of Internal Medicine and Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA, <sup>c</sup> Department of Pathology, Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France, <sup>d</sup> Institut des Maladies Métaboliques et Cardiovasculaires, Université de Toulouse, INSERM, UPS, Toulouse, France, <sup>e</sup> Department of Pathology, Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France, <sup>f</sup> Institut des Maladies Métaboliques et Cardiovasculaires, Université de Toulouse, INSERM, UP, <sup>g</sup> Department of Urology and Renal Transplantation, Toulouse University Hospital ; Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France., <sup>h</sup> Institut RESTORE, Université de Toulouse, CNRS U-5070, EFS, ENVT, INSERM U1301, UPS, Toulouse, France., <sup>i</sup> Department of Urology and Renal Transplantation, Toulouse University Hospital ; Institut Universitaire du Cancer de Toulouse, Oncopole, Toulouse, France, <sup>i</sup> Touchstone Diabetes Center, Department of Internal Medicine and Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA.

Periprostatic adipose tissue (PPAT) abundance correlates with prostate cancer progression, but the mechanism remains unexplained. Here, we used a statistical approach to define abundant PPAT by normalizing PPAT volume to prostate volume in a cohort of 351 patients with a linear regression model. Applying this definition, we find that tumors specifically from patients with abundant PPAT exhibit several hallmarks of aggressiveness, suggesting that PPAT abundance might be used to improve risk stratification. We show that abundant PPAT expands by adipocyte hypertrophy but this does not result in inflammation. Extensive extracellular matrix remodeling, notably of the collagen network, and decreased expression of mechano-sensing proteins in adipocytes explains this inflammation-free expansion by decreasing the mechanical constraints on the adipocytes. Moreover, collagen VI degradation in abundant PPAT is associated with production of endotrophin, a matrikine that promotes cancer progression. We find high levels of endotrophin specifically in the urine of patients with abundant PPAT, indicating the clinical relevance of our findings.



## LARGE SCALE CELL LYSATE PROTEOMICS: HIGH-THROUPT SAMPLE PREPARATION AND FAST LC-MSMS

Dr Joanna LIPECKA<sup>a</sup>, Ines METATLA<sup>a</sup>, Sara CECCACCI<sup>a</sup>, Vincent JUNG<sup>a</sup>, Kévin ROGER<sup>a</sup>, Ida Chiara GUERRERA<sup>a</sup>

<sup>a</sup> PROTEOMICS PLATFORM NECKER, UNIVERSITÉ PARIS CITÉ - STRUCTURE FÉDÉRATIVE DE RECHERCHE NECKER, INSERM US24/CNRS UAR3633, PARIS 75015, FRANCE.

### Introduction

Simplifying and automating sample preparation in large scale proteomics analysis is crucial to improve speed and reproducibility. Sample preparation is a critical step that greatly impacts data quality. While detergent-based digestion methods are commonly used for efficient cell lysis and protein solubilization, they may not be compatible with downstream liquid chromatography-coupled mass spectrometry, which can require additional purification and buffer-exchange steps. However, S-Trap sample preparation technology is an effective method that enables the use of detergent, ensures efficient protein digestion, and provides high peptide recovery.

In this study we evaluated the S-Trap digestion of 1- 30 µg on two formats: micro spin columns and 96-well plates. The micro spin columns, suitable for < 100 µg samples are very performant even for low protein samples (<1µg) however they are not easily automatable and thus less suitable for large-scale studies. On the other hand, the 96-well plates are designated for the use with liquid handling systems and allow faster processing but are designed to be used for 100-300 µg samples.

### Methods

The study used HEK-293 cell lysates for proteomics analysis. S-Trap<sup>TM</sup> micro spin columns and 96-well plate mini (Protifi, Huntington, USA) were utilized for sample digestion according to the manufacturer's instructions, with an increased number of washes (6x with 250µL and 3x with 400µL S-Trap buffer). The samples were then processed by centrifugation at 1500g for 2min or using a positive pressure manifold (Resolvex M10, Tecan, USA). 200ng of peptides were injected on a nanoElute system 400nL/min using a packed emitter column (C18, 25cm×75µm 1.6µm) (Ion Optics, Australia) using a 15min gradient coupled to a timsTOF Pro (Bruker Daltonics, Germany) MS. The mass spectrometry data was acquired in diaPASEF mode, with an m/z range from 475 to 1000 Th, and an ion mobility range from 0.85 to 1.27V s/cm<sup>2</sup> (1/k0). The total cycle time was set to 0.95s. Data analysis was performed using DIA-NN software version 1.8.1 against the human UniProtKB/Swiss-Prot Homo sapiens database (v01/2023, 26619 entries) with a false discovery rate (FDR) of 0.01, and match between runs was allowed.

### Results

Firstly, the objective of our study was to investigate the feasibility of using 96-well S-Trap plates for protein samples under 100 µg. We compared the number of identified protein groups using centrifugation-based micro spin columns and 96-well S-Trap plates for 1 µg, 10 µg, and 30 µg samples. Overall, both formats yielded over 6500 proteins, but we observed higher variability in the results obtained with the 96-well S-Trap plates. To improve yield and repeatability, we made minor modifications to the S-Trap plate protocol, including increasing trypsin amounts, washing and digestion volumes, and eliminating AF from the elution.

In our efforts to automate the sample preparation workflow, we compared the performance of the 96-well plate using a manual protocol with centrifugation steps and a semi-automatic positive pressure system. The automated protocol resulted in the identification of 7393 ± 53 protein groups (when starting from 10 µg and 30 µg) and 6693 ± 613 protein groups when starting from 1 µg, indicating a 10% improvement in protein group identification using positive pressure compared to centrifugation.

In summary, our study highlights the advantages of the 96-well plate format for large-scale proteomics studies and demonstrates the potential use of S-Trap plates for samples between 10-100 µg with minor protocol modifications. Furthermore, our findings indicate that using a positive pressure system instead of centrifugation for the 96-well plate format can lead to better repeatability and identification of more protein groups, possibly due to improved peptide recovery.

## Optimized and scalable workflow for high sensibility urine proteomics analysis

Vincent JUNG<sup>a</sup>, Ines METATLA<sup>a</sup>, Kevin ROGER<sup>a</sup>, Ida Chiara GUERRERA<sup>a</sup>

<sup>a</sup> Proteomics platform Necker, Université Paris Cité - Structure Fédérative de Recherche Necker, INSERM US24/CNRS UAR3633, Paris 75015, France

### Introduction

Urine is a readily available and non-invasive sample that has been increasingly used in proteomics research, particularly for the study of kidney diseases. However, the proteomic analysis of urine samples can be challenging due to their high variability, low protein concentration, variability in sample collection and storage, and presence of interfering compounds, like salts. These challenges need to be addressed before running cohort analysis.

In this work, we present our optimized workflow for deep urine proteomics analysis. We compared urine sample collections, preparation protocols, chromatographic separation gradients, DDA and DIA MS analysis.

### Methods

For precipitation with TCA, aliquots of 1 mL urines were precipitated by adding TCA to a final concentration of 10%, centrifuged 40min at 4 °C, and the resulting pellets were washed three times with 1 mL of cold acetone for 20min at 14,000g (n=5). For the precipitation in acetone, aliquots of 1mL urines were precipitated by adding 4 volumes of acetone, centrifuged 1h at 14,000 g. The precipitated proteins were suspended in 5% SDS and digested using the S-Trap (Protifi, Hutington, USA). 200ng of peptides were injected into a nanoElute HPLC system (25cm IonOptiks) coupled to a TimsTOF Pro MS (Bruker Daltonics, Germany). Peptides were separated in 40min and acquired in DDA-PASEF mode setting m/z range from 100 to 1700Th, ion mobility range from 0.75 to 1.25V s/cm<sup>2</sup>(1/k0), total cycle time to 1.1s and number of PASEF MS/MS scans to 10. Data were also acquired in dia-PASEF over a 40min or a 15min elution gradient, (m/z range=475 to 1000 Th, IM=0.85 to 1.27 V s/cm<sup>2</sup>(1/k0), total cycle=0.95s; PASEF MS/MS scans=10). DDA raw files were processed using MaxQuant version 2.1.3.0 and DIA raw files were processed using DIANN v1.8.1 with a Homo sapiens database (release 02/2022, 20398 entries). Match between run was allowed, except when comparing TCA and acetone precipitation protocols.

### Results

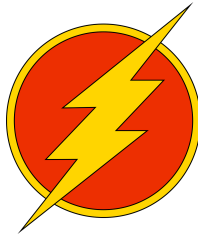
Firstly, we evaluated the stability of proteins in urine samples stored for up to 48h at 4 °C before being frozen at -80 °C. Our results showed that number of proteins groups identified remained stable for 48 hours, with minor increase in protein abundance variability after 48h.

Next, we compared two of the most frequently methods used to concentrate urinary proteins: TCA and acetone precipitation. Acetone precipitation resulted in a higher overall protein yield (20.98+/-3.88µg) vs TCA (8.05+/-0.98µg), although, the number of proteins identified was similar leading to ca 2000 protein groups, with a higher variability in acetone precipitated samples (DDA-mode). Our data showed that TCA precipitation allowed for the enrichment of CD9, CD63 and CD81, markers of exosome, which are informative for the monitoring kidney intracellular state.

Finally, we compared different LC-MSMS conditions and evaluated the number of protein groups identified using DDA and DIA modes on a 40min gradient. We obtained 2081+/-13 proteins using nanoElute in DDA mode, while switching to DIA mode, boosted the identification 4460+/-14 protein groups (n=5). Furthermore, in anticipation of large cohorts' analysis, we reduced the gradient time down to 15 min in order to shorten the acquisition time, allowing the identification of 3479+/-8 urinary proteins from the same samples (n=5).

A large cohort of patients affected by a rare nephropathy is currently being analyzed using this protocol and the results will be shared at the conference.

In conclusion, this study provides insights into the optimization of urine sample preparation and chromatographic separation for large scale proteomics analysis. TCA precipitation can enrich for exosome and extracellular vesicles and implementation of DIA increases 2.5x the depth and reproducibility of urinary proteome analysis.



## Identification of potential biomarkers related to A $\beta$ deposition through proteomic profiling in patients with Mild Cognitive Impairment

Vera M. MENDES<sup>a</sup>, Daniela MOUTINHO<sup>b</sup>, Sara C. MADEIRA<sup>c</sup>, Inês BALDEIRAS<sup>d</sup>, Manuela GUERREIRO<sup>b</sup>, Sandra CARDOSO<sup>b</sup>, Helena AIDOS<sup>c</sup>, Isabel SANTANA<sup>e</sup>, Alexandre MENDONÇA<sup>b</sup>, Bruno MANADAS<sup>a</sup>

<sup>a</sup> CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal ; Centre for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Coimbra, Portugal, <sup>b</sup> Faculty of Medicine, University of Lisbon, Lisboa, Portugal, <sup>c</sup> LASIGE, Faculty of Sciences, University of Lisbon, Lisboa, Portugal, <sup>d</sup> CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal ; Centre for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Coimbra, Portugal ; Faculty of Medicine, University of Coimbra, Coimbra, Portugal, <sup>e</sup> Faculty of Medicine, University of Coimbra; Department of Neurology, Hospital and University Centre of Coimbra; Center for Neuroscience and Cell Biology ; Centre for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Coimbra, Portugal

Alzheimer's disease is the leading type of dementia, characterized by a gradual decline in cognitive and behavioral function. It usually starts with mild memory loss, progressing over time to difficulties with communication and response to the environment. Gaining insight into the underlying causes of this disorder is critical for enabling early and precise diagnosis, as well as developing effective treatments. Mild cognitive impairment (MCI) is a condition that often precedes AD, and studying MCI can help identify early signs and risk factors for AD, allowing for earlier intervention and potential prevention of the disease. A comparative proteomic study was carried out between an MCI group of patients without any biomarkers of A $\beta$  deposition or neuronal injury (MCI<sub>A $\beta$ -</sub>) and an MCI group of patients fulfilling criteria of high likelihood for AD (MCI<sub>A $\beta$ +</sub>). In order to enhance the reliability of the results and avoid limitations to a specific patient group, this study employed analysis and comparison of samples obtained from two distinct cohorts. The protein content from 60 CSF samples from MCI<sub>A $\beta$ -</sub> patients and 68 from MCI<sub>A $\beta$ +</sub> patients was precipitated with methanol and analyzed using the Short-GeLC-SWATH approach. Pooled samples were prepared to generate the peptide ion libraries by Data Dependent acquisition (DDA) and each sample was analyzed by SWATH/DIA for relative protein quantification.

Results show that, out of the 617 proteins in the library, 517 were quantified and subjected to statistical analysis to identify the most significant proteins that distinguished between the two MCI groups. This analysis was conducted separately for each cohort and later for all samples. Differences between the two cohorts were identified through PCA analysis, and to mitigate any potential technical variations, batch effect correction was conducted. This facilitated a reassessment of the observed disparities between the two cohorts.

The most interesting proteins related to A $\beta$  deposition (MCI<sub>A $\beta$ +</sub>) were selected based on a VIP score above 1, extracted from the PLS-DA analysis. Although the number of selected proteins was similar between both cohorts (around 160 proteins), the proteins themselves differed significantly. Gene Ontology enrichment analysis was conducted on the identified sets of interesting proteins. While some biological processes, such as cell adhesion and post-translational protein modification, were common to both cohorts, a large number of biological processes were found to be enriched in only one of the cohorts.

Despite the differences among the selected proteins from each cohort, the combination of all the analyses resulted in a group of proteins that distinguishes MCI<sub>A $\beta$ -</sub> from MCI<sub>A $\beta$ +</sub> and that can be considered potential biomarkers for Alzheimer's disease.

## Proteomics-based stratification of bariatric patients through non-invasive samples allows a good matching with biochemically assayed parameters

Raffaello VIGANÒ<sup>a</sup>, Jonica CAMPOLO<sup>b</sup>, Francesca BRAMBILLA<sup>c</sup>, Dario DI SILVESTRE<sup>c</sup>, Ettore CORRADI<sup>d</sup>, Marina PAROLINI<sup>b</sup>, Patrizia TARLARINI<sup>d</sup>, Francesco SCAGLIONE<sup>e</sup>, Pierluigi MAURI<sup>c</sup>

<sup>a</sup> Institute for Biomedical Technologies, National Research Council (ITB-CNR), <sup>b</sup> CNR Institute of Clinical Physiology, ASST GOM Niguarda, Milan, ITALY, <sup>c</sup> Institute for Biomedical Technologies, National Research Council (ITB-CNR), Segrate (Milan), ITALY, <sup>d</sup> Clinical Nutritional Unit, ASST GOM Niguarda, Milan, ITALY, <sup>e</sup> Clinical Nutritional Unit, ASST GOM Niguarda, Milan, ITALY; Department of Chemical-Clinical and Microbiological Analyses, GOM Niguarda, Milan, ITALY.

Aging is characterized by a progressive functional impairment of tissues and organs with different pathological consequences. In particular, inflammatory and metabolic abnormalities are at the basis of several typical diseases of the elderly, such as neurodegenerative diseases and type 2 diabetes, with high co-morbidity. Among risk factors for neuronal and cognitive impairment, various types of stress inducing cellular senescence and metabolic disorders such as obesity and diabetes must be considered. These conditions are of particular public health relevance given the Italian context of overweight (39.4%), obesity (14.1%) and diabetes (8.6%) for people aged 50-69 [1]. They share insulin resistance (IR) as common mechanism and are potentially modifiable by lifestyle interventions. Mitochondrial dysfunction, oxidative and endoplasmic reticulum stress have been observed at the cellular level following IR [2,3]. These events induce a pro-inflammatory condition and cellular senescence [4,5]. In this context, it is crucial to identify molecular mechanisms associated with metabolic disorders before the onset of clinically detectable pathologies such as diabetes or organ dysfunctions, to rapidly apply preventive interventions focused on lifestyle (nutrition, physical activity, sleep) which may avoid most severe symptoms derived from metabolic disorders.

To identify and characterize molecular markers and pathways related to metabolic dysfunction, we applied a high-resolution nano-liquid chromatography coupled to tandem mass spectrometry (hr nLC-MS/MS) workflow to a cohort of 45 subjects recruited at the CNR Institute for Clinical Physiology (CNR-IFC) of Milan, in collaboration with the Clinical Nutritional Unit of Niguarda Hospital (Milan). The inclusion criteria were: age  $\geq 50$  and  $\leq 70$  years and body mass index (BMI)  $\geq 30$  and  $< 40$  kg/m<sup>2</sup>. We excluded patients with organ dysfunction, active cancer, or overt diabetes. Extracellular Vesicles (EVs) from their urine were analysed by LC-MS/MS and the 3914 identified proteins were used to perform a linear discriminant analysis (LDA) to cluster the subjects and identify descriptors ( $P < 0.05$ ) able to discriminate the different profiles [6]. A label-free quantification and differential analysis were applied to selected proteins, through MProMa platform [7]. A functional evaluation on enriched pathways was also performed to identify the molecular processes that may be involved in metabolic disorders, causing alterations that may lead to severe pathologies.

This workflow allowed to classify patients into two main molecular clusters; the two groups obtained from proteomics stratification showed remarkable differences in the non-enzymatic glycation grade of the proteins identified. Moreover, we verified the correlation between plasma redox parameters associated to metabolic disorders measured previously by CNR-IFC and the two main molecular clusters. In particular, the plasma levels of total and oxidized glutathione and cysteine, expression of an oxidant balance, were significantly higher in one group related to the other, in agreement with the non-enzymatic glycation level. Functional evaluation of proteins identified overlapped with proteomic stratification, showing a differential representation of pathways related to oxidative stress, such as and "glutathione, cysteine and methionine metabolism" and "ferroptosis" (KEGG). The results obtained suggest that cellular pathways related to oxidative stress condition may be affected since the very first stages of metabolic disorder itself. In this context, the proteomic profiling may be helpful to target this phenomenon before the clinical onset of metabolism-related pathologies and can be a useful tool to early identify subjects exposed to a higher risk of development of most severe diseases related to this condition.

### Bibliography

1. <http://www.epicentro.iss.it/passi/infoPassi/infoGen.asp>, 2021.
2. De la Monte SM., 2009. doi: 10.5483/bmbrep.2009.42.8.475.
3. Ma L. et al., 2015. doi: 10.1016/j.cca.2015.01.027.
4. Brozzi F. et al., 2015.. doi: 10.1007/s00125-015-3669-6.
5. Zhu YF. et al., 2015. doi: 10.1016/j.ejphar.2015.06.002.
6. Doncheva NT. et al., 2019. doi: 10.1021/acs.jproteome.8b00702.
7. Di Silvestre D. et al., 2013; doi: 10.1007/978-1-62703-386-2\_3.



## Exploring the potential of urinary proteomics for non-invasive diagnosis of head and neck cancer

Patrícia TAVARES<sup>a</sup>, Joaquim CASTRO SILVA<sup>b</sup>, Rita FERREIRA<sup>a</sup>, Ana Carolina CASTRO<sup>b</sup>, Daniel MOREIRA-GONÇALVES<sup>c</sup>, Susana AVEIRO<sup>d</sup>, Pedro DOMINGUES<sup>a</sup>, António BARROS<sup>e</sup>, Francisco AMADO<sup>a</sup>, Rui VITORINO<sup>f</sup>, Lúcio LARA SANTOS<sup>g</sup>

<sup>a</sup> LAQV-REQUIMTE, Associated Laboratory for Green Chemistry of the Network of Chemistry and Technology, Department of Chemistry, University of Aveiro, Aveiro, Portugal, <sup>b</sup> Instituto Português de Oncologia do Porto Francisco Gentil, Porto, Portugal, <sup>c</sup> CIAFEL, Center for Research in Physical Activity, Health and Leisure, University of Porto, Faculty of Sports, University of Porto, Porto, Portugal, <sup>d</sup> GreenCoLab-Green Ocean Association, University of Algarve, Faro, Portugal, <sup>e</sup> UCIBIO@REQUIMTE/Laboratório de Toxicologia, Departamento de Ciências Biológicas, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal; Department of Cardiothoracic Surgery and Physiology, Faculty of Medicine, Porto, Portugal, <sup>f</sup> iBiMED, Department of Medical Sciences, Institute of Biomedicine, University of Aveiro, <sup>g</sup> Experimental Pathology and Therapeutics Group Research Centre, Portuguese Oncology Institute of Porto (IPO-Porto), Porto, Portugal

Head and neck squamous cell carcinoma (HNSCC) is a challenging disease to diagnose and treat due to its location and complexity. Urinary proteomics has emerged as a promising non-invasive diagnostic tool for the early detection and monitoring of various diseases. In this study, we investigated the potential of urinary proteomics for the diagnosis of HNSCC. Urine samples were collected from patients with HNSCC and healthy controls. Using mass spectrometry-based proteomic analysis, we identified several inflammation- and cancer-related proteins that could serve as tumor biomarkers. Additionally, we observed changes in the urinary proteome of HNSCC patients, including a reduction in Galectin-3 Binding Protein, CD44, and GM2A. Our results suggest that urinary proteomics could be a valuable tool for the early detection and monitoring of HNSCC. Furthermore, our study provides proof of principle that urinary proteomics could be used to track proteasome changes during complex treatments based on urine samples.

## Exploring the X Factor using Proteomics: Enhancing Health Outcomes in Pulmonary Arterial Hypertension Rats with Plasma from Trained Rats

Prof Rita FERREIRA<sup>a</sup>

<sup>a</sup> UNIVERSITY OF AVEIRO

Rita Ferreira<sup>1</sup>, Rui Vitorino<sup>2</sup>, Bruno Manadas<sup>3</sup>, Francisco Amado<sup>1</sup>, Daniel Moreira-Gonçalves<sup>4</sup>

<sup>1</sup> LAQV-REQUIMTE, Department of Chemistry, University of Aveiro

<sup>2</sup> iBiMED, Department of Medical Sciences, University of Aveiro

<sup>3</sup> CNC-Center for Neuroscience and Cell Biology, University of Coimbra

<sup>4</sup> CIAFEL, Faculty of Sports, University of Porto

Pulmonary arterial hypertension (PAH) remains a major therapeutic challenge that requires continued research into more effective therapeutic options. The preventive and therapeutic effects of exercise training in various chronic diseases, including PAH, are well recognized. However, certain patients have health limitations that prevent them from exercising regularly. For such patients, the development of an exercise mimetic is a promising therapeutic option.

To gain insights into the molecular mechanisms underlying the beneficial effects of exercise training at PAH, a study was conducted using male Wistar rats injected with a single dose of monocrotaline (MCT; 60 mg/Kg) to induce PAH. Subsequently, the PAH rats were treated with plasma from healthy exercised or sedentary rats for a period of 25 days (6 injections every four days). The exercised rats trained on a treadmill exercise for 5 days/week, 60 minutes/day, at a speed of 25 m/min, for a total of 10 weeks. To delve into the molecular composition of injected plasma, GeLC- MS/MS was performed for protein separation, identification, and quantification. Bioinformatics was employed to annotate identified proteins, functional enrichment analysis and explore protein-protein interactions.

Data showed that PAH rats treated with plasma from trained rats exhibited significant improvement in maximal oxygen consumption, a measure of cardiac fitness. In addition, these rats exhibited higher survival and a reduction in MCT-induced hypertrophy and fibrosis of the right ventricle, which are characteristic features of PAH. GeLC-MS/MS data of the plasma injected into the MCT-induced rats revealed that exercise-induced benefits were related with the upregulation of proteins involved in the biological processes of "vasodilation," "response to selenium," and "response to lipids," and downregulation of "wound healing," "VLDL particle remodeling," and "sterol metabolic processes". Three of the proteins modulated by exercise have previously been identified as involved in remodeling of the lung tissue proteome in PAH patients. One of the proteins upregulated in plasma from exercised animals and downregulated in lung tissue from PAH patients was zinc- $\alpha$ -2-glycoprotein, which is known to be associated with lipid metabolism and vascular fibrosis. In other organs such as the heart, zinc- $\alpha$ -2-glycoprotein is thought to have an antifibrotic effect.

Overall, this study not only supports the benefits of exercise training in the context of PAH, but also highlights potential targets for further investigation and therapeutic development.



## Photoactivation of (pro-)activity-based proteome profiling probes to investigate the interactome of the antimalarial drug plasmodione

Dr Agnès HOVASSE<sup>a</sup>, Ilaria IACOBUCCI<sup>b</sup>, Christine SCHAEFFER<sup>a</sup>, Stéphanie BLANDIN<sup>c</sup>, Sarah CIANFERANI<sup>a</sup>, Elisabeth DAVIOUD-CHARVET<sup>b</sup>

<sup>a</sup> BioOrganic Mass Spectrometry laboratory (LSMBO), Institut Pluridisciplinaire Hubert Curien (IPHC), UMR 7178 CNRS, ECPM, F-67087 Strasbourg, Proteomics French Infrastructure (ProFI), <sup>b</sup> UMR7042 CNRS-Unistra-UHA, Laboratoire d'Innovation Moléculaire et Applications (LIMA), ECPM, F-67087 Strasbourg, <sup>c</sup> UPR9022 CNRS-Unistra / U1257 INSERM, Modèles Insectes de l'Immunité Innée (MIR/M3I), IBMC, F-67000 Strasbourg

In 2021, malaria caused an estimated 247 million clinical episodes and 619,000 deaths (80% of them children under 5 years). The causative agents of malaria are parasites belonging to the *Plasmodium* genus. During years, several antimalarial drugs have been produced, but the parasite quickly develops drug-resistances to all of them.

We identified the lead antimalarial plasmodione (PD), which has been demonstrated to kill the malaria parasite in a mechanism of action based on redox cycling. Experimental data indicate that this early lead drug, could exhibit several modes of action (MoA) via its metabolite (PDO), depending on parasitic stages, and protein targets [1-3]. Accordingly, recent studies in yeast identified Nde1 of the mitochondrial respiratory chain, as the main target of PD, and other flavoenzymes as minor targets. We expect that other flavoenzymes are likely to contribute to PD activity. To identify them and fully characterize PD MoA we propose to use an activity-based protein profiling (ABPP) strategy in *Plasmodium* parasites. For this purpose, we synthesized PD- and PDO-derived ABPP probes and used the PDO-based probes, which possess intrinsic benzophenone-like moiety, to crosslink to proteins upon photoirradiation [4].

Activity-based probes were first tested on a model protein, the recombinant model glutathione reductase. Each step was optimized: photoaffinity labeling, tagging photolabeled proteins by click chemistry with a biotin probe and enrichment of biotin tagged proteins using streptavidin beads. Reactions were carried out in physiological media paying attention to *in vivo* issues and compatibility.

The goal is now to transfer this strategy to the whole proteome in order to identify the relevant targets of plasmodione in living *P. falciparum*. Because of the low yield of the whole protocol and the low quantity of malaria parasite in red blood cells, we investigated model proteomes from *S. cerevisiae* and *E. coli*. Currently, work is underway to optimize the protocol using a proteolytic digestion and identification of the targeted enzymes by label-free quantitative MS analysis. In a first approach, unlabeled lysates were used as control. The workflow is performed on knock out (yeast) and overexpression (*E. coli*) mutant models for proteomics differential studies.

Thus, the strategy is going to be applied in a first attempt to model systems to both standardize it before moving on the parasite and to find potential targets that need to be validated in transgenic *P. falciparum* parasites.

[1] Müller T. et al. **J. Am. Chem. Soc.** 2011, 133, 11557-71.

[2] Ehrhardt K. et al. **Antimicrob. Agents Chemother.** 2016, 60, 5146-5158.

[3] Feng L. et al. **Org. Biomol. Chem.** 2018, 16, 2647-2665.

[4] Cichocki B.A. et al. **JACS**<sup>au</sup> 2021, 1, 669-689.



## Plasma ALS and LG3BP predict liver fibrosis in NAFLD patients

David PÉREZ COMPTE<sup>a</sup>, Lucas ETOURNEAU<sup>a</sup>, Anne-Marie HESSE<sup>a</sup>, Alexandra KRAUT<sup>a</sup>, Justine BARTHELON<sup>b</sup>, Nathalie STURM<sup>b</sup>, Hélène BORGES<sup>a</sup>, Salomé BIENNIER<sup>a</sup>, Marie COURÇON<sup>a</sup>, Marc DE SAINT LOUP<sup>c</sup>, Charlotte COSTENTIN<sup>b</sup>, Thomas BÜRGER<sup>a</sup>, Yohann COUTÉ<sup>a</sup>, Christophe BRULEY<sup>a</sup>, Thomas DECAENS<sup>b</sup>, Michel JAQUINOD<sup>a</sup>, Jérôme BOURSIER<sup>d</sup>, Virginie BRUN<sup>a</sup>

<sup>a</sup> CEA GRENOBLE, <sup>b</sup> Service d'hépatogastroentérologie, Pôle Digidune, F-38700 La Tronche, France, <sup>c</sup> Hepato-Gastroenterology Department, University Hospital, Angers, <sup>d</sup> Hepato-Gastroenterology Department, University Hospital, Angers, France

Liver diseases are a major global health problem, causing more than 2 million deaths per year worldwide<sup>1</sup>. Most liver diseases result from viral infections, alcoholism and NAFLD (Non-Alcoholic Fatty Liver Disease). NAFLD is highly prevalent (25% of the global population worldwide<sup>1,2</sup>) and closely associated to the rise of obesity in the world<sup>1</sup>. Patients with NAFLD may progressively develop NASH (Non Alcoholic Steato-Hepatitis) requiring liver transplantation in the most advanced stages<sup>3,4</sup>. Liver biopsy remains the reference examination to detect NAFLD, but it is an expensive, invasive and inappropriate tool for massive screening<sup>3</sup>. Although a large number of potential biomarkers has been discovered for NAFLD patient stratification, none of them is presently able to reach the performance of liver biopsy and histopathological examination. The development of non-invasive tests to diagnose NAFLD patients from the earliest stage is still a crucial need.

As a contribution to the field, we carried out a LC-MS/MS proteomic discovery study using plasma samples (n=160) from NAFLD patients diagnosed at Grenoble hospital hepatology department and classified into five groups of disease severity (fibrosis progression), according to liver biopsy and histopathological examination. This discovery study led to the identification of 114 plasma proteins with differential abundances between two fibrosis groups (early versus advanced fibrosis) in NAFLD patients. In order to verify the relevance of our results, a validation study targeting two plasma biomarkers by ELISA has been launched in an independent cohort of NAFLD patients (n=200).

Finally, we created a Fibrotest-inspired panel including the two validated proteins. The obtained results improve fibrosis detection in NAFLD patients and could be rapidly implemented in hospitals.

References:

- [1] Asrani, S. K., Devarbhavi, H., Eaton, J. & Kamath, P. S. Burden of liver diseases in the world. *J. Hepatol.* **70**, 151-171 (2019).
- [2] Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* **64**, 73-84 (2016).
- [3] Hashimoto, E., Taniai, M. & Tokushige, K. Characteristics and diagnosis of NAFLD/NASH. *J. Gastroenterol. Hepatol.* **28 Suppl 4**, 64-
- [4] Malaguarnera, M., Di Rosa, M., Nicoletti, F. & Malaguarnera, L. Molecular mechanisms involved in NAFLD progression. *J. Mol. Med.* **87**, 679-695 (2009).

## 10 years of troubleshooting on a nanoLC-MS/MS coupling in proteomics

Magali ROMPAIS<sup>a</sup>, Jean-Marc STRUB<sup>a</sup>, Agnès HOVASSE<sup>a</sup>, Hélène DIEMER<sup>a</sup>, François DELALANDE<sup>a</sup>, Alexandre BUREL<sup>a</sup>, Fabrice VARRIER<sup>a</sup>, Christine SCHAEFFER<sup>a</sup>, Sarah CIANFERANI<sup>a</sup>, Christine CARAPITO<sup>a</sup>

<sup>a</sup> INSTITUT PLURIDISCIPLINAIRE HUBERT CURIEN - IPHC\_UMR7178

### Introduction

High-throughput bottom-up proteomics analysis requires cutting-edge technologies and new instruments are launched annually on the market. Troubleshooting of new couplings is not trivial, nevertheless one can learn from previous generation instruments. In 2013, a nano-UPLC system (nanoAcquityUPLC, Waters, USA) coupled to a quadrupole-Orbitrap hybrid mass spectrometer (Q-Exactive Plus, Thermo Scientific, San Jose, CA) was installed in our lab and was at that time a cutting-edge instrument. Ten years later this robust coupling is not the most performant anymore but is still used for day-to-day proteomics as its robustness is quite remarkable. Within a decade of use, various issues occurred either on the chromatographic side, on the spectrometer side or at the spray interface. This poster gathers 10 years of troubleshooting feedback.

### Methods

Every event that required to stop the sample acquisition except for weekly maintenance was considered as an issue. These issues were recorded in our internal quality system for a decade including issue symptoms, solving process and associated cost (loss of days of use, maintenance, spare part replacement).

### Results

Hundreds of issues occurred (~2 per month). Most of them were minor issues and were solved locally, i.e. loss of performances due to chromatographic aging, spray instabilities, instrument source contaminations. Around 40 issues (~4 per year) however implied external help from either the nanoUPLC or mass spectrometer manufacturers.

### Conclusions

This poster is showing a didactic overview of 10 years of troubleshooting with the aim of sharing feedback with the proteomics community and helping new proteomists in their troubleshooting.

## SpecGlobX: a fast tool for aligning mass spectra in large proteomics datasets, capturing dissimilarities arising from multiple complex peptides modifications

Dr Mehdi CHERKAOUJ<sup>a</sup>, Grégoire PRUNIER<sup>a</sup>, Albane LYSIAK<sup>b</sup>, Olivier LANGELLA<sup>c</sup>, Mélisande BLEIN-NICOLAS<sup>c</sup>, Virginie LOLLIER<sup>a</sup>, Emile BENOIST<sup>b</sup>, Géraldine JEAN<sup>b</sup>, Guillaume FERTIN<sup>b</sup>, Hélène ROGNIAUX<sup>a</sup>, Dominique TESSIER<sup>a</sup>

<sup>a</sup> INRAE, PROBE research infrastructure, BIBS facility F-44300, Nantes, France, <sup>b</sup> LS2N, Nantes Université, Nantes Cedex 3, France, <sup>c</sup> Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, PAPPSO, F-91190, Gif-Sur-Yvette, France

### Introduction :

Mass spectra representing peptides carrying multiple unknown modifications are particularly difficult to interpret. Yet, identifying post-translational and chemical modifications remains challenging and is currently limited both by the number of potential modifications that can be considered during a single analysis and by the need to know them in advance. To address this issue, we implemented SpecGlobX, a stand-alone Java application that quickly aligns a (possibly large) list of peptide-to-spectrum matches (PSMs) provided by any open modification search (OMS) method. SpecGlobX takes as input spectra in mgf or MzML format and a comma-delimited spreadsheet describing the list of PSMs, and outputs its best alignment to the candidate, splitting the mass difference between spectra and peptides into one or several shifts.

### Methods :

SpecGlobX relies on a dynamic programming algorithm and finds the best alignment of each input peptide/spectrum pair. SpecGlobX looks for an alignment that maximizes a user-defined score, and in the process possibly splits the mass difference between peptide and spectrum into several mass shifts.

Once the best score for a PSM is obtained, a first alignment is suggested, with zero, one, or more mass shifts. Next, a post-processing step is applied, aiming at optimizing the locations of the mass shifts (if any) in order to increase the number of shared peaks, notably in order to highlight the presence of a 'not-aligned mass' if any (neutral loss, charge error, dimers, etc.).

### Preliminary Data :

SpecGlobX was tested on a simulated experimental dataset Dsim generated from 50,000 tryptic unique peptides randomly selected from the human proteome (Ensembl99, release GrCh38), containing 12 to 25 amino acids. Each peptide was transformed into a doubly charged simulated spectrum with the following modifications: a deamidation on each asparagine (N+0.984016) and a sodium adduct on each aspartic acid (D+21.981943). We randomly removed 20% of the peaks to simulate missing peaks, added noise by creating random peaks (up to 60), and modified the mass so that a neutral loss (17.03 Da) is simulated on each spectrum. Consequently, each spectrum contains one to five modifications. Our tests show that SpecGlobX performs well and detects the neutral loss on a large proportion of spectra (nearly 70%), while 53% of all the expected modifications are correctly identified.

We also evaluated SpecGlobX downstream to three OMS software (MODPlus, MSFragger, SpecOMS), and showed that SpecGlobX significantly improves the percentage of correct modifications when correct PSMs are provided as input by these software: notably, a large part of the misinterpreted neutral losses can be recovered by SpecGlobX. Similarly, the percentage of correct modifications relative to the number of expected modifications (based on the number of N and D in Dsim) is multiplied by two to three depending on the software. SpecGlobX can thus be viewed as a decision-support tool highlighting unexpected and complex peptide modifications.

We finally show, on a smaller dataset, that experimental spectra behave similarly as our simulated spectra. The spectra dataset used was generated from HEK293 cells. We focused on the interpretation of a subset of 77 spectra identifying peptide DATNVGDEGGFAPNIENK according to SpecOMS.

Importantly, our experiments show that SpecGlobX is very fast: one million PSMs is processed every 1.5 minutes on an off the shelf desktop.



## Integrating sample related metadata to automated statistical analysis of quantitative proteomics analysis using wrProteo

Wolfgang RAFFELBERGER<sup>a</sup>, Bastien MORLET<sup>a</sup>, Frank RUFFENACH<sup>a</sup>, Luc NEGRONI<sup>a</sup>

<sup>a</sup> IGBMC

Numerous algorithms and software options for data analysis of quantitative proteomics experiments have emerged over the past years. Human operators still need to interact with computing infrastructure, this remains a frequent bottleneck in efficient data-analysis. With recent efforts on metadata standards the 'sample to data file format' (sdrf) has been proposed and was integrated to the major data-bases like Pride. To exploit the novel possibilities of automating data-analysis pipe-lines, the R package wrProteo now offers integrating such meta-data at the import of quantitative data. The sdrf meta-data describing the sample setup are automatically extracted from a given Pride accession number or an equivalent user-provided table and integrated to all subsequent steps like QC, normalization, NA-imputation and automated statistical data-analyses. In absence of sdrf-data, meta-data exported from MaxQuant or ProteomeDiscoverer can be mined for understanding the experimental setup and integrated in the same way to following analysis. They may also be transformed to a sdrf-like format for integration to Pride.

## Genome assembly and annotation using Brownotate, a newly developed automated tool

Adrien BROWN<sup>a</sup>, Alexandre BUREL<sup>a</sup>, Sarah CIANFERANI<sup>a</sup>, Christine CARAPITO<sup>a</sup>, Fabrice BERTILE<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO), IPHC UMR 7178, Université de Strasbourg, CNRS, 25 rue Becquerel, 67087 Strasbourg, France. Infrastructure Nationale de Protéomique ; ProFI - FR2048, 67087 Strasbourg, France

### Introduction

Mass spectrometry (MS)-based proteomics most often relies on database searching to identify peptides or proteins. Ideally, protein sequences come from publicly available and reviewed reference databases, such as NCBI RefSeq or UniProtKB-Swissprot. However, 98.1% of the 1,758,200 sequencing datasets (with full genome representation; April 2023) for eukaryotes in NCBI are not assembled, and only 0.4% are fully assembled and annotated (0.1% only in Refseq). Even for prokaryotes, 20-30% of the available sequencing datasets are not assembled or annotated. To overcome this bottleneck and enable the proteomist to assemble and annotate any genome on their own, we are developing a bioinformatics pipeline named Brownotate (BR).

### Material and methods

BR gathers existing tools to download DNA sequencing data, filter out unreliable reads, assemble the retained reads, predict protein-coding genes, perform translation into protein sequences, and name them. An evaluation of the completeness of BR-derived assemblies or annotated genes is performed at runtime using the single-copy orthologs expected for given species (OrthoDB). To assess BR performance, BR-derived assemblies and annotations for 24 species belonging to 8 different taxons (bacteria, fungi, arthropods, birds, fish, reptiles, mammals and plants) were compared to reference (REF) data extracted from NCBI SRA and RefSeq. MS raw data downloaded from PRIDE for these same species were also processed using Maxquant (v2.0.3.1) against either the BR-derived or the REF protein database, and datasets of identified proteins (FDR 1%) were compared.

### Results and discussion

Except for four species, BR assemblies appeared 1-26% longer than REF assemblies, even more for the fungi *F. oxysporum* and for plant species (+58% to +270%), but with on average 90% of the REF assemblies being identical to BR assemblies. The completeness (ortholog groups) of REF and BR assemblies and annotations was comparable, with on average an 86% and 81% overlap, respectively, but for reptiles where the values reached only 36-51% and 24-42%, respectively. This suggests equivalent performances for Brownotate and the NCBI pipeline. Except for *S. aureus*, *E. nidulans*, *F. oxysporum* and *N. naja*, more than half of the proteins in REF databases were present in BR-derived databases. We also observed that except for *L. brevis*, *M. xanthus* and *A. gambiae*, more than 40% of BR-annotated proteins were absent from REF protein databases (until 97% for *N. naja* and *C. sativa*). Longer assemblies using BR may explain part of these differences.

The analysis of PRIDE-extracted raw MS/MS data using BR and REF protein databases reveals that, except for reptiles, the use of the REF protein database resulted in an average of 14% more identified peptides compared to the BR-derived database. Regarding the number of protein groups identified, the use of the BR database resulted in an average of 20% more identifications compared to the use of REF, except for two plant species. These results suggest a higher protein sequence coverage using the REF database. However, we also generally observed a lower number of unique proteins when using the BR-derived database, suggesting that BR and REF annotations may differ at the level of the peptide sequences that allow to specifically discriminate isoforms/variants.

### Conclusion

BR proved to generate high-quality protein sequence databases, well-adapted for MS/MS data interpretation. The development of BR is still ongoing, e.g. to adapt it for RNA sequence data and use it in a friendly graphical user interface. Ultimately, BR should help for studies on species for which only DNA or sequence data are available or in the context of personalized medicine.





## Pathway Grabber: Exploring proteomics data with KEGG pathways

Alexandre BUREL<sup>a</sup>, Patrick GUTERL<sup>a</sup>, Sarah CIANFERANI<sup>a</sup>, Christine CARAPITO<sup>a</sup>, Fabrice BERTILE<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO), Institut Pluridisciplinaire Hubert Curien (IPHC), UMR7178, Strasbourg, France ; Infrastructure Nationale de Protéomique ProFI - FR2048, 67087 Strasbourg, France

### Introduction

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource for understanding high-level functions of biological systems from large-scale molecular datasets. It has been developed together with a collection of tools for mapping molecular objects to KEGG pathway maps by the Kanehisa Laboratories ([www.kegg.jp](http://www.kegg.jp))

[www.kegg.jp](http://www.kegg.jp)

### Methods

Pathway Grabber was developed with Julia 1.8.5 and some Javascript. From an Excel file containing UniProt or KEGG identifiers and associated statistical scoring values, including as well p-values from single (e.g. t-tests) or multiple comparisons (e.g. Anova and post-hoc tests) and fold changes, and the definition by the user of thresholds for each parameter, proteins are categorized as “non-significant” or “significant” and the information “upregulated” or “downregulated” is kept. In parallel, KEGG annotations are downloaded for each protein, which allows proteins from the dataset to be distributed among Pathway maps. A cache mechanism has been added to reduce the amount of data transfer, hence making the tool faster and decreasing the impact on the bandwidth. Extracted data is provided in the form of an Excel file and a list of HTML files.

### Results

Output HTML files correspond to all Kegg Pathway maps that contain at least one protein from the dataset. On these maps, the information that has been made available is highlighted using a colour code relative to the ‘statistical category’ (for any item of interest, whether it is, e.g., a module, a protein/gene, a compound, a relation). Additional details about statistical scoring is also given as tooltip text items. All the items in the HTML files are clickable and reroute the user to the corresponding KEGG entries, but these files can also be used completely offline, once they are generated they do not require an Internet connection anymore.

The output Excel file summarizes the whole information that is highlighted on maps, one sheet listing all the pathways associated with each with protein, another sheet listing all proteins associated to each revealed pathway.

### Conclusions

Pathway Grabber is a user-friendly tool for biologists and proteomists, helping them to get a comprehensive view of the molecular regulations from large lists of proteins, including the statistical metrics of differential analysis. Therefore, it allows to greatly accelerate omics data mining and it helps to nicely draw graphical representations of omics results (1, 2). Today, Pathway Grabber constitutes a solid basis for many future improvements (e.g. network analysis, extension to other resources like Reactome).

### References

1. Chazarin, B.; Benhaim-Delarbre, M.; Brun, C.; Anzeraey, A.; Bertile, F.; Terrien, J. Molecular Liver Fingerprint Reflects the Seasonal Physiology of the Grey Mouse Lemur (*Microcebus murinus*) during Winter. *Int. J. Mol. Sci.* **2022**, *23*, 4254. <https://doi.org/10.3390/ijms23084254>
2. Tascher, G.; Burban, A.; Camus, S.; Plumel, M.; Chanon, S.; Le Guevel, R.; Shevchenko, V.; Van Dorsselaer, A.; Lefai, E.; Guguen-Guillouzo, C.; Bertile, F. In-Depth Proteome Analysis Highlights HepaRG Cells as a Versatile Cell System Surrogate for Primary Human Hepatocytes. *Cells* **2019**, *8*, 192. <https://doi.org/10.3390/cells8020192>

## Comparison of DDApasef and DIAPasef acquisition methods for affinity-purification analyzes.

Virginie SALNOT<sup>a</sup>, Johanna BRUCE<sup>a</sup>, Morgane LE GALL<sup>a</sup>, Marjorie LEDUC<sup>a</sup>, Emilie-Fleur GAUTIER<sup>a</sup>

<sup>a</sup> Institut Cochin

DIA acquisition method is progressively implemented in proteomics protocols because it could allow more identifications and fewer missing values with faster analysis. Affinity purification samples are often analyzed in proteomics. Indeed this approach allows to identify protein partners of a protein of interest. The objective of our study is to compare the relevance of the results obtained in DDApasef (Data Dependent Acquisition) and in DIAPasef (Data Independent Acquisition) for affinity purification analysis in proteomics.

We carried out this comparison on different projects. In all the studies, denaturation of proteins was carried out on samples resulting from affinity purification using a buffer containing a detergent, SDS (Sodium Dodecyl Sulfate) with a reducing reagent and an alkylating one, the samples were then heated for 5 min at 95 °C. Then, the samples were digested using S-Trap (Suspension Trapping) by trypsin, a suspension trapping <sup>1</sup> that consists of a layer of quartz phase on a C18 phase which allows to eliminate detergents, here, the SDS. The samples were analyzed by LC-MS/MS on a U3000 RSLC nano-LC-system chromatography (Thermo Scientific) coupled to a timsTOF Pro (Bruker) with a 60 minutes overall run-time gradient ranging from 99% of solvent A containing 0.1% formic acid in milliQ-grade H<sub>2</sub>O to 40% of solvent B containing 80% acetonitrile, 0.085% formic acid in mQH<sub>2</sub>O by injecting 1 μL in DDApasef (Data Dependent Acquisition) and in DIAPasef (Data Independent Acquisition). The results from the DDA injection and DIA injection were analyzed with the Maxquant <sup>2</sup> or the DIA-NN software <sup>3</sup> respectively. The parameters of the software were without MBR (Match Between Runs) and without normalization. We compared the number of proteins identified as well as the number of known target proteins identified. A functional analysis was also performed to assess the relevance of the identified proteins regarding targets of the protein of interest and function or compartment of interest.

The depth of analysis in DIAPasef made it possible to identify a greater number of proteins than in DDApasef. The two methods were then compared according to target's detection and proportion of identified known partners with functional analysis.

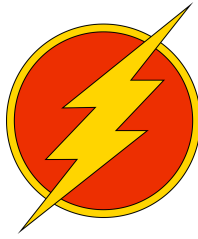
Keywords: AP-MS, Bottom-up, LC-MS/MS, DDA, DIA

<sup>1</sup>Katelyn Ludwig *et al.*, 2018

<sup>2</sup>Cox *et al.*, 2014

<sup>3</sup>Demichev V. *et al.*, 2022





## Optimization of sample preparation and nanoLC-MS/MS methods for single cell proteomic analysis

Pauline PERDU-ALLOY<sup>a</sup>, Laurence SABATIER<sup>a</sup>, Christine CARAPITO<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique, IPHC (UMR 7178), CNRS, 67000 Strasbourg, France

Understanding the various mechanisms in which proteins are involved is crucial for comprehending biochemical phenomena, including the onset and development of diseases<sup>1</sup>. Conducting qualitative and quantitative analysis of the entire proteome can provide insights into the response of an organ or a population of cells to treatments or diseases. However, analysis of proteins from an average cell population often neglects cellular heterogeneity and does not allow to distinguish healthy cells from infected cells, which cohabit in the same tissue<sup>2</sup>. The rise of single-cell proteomics is therefore meaningful because it allows the identification of the proteome expressed in individual cells with increasing depth at a given time<sup>3</sup>, with applications ranging from disease diagnosis to the response of a chemical treatment. However, despite its importance, the proteome analysis at the single-cell level remains a major analytical challenge due to the very low abundance of protein material available within a cell (<ng). Thus, every step of the workflow needs to be revisited to minimize losses during sample preparation and achieve highest sensitivity in liquid chromatography tandem mass spectrometry analysis (LC-MS/MS).

In recent years, the development of trapped ion mobility spectrometry (TIMS) coupled to a high sensitivity mass spectrometer (TimsQTOF) as well as the development of automated methods for sample preparation and sophisticated data software constitute an anchor for the analysis of protein material at the single cell scale<sup>4</sup>.

On the one hand, Hela cell digests (Pierce, Ref. 88329) were used to prepare a dilution series from 0.1 to 10 ng of injected material. On the other hand, HeLa cells were sorted, lysed and digested using the CellenONE instrument at Cellenion (Lyon). Varying cell numbers, from 1 to 20 cells were sorted, lysed, digested prior to nanoLC-MS/MS analysis.

Peptide mixtures were separated on an Aurora C18-RP 120Å (75µm x 250mm, 1.9µm) (IonOpticks) column using a nanoElute (Bruker Daltonics) coupled to a timsTOFpro2 (Bruker Daltonics) mass spectrometer. Separation was performed on a gradient from 2 to 30% ACN in 0.1% formic acid over 30min at 0.2µL/min. Data Dependent Acquisition - Parallel Accumulation and Serial Fragmentation (DDA-PASEF) and Data Independent Acquisition - PASEF (DIA-PASEF) methods were used. A comparative analysis was performed to assess the performance of two data analysis tools, MaxQuant and Proline for DDA data analysis, and Spectronaut and DIA-NN for DIA data interpretation.

In DDA-PASEF mode, around 300 proteins were identified in the 0.1 to 0.3ng region, while more than 1000 proteins could be identified in the same region in DIA-PASEF. These first results were encouraging to further explore the instrument's performances for very low amounts as these protein amounts are close to single cell protein contents (approx. 0.3ng). In isolated cells, an average of 540 proteins were identified in single cells in DDA-PASEF mode, while more than 2400 proteins were identified using a DIA-PASEF method. This remarkable outcome highlights the great capacities of DIA-PASEF for single-cell analysis.

Investigating the data processing aspect of DIA, comparing the two data processing software (Spectronaut and DIA-NN) revealed close average numbers of quantified proteins for single cells and up to 5 isolated cells.

While those initial results are promising, further optimizations are still required from sample preparation to nanoLC-MS/MS acquisition methods optimization and finally data processing to be able to unlock deeper insights into cellular heterogeneity and disease mechanisms.

<sup>1</sup> Minakshi, *Single-Cell Proteomics: Technology & Applications*. (2019).

<sup>2</sup> Marino, Zito, Accardo, Ronchi, Cozzolino, Morgillo, Rossi, Franco, *Molecular heterogeneity in lung cancer*. (2019).

<sup>3</sup> Templer, Ces, *New frontiers in single-cell analysis*. (2008).

<sup>4</sup> Meier, Park, Mann, *Trapped Ion Mobility Spectrometry and Parallel Accumulation-Serial Fragmentation in Proteomics*. (2021).

## DIAgui a shiny app to process the output from DIA-NN

Marc Antoine GERAULT<sup>a</sup>, Samuel GRANJEAUD<sup>a</sup>, Luc CAMOIN<sup>a</sup>

<sup>a</sup> INSERM

**Background:** Data-independent acquisition (DIA) proteomics is a recently-developed global mass spectrometry (MS)-based proteomics strategy. In DIA method, precursor ions are isolated into pre-defined isolation windows and then fragmented; unlike in data-dependent acquisition (DDA) where they are isolated for a specific  $m/z$ . Fragmented ions in each window are then analyzed by a high-resolution mass spectrometer. DIA technology has been increasingly utilized in various proteomics studies since it offers a broad protein coverage, high reproducibility, and accuracy. The analysis of these chimeric MS2 spectra is now facilitated thanks to AI. Many software dedicated to DIA are now available including DIA-NN which is increasingly popular in the proteomics community [1]. However, it is strongly advised to filter the DIA-NN output under R to get better quantification. For this, users must be comfortable with the R language. In addition, for the moment DIA-NN does not offer the possibility to get a TOP3 or iBAQ quantification. To overcome these problems, we offer DIAgui an R package based on Demichev's diann-rpackage that contains a user-friendly interface to process the output of DIA-NN [2].

**DIAgui R package:** DIAgui contains two main functions: *report\_process* which allows to process your report file output from DIA-NN with one R command and *runDIAgui* which launch the shiny app to process your file in an interactive way. After loading the report file, you can change the names of your fractions which are by default the path to the raw files used. Then, you can choose to extract the precursor, peptide, protein group or gene datasets from your report. You will have the possibility to filter according to some  $q$  values, to keep only proteotypic and to eliminate or not the modified peptides. For the protein group file, the Max-LFQ algorithm will be used to quantify proteins. You have two options: use the diann-rpackage or iq package method [2-3]. You can also choose to get the Top3 and iBAQ quantification. For iBAQ, you can either load a FASTA file or use seqinr package which will make a query to SwissProt database. Use a FASTA file is way faster since the shiny app doesn't have to make a query for each protein. For the other datasets, the app uses the function *diann\_matrix* which is based on the function of the same name from diann-rpackage. Nevertheless, we added the features offers by the shiny app which are the possibility to obtain the Top3 quantification and the number of peptides used for the quantification (for the gene centred dataset), and also to take the sum or the max of the intensities of same ID. In the last tab of the app, you can visualize your data with an interactive heatmap, a density and MDS plot or others, and this for each your dataset (either the ones obtained in the app or one you uploaded).

**Discussion:** DIAgui is already available on github at <https://github.com/marseille-proteomique/DIAgui> and only require from the user to know how to install R and R package, no R code skills are required. As DIA-NN is more and more used nowadays in quantitative proteomics this user-friendly interface could be very useful to the proteomics community.

### References

1. Gotti et al. J. Proteome Res., 20, 4801–4814, (2021).
2. Demichev et al. Nat Methods 17, 41-44, (2020).
3. Pham et al. Bioinformatics, 36, 8, 2611-2613 (2020).

## A new way to find drug protein targets through thermal proteome profiling (TPP) experiment

Samuel GRANJEAUD<sup>a</sup>

<sup>a</sup> CRCM, INSERM U1068

GERAULT M. (1,2,3), NORDLUND P. (3), GRANJEAUD S. (1), CAMOIN L. (1)

1 Aix-Marseille University, INSERM, CNRS, Institut Paoli-Calmettes, CRCM, Marseille Protéomique, Marseille, France

2 Centrale Marseille School, Marseille, France

3 Karolinska Institutet, Stockholm, Sweden

**Background:** TPP or Cellular Thermal Shift Assay (CETSA)[1] is a biophysical assay based on the principle of ligand-induced thermal stabilization of target proteins. Protein's melting temperature will change upon ligand interaction. By heating samples to different temperatures and quantifying proteins in solution, we can detect altered protein interactions after drug treatment. The thermal profile of each protein is established by fitting a sigmoid curve, which allows to determine its melting temperature. Such an experiment is complex and time consuming to set up. A simpler way described in [2] is to consider only the expression fold change between the treatment and the control on few temperatures points. However, the numerical analysis still needs to be validated and a robust scoring has to be determined in order to focus on proteins with the largest change in melting temperature. In addition, it would be interesting to identify proteins with similar CETSA profiles because they may belong to the same protein complex. Finally, to make these analyses easily accessible an interactive graphical exploration interface would be useful.

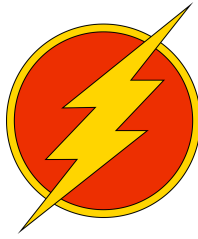
**Methods:** To determine a score for ranking the best hits, we first evaluated the mean of differences as in [3], although the mean can lower the score if a single temperature has a significant fold change. However, a hit can only have one significant fold change. Another way is to compute a weighted least-squared regression on the absolute value of the fold changes, ranked in the decreasing order with larger weights on the two largest fold changes. From this regression, we obtain the intercept which will be the score, called Stability Rate (SR). In this way, the score is always positive, which overcomes the problem explained earlier. We can plot the combined p-value (Fisher's test p-value of the two p-values of the two biggest fold changes) against the SR. The more the protein is in the top right corner of the plot, the more confident we can be that the drug is binding to this protein. To find proteins with similar thermal profiles, we also chose to compute a similarity score. A simple score is the Euclidean distance score, which is between 0 and 1, allowing these proteins to be filtered. Finally, to evaluate the relevance of their associations or their belonging to a complex, we can query the STRING database.

**Results:** As an example, we used the elutriation dataset[2]. With a combined p-value cut-off of 0.01 and an SR cut-off of 0.5 we found 501 hits in total with ten times fewer hits in G2 phase compare to S phase (as in [2]). For example CCNB1 comes in first position followed by CDK1 as shown in [2]. In S phase it's RFC3. By searching profiles similar to RFC3 (cut-off 0.3) we recovered the RFC complex. For EXOSC4, the similarity search recovered only EXOSC1. With the vimentin protein (VIM) we recovered a STRING network of 6 interactions, a potential complex formed by VIM, VCP, ACTN1 and YWHAG. This technique could be a new way to focus faster on proteins revealed by CETSA and to identify protein complex. To ease the analysis of CETSA experiments in a single and user-friendly way, we developed an R package mineCETSAapp (<https://github.com/marseille-proteomique/mineCETSAapp>) that offers a Shiny application.

### References

1. Martinez Molina et al. Science, 341(6141):84-87, (2013).
2. Dai et al. Cell 173, 1481-1494, (2018).
3. Martinez-Val et al. Nature Communications, 12:7113, (2021).





## Promising Salivary Protein Markers for Head and Neck Squamous Cell Carcinoma Prognosis

Dr Rui VITORINO<sup>a</sup>

<sup>a</sup> UNIVERSITY OF AVEIRO

Oriana Barros<sup>1,2</sup>, Vito G. D'Agostino<sup>3</sup>, Lucio Santos<sup>2</sup>, Rita Ferreira<sup>4</sup>, Rui Vitorino<sup>1,2,5\*</sup>

<sup>1</sup> Department of Medical Sciences, Institute of Biomedicine iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal

<sup>2</sup> Experimental Pathology and Therapeutics Group, Research Center of IPO Porto (CI-IPOP) / RISE@CI-IPOP (Health Research Network) and Surgical Department of Portuguese Oncology Institute of Porto (IPO Porto) / Porto Comprehensive Cancer Center (Porto.CCC)

<sup>3</sup> Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Via Sommarive 9, 38123, Trento, Italy

<sup>4</sup> LAQV-REQUIMTE, Department of Chemistry, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.

<sup>5</sup>UnIC, Department of Surgery and Physiology, Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal

*Head and neck squamous cell carcinoma (HNSCC) is one of the most common malignant cancers in the world. In advanced stages, the mortality rate is high. Stratification of HNSCC patients by molecular characteristics allows optimization of clinical management of these patients. Here, we explored the GEO and PRIDE databases to identify putative prognostic biomarkers for HNSCC. Thus, an integrated transcriptome analysis was performed to identify the major differentially expressed genes in HNSCC tissues (using GSE12452, GSE13597, GSE31056, GSE6631 and GSE3524 GEO datasets) in combination with a proteomic analysis to identify the overrepresented proteins in saliva samples from HNSCC patients (using PRIDE PXD012436 dataset). The panel of identified biomarkers was characterized using web tools, namely UALCAN, ToPP and PINAv3.0. From the combinatorial analysis of these Omics data, the salivary biomarkers with the greatest potential for clinical application for prognostic stratification were identified: ADH7, MMP9, and S100A14. Overall, this Multi-Omics approach provides comprehensive information on the potential value of combining ADH7, MMP9, and S100A14 as prognostic biomarker panel for risk stratification. Future studies should validate this panel of biomarkers in saliva samples from a large cohort of patients using targeted approaches, envisioning its translation into the clinical setting.*

## Merging SpecOMS and X!Tandem identification results using machine learning algorithm in i2MassChroQ

Olivier LANGELLA<sup>a</sup>, Thierry BALLIAU<sup>a</sup>, Marlène DAVANTURE<sup>a</sup>, Filippo RUSCONI<sup>a</sup>, Mélisande BLEIN-BICOLAS<sup>a</sup>, Dominique TESSIER<sup>b</sup>

<sup>a</sup> PAPPSO, Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE-Le Moulon, <sup>b</sup> INRA UR1268 Biopolymères Interactions Assemblages, F-44316 Nantes, France

### Introduction

SpecOMS is an Open Modification Search method published originally in David et al. 2017. SpecOMS compares experimental spectra generated by a discovery proteomics experiment to a whole set of theoretical spectra deduced from a protein database in a few minutes on a standard workstation. The procedure yields identification results that might comprise peptides bearing unknown modifications.

X!Tandem is a highly regarded classical identification engine, very efficient but unable to find peptide spectrum matches (PSMs) if the precursor ion's mass does not corresponds to the theoretical peptide's mass.

Combining the results obtained by running both SpecOMS and X!Tandem reduces the proportion of undecided PSMs compared to a classical approach. However this combination approach raises a number of questions because the two engines do compute PSM scores very differently. We address this problem in this report.

### Methods

Our i2MassChroQ quantitative proteomics software has been enhanced by integrating a machine learning algorithm called TIDD (tool-independent and data-dependent machine learning for peptide identification).

TIDD is a universal post-processing tool which supports confident peptide identifications regardless of the search engine (Li et al. 2022). i2MassChroQ computes a collection of 30 features describing the PSM (signal:intensity proportion matching theoretical peptide fragments, expected isotope ratio, amongst others). Each PSM is tagged either "target" if the determined peptide sequence was found on the targeted protein database or "decoy" if not. Based on a data subset, TIDD uses a classification Support Vector Machine (SVM) algorithm to learn how to use the features to confidently classify each PSM in the "target" or "decoy" category. TIDD computes an SVM probability that i2MassChroQ uses as a new PSM score in order to filter, group and infer peptides and proteins in the sample.

### Results

We have elaborated an efficient tool to merge identification results from different identification engines.

Depending on the sample, the number of PSMs in an MS run identification may be greatly increased.

As i2MassChroQ is able to handle results from various identification engines (pepXML, mzIdentML, Mascot dat formats are supported), the procedures described above might be applied using other identification engines.

## Evaluating Phosphopeptides Quantification Across Various Acquisition Methods and Bioinformatic Tools: Novel Standards and Their Utility in Algorithm Development

Pinar ALTINER<sup>a</sup>, Anne GONZALEZ DE PEREDO<sup>b</sup>, Carine FROMENT<sup>b</sup>, David BOUYSSIÉ<sup>b</sup>, Odile SCHILTZ<sup>b</sup>

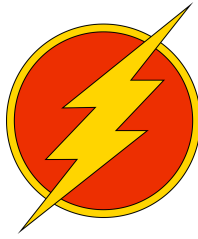
<sup>a</sup> Institut de pharmacologie et biologie structurale (IPBS), <sup>b</sup> Institut de pharmacologie et biologie structurale (IPBS), Toulouse France

Phosphoproteomics is a crucial methodological approach focused on identifying and quantifying phosphorylated peptides from complex protein samples. Indeed, quantification of phosphorylation sites across different runs and samples is important to monitor signaling events involved in biological processes. However, achieving accurate localization of phosphorylation sites can be challenging due to the ambiguity of the experimental fragmentation pattern in an MS/MS spectrum, particularly when the fragmented peptide contains multiple possible phosphorylation sites. To address this issue, several localization algorithms have been developed to assign a significance score to potential sites observed in individual MS/MS spectra. Despite the ongoing efforts to improve these algorithms, they may however not always succeed in determining the precise location of phosphorylations in each analytical run, leading to mismatching of MS signals across runs, missing values, and quantitative errors. Here, we present a computational approach aiming at mitigating such errors, as well as a standard phosphoproteomic dataset designed to benchmark bioinformatic tools in their ability to properly quantify phospho-data.

To improve the accuracy of phospho-localization, we developed a novel algorithm based on a peptide-centric approach, which gathers site-determining ions information from all MS/MS spectra related to a given phospho-peptide across all analytical runs. It generates in that way a phosphorylation evidence matrix (PEM), which summarizes fragment ion intensities for alternative phosphate positions in all the compared samples. The PEMs are then processed to cluster phospho-sites with similar fragment ion signatures across the analyzed dataset, allowing more accurate matching of MS signals. To compare the accuracy and precision of our method with existing algorithms, we conducted a benchmarking study using two different phosphorylated standards, based either on a library of 180 synthetic phospho-peptides with known localizations or on phosphopeptides enriched from mouse T-cells, spiked at various concentrations into *E. coli* background. We analyzed those samples using different MS instruments (Thermo Exploris and Bruker timsTOF) and various acquisition methods (Data Dependent Acquisition, Data Independent Acquisition), with or without ion mobility separation.

We present here initial results based on the use of PEM heatmaps, to determine the correct localization of phosphosites by clustering fragment ion signatures of several MS/MS spectra. As a next step, a scoring function and a target-decoy method will be implemented, enabling automatic determination of phospho-sites position for a given FDR. Although more development is still needed for a direct comparison of this method with existing algorithms, we also show here the first results obtained when comparing quantitative phosphoproteomic approaches using our calibrated phospho-standards. We show the utility of such samples and standard datasets to benchmark different acquisition methods and different bioinformatic tools (Proline, MaxQuant, ProteomeDiscoverer, DIA-NN, Spectronaut) in terms of phosphosites localization accuracy and phosphopeptides quantification accuracy.





## Comparison of tissue homogenization protocols for optimization of quantitative proteomics and phosphoproteomic analysis

Marta ROJAS-TORRES<sup>a</sup>, Ana MARTINEZ-VAL<sup>b</sup>, Lucia BELTRAN-CAMACHO<sup>c</sup>, Sara ESLAVA-ALCON<sup>a</sup>, Jesper V. OLSEN<sup>b</sup>, M<sup>ª</sup> Carmen DURAN-RUIZ<sup>a</sup>

<sup>a</sup> Biomedicine, Biotechnology and Public Health, University of Cadiz / Inhibica, <sup>b</sup> Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, <sup>c</sup> Cell biology, Physiology and Immunology, Imibic / University of Cordoba

Tissue homogenization and cell lysis are the first steps in virtually all molecular biology and diagnostic techniques. The reproducible and efficient extraction of proteins from tissue biopsies for quantitative analysis is a critical step in biomarker and translational research, as it allows maximal recovery from limited quantity specimens and minimize the effects of intratissue heterogeneity. Tissues are specially challenging to phosphoproteomics analysis, since it is crucial to inactivate phosphatases that might be active during the homogeneization process. Nowadays, there are several methods and specific devices focused on the optimization of the abovementioned. In this regard, some of the most used are automatic dry pulverizer (Covaris <sup>®</sup>) and Precellys <sup>®</sup> 24 (Bertin technologies), alone or in combination with proteome stabilizer (Denator AB).

In the current study we have evaluated, by application of a label-free quantitative (LFQ) proteomic approach, the MS profiles obtained from different tissue samples (liver, kidneys, spleen, lungs, heart and brain) after homogenization using Precellys, Denator with Precellys and Covaris, in order to identify the optimal approach for each tissue when performing a total proteome and phosphoproteome quantitative and qualitative analyses. After the homogenization and protein extraction step, proteins were digested using Protein Aggregation Capture in the KingFisher platform. Moreover, phospho-peptide enrichment was performed in the KingFisher platform using Ti-IMAC HP beads. Samples were analysed on a Orbitrap Exploris (Thermo Scientific) coupled to an EvosepOne LC platform, operating in Data Independent acquisition (DIA). Spectronaut<sup>™</sup> and Perseus software were used for data processing and statistical analysis.

We evaluated the performance on each homogeneization protocol in terms of overall protein and phospho-peptide coverage, as well as enrichment efficiency and applicability of the method for large scale studies. All methods were comparable in terms of protein coverage, as well as on phospho-peptide enrichment efficiency. On average, we quantified 5,000 proteins per tissue. The phospho-peptide coverage, as well as other PTMs such as oxidation, revealed higher tissue specificity reflecting the different physiology and signaling of each tissue. However, we did not find any consistent bias on any of the pipelines studied here. Several differences were found depending on the tissue and homogenization approach, especially from the processing time required in each pipeline. Quantification results regarding total proteome and phosphoproteome will be further discussed.



## How to identify, at a lower cost, 3 to 4 times more peptides from agri-food hydrolysates?

Dr Christophe FLAHAUT<sup>a</sup>, Tahirimanjaka ANDRIAMANARIVO<sup>b</sup>, Barbara DERACINOIS<sup>b</sup>, Benoit CUDENNEC<sup>b</sup>, Rozenn RAVALLEC<sup>b</sup>

<sup>a</sup> UNIVERSITÉ D'ARTOIS, <sup>b</sup> UNIVERSITÉ DE LILLE

The valorization of by-products from the agro-food industry represents a promising approach for a transition to a circular economy due to the large quantities of agro-food by-products generated each year [1]. Many agro-food by-products are an excellent source of proteins that can be valorized by enzymatic hydrolysis process to obtain peptides with high added value such as nutritional-, biological-, organoleptic- or techno-functional-properties [2,3]. However, although many methods of peptidome analysis can be used, the complete identification of peptide (determination of the amino acid sequence) content from these by-products still represents a challenge in the field of peptidomics due to the complexity of biological matrices and the use of non-specific enzymes[4].

To address this challenge, this study aims to develop a two-dimensional (2D) peptide separation/fractionation strategy consisting of a cation-exchange chromatography followed, off-line, with a reversed-phase high performance liquid chromatography (RP-HPLC) coupled to a high-resolution mass spectrometer (HR-MS). The optimization of the separation/fractionation operating conditions was carried out using *i*) a milk protein hydrolysis and *ii*) a mixture of fifteen proteins from different sources subjected to an *in vitro* simulated gastrointestinal digestion. Mass spectrometry data and their bioinformatics retreatment highlight the interest of a 2D-chromatographic strategy in comparison to a classical strategy by simple RP-HPLC-HR-MS analysis.

### References:

- [1] B. Socas-Rodríguez, G. Álvarez-Rivera, A. Valdés, E. Ibáñez, A. Cifuentes, *Trends Food Sci. Technol.* **2021**, *114*, 133-147.
- [2] B. Bhaskar, L. Ananthanarayan, S. N. Jamdar, *Food Sci. Biotechnol.* **2019**, *28*, 43-52.
- [3] Md. S. Islam, H. Wang, H. Admassu, A. A. Sulieman, F. A. Wei, *Process Biochem.* **2022**, *116*, 116-125
- [4] S. Martini, L. Solieri, D. Tagliazucchi, *Curr. Opin. Food Sci.* **2021**, *39*, 51-59.

## Comparison of histidine phosphorylation protocols in eukaryotes

Maryam ABOUTIMAN<sup>a</sup>, Virginie SALNOT<sup>b</sup>, Arije MESRI<sup>b</sup>, Michel AUTIN<sup>b</sup>, Johanna BRUCE<sup>b</sup>, Cédric BROUSSARD<sup>b</sup>, Marjorie LEDUC<sup>b</sup>, François GUILLONNEAU<sup>b</sup>, Olivier HERMINE<sup>c</sup>, Emilie-Fleur GAUTIER<sup>b</sup>

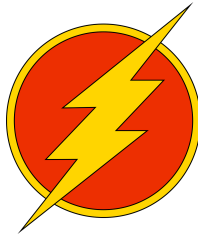
<sup>a</sup> INSTITUT COCHIN/ INSTITUT IMAGINE, <sup>b</sup> INSTITUT COCHIN, <sup>c</sup> INSTITUT IMAGINE

Phosphorylations are among the most common and the most studied post translational modifications. Commonly studied phosphorylations are on serine, threonine, and tyrosine because of their lower lability level compared to other phosphorylated amino acids. As, in eukaryotes, 9 amino acids can be phosphorylated, forming in total four types of phosphates-protein link. Among them, we can find histidine phosphorylations. This phosphorylation is estimated to represent 6% of the total phosphorylations in eukaryotes. However, until recently, there has been few research on this phosphorylation, in eukaryotes. This is principally due to the lack of phosphoprotective techniques, that would prevent the high lability of the phosphorylation. Yet, since 2015, new protocols and tools have been developed allowing histidine phosphorylation study. Our objectives were 1) to set up different protocols of histidine phosphorylation analysis by mass spectrometry 2) to compare them to chose the most performant 3) to improve this protocol.

In this study, we compared two different histidine phosphorylation enrichment protocols to protocols using iMAC and TiO<sub>2</sub> usually employed for O-phosphorylation analyses in proteomics. Doing so, we compared a protocol of iMAC in non-acidic condition adapted from Potel CM et al., MCP, 2018, and SAX protocol adapted from Hardman G et al., EMBO, 2019. In a UT7epo cell line, we deprived and restimulated during 15 minutes with EPO the cells to induces phosphorylations. We then performed each enrichment approach on 4 independent replicates. All the analyses were performed on an Orbitrap Fusion mass spectrometer hyphenated with a RSLC U300 nanochromatography chain (Thermo Fisher) and Maxquant software for the identification and quantification of phospho-sites. We compared the number of identified phosphorylation on serine, threonine, tyrosine (STY), and histidine, but also the 5 other non-canonical phosphorylations.

With a threshold at 0.9 for the localization probability and 0.1 for the Posterior Error Probability (PEP), we were able to identify a mean of 3257 phospho-sites for the SAX protocol, 1406 phospho-sites for the iMAC in non-acidic condition protocol, 2715 phospho-sites for the iMAC in acidic condition protocol, and 3095 phospho-sites for the TiO<sub>2</sub> in acidic conditions protocol. Among those phospho-sites, the SAX protocol allowed us to identify a mean of 821 phospho-sites on STY, the TiO<sub>2</sub> and iMAC protocol in acidic conditions allowed us to identify, respectively, a mean of 2253 and 2815 phospho-sites on STY. Among them, for the SAX protocol, we find known STY phosphorylations induced by EPO stimulation like the tyrosine phosphorylation of STAT5 or ERK-2 protein, which confirm our data by comparison to previously published one. In non-acidic condition we were able to yield a mean of 251 histidine phospho-sites for SAX, 39 for iMAC in non-acidic conditions, 36 and 30 for iMAC and TiO<sub>2</sub> in acidic condition. However, even if the SAX protocol allowed the identification of several hundred of histidine phosphorylation sites, we observed a general lack of reproducibility of the results between replicates. Consequently, we improved our protocol. This improvement allowed to maximize the localization probability with a mean of 66% of identified histidine phospho-site with a localization probability over 0.9 when the previous SAX protocol only scored 30% of its identified histidine phospho-site with the localization probability over 0.9. We were also able to confirm two histidine phosphorylations of interest by immunoprecipitating them and revealing the histidine phosphorylation by western blot.

To conclude, we have set up a SAX protocol that we improved allowing the identification of several hundred of histidine phospho-sites. However, to this point we only have few technical replicates, and these data need to be confirmed on biological replicates. Finally, the reproducibility of this new protocol needs to be assessed.



## A novel DIA label-free proteomics workflow to study protein glycation in plasma samples

Sarahi JARAMILLO ORTIZ<sup>a</sup>, Charlotte BRUN<sup>b</sup>, Fabrice BERTILE<sup>b</sup>, Francois CRISCUOLO<sup>c</sup>, Christine SCHAEFFER<sup>b</sup>

<sup>a</sup> IPHC-LSMBO - Université de Strasbourg/CNRS, 25 rue Becquerel 67087, Strasbourg, France/Infrastructure Nationale de Protéomique ProFI - FR2048, 67087, <sup>b</sup> IPHC-LSMBO - Université de Strasbourg/CNRS, 25 rue Becquerel 67087, Strasbourg, France/ Infrastructure Nationale de Protéomique ProFI - FR2048, 67087 Strasbourg, France, <sup>c</sup> IPHC-DEPE - Université de Strasbourg/CNRS, 23 rue Becquerel 67087, Strasbourg, France

Glycation is a post-translational modification (PTM) induced by the non-enzymatic reaction between reducing sugars and primary amino groups of proteins, resulting in the formation of Amadori compounds and advanced glycation end products (AGEs). Protein glycation has been associated with diabetes and aging-related pathologies<sup>1</sup>. Although very few proteomics studies have been conducted on glycation research, the preferred recurrent analytical platform to study this PTM is mass spectrometry with enzymatic digestion (bottom-up approach)<sup>2</sup>. Glycated proteins in plasma are relatively low abundant and comprehensive glycation studies require the enrichment of modified peptides coupled with advanced mass spectrometry-based methods. To date, no glycated proteome studies have been published in which data-independent acquisition (DIA) approach has been applied.

In this study, a label-free quantitative approach was developed to study glycations in undepleted plasma. Two sample preparation workflows were optimized. The first approach involved isopropanol gradient peptide fractionation (CIF) with paramagnetic beads to decomplexify the proteome prior to LC-MS/MS analysis. The second one sequentially combined CIF and automated boronate affinity chromatography (BAC) enrichment on an AssayMAP Bravo platform. Herein, we evaluated dia-PASEF methods on a timsTOF Pro platform.

Commercial human plasma proteins (450 µg) were digested overnight using trypsin/Lys-C. In the first approach, CIF fractionation was performed on tryptic peptides and then analyzed by LC-MS. In parallel, for the second approach, digested peptides were sequentially fractionated and glycated peptides were enriched prior to LC-MS/MS analysis. CIF was performed using paramagnetic beads grafted with carboxylate groups; peptides were eluted into five fractions (F<sub>1</sub> ACN; F<sub>2</sub> 80, F<sub>3</sub> 60, F<sub>4</sub> 30, and F<sub>5</sub> 0% isopropanol). Custom BAC cartridges high capacity (BeneLab, Zagreb, Croatia) were used to perform automated enrichment on the AssayMap Bravo platform (Agilent). All samples were prepared in triplicate and analyzed by nanoLC-MS/MS using a TimsTOF Pro (Bruker) in DIA mode. Data were processed with Spectronaut 17.4 software (Biognosys) and DIA-NN 1.8.1<sup>3</sup>.

Our results showed that, in terms of total unmodified proteins and peptides, using DIA mode increases the number of identifications by a factor of two compared to previous experiments performed in DDA mode, albeit not when using sample fractionation.

Regarding modified peptides and proteins in control samples (non-fractionated and non-enriched); 7 times more peptides were identified in DIA compared to DDA acquisition mode and 78% more proteins. The DIA acquisition mode analysis of fractionated CIF samples enabled a 45% increase in the number of modified proteins and a 47% increase in the number of modified peptides and glycation sites recognized. The combined implementation of CIF fractionation and BAC enrichment of glycated peptides in DIA mode acquisition achieved to increase in the number of glycated peptides and glycation sites (twofold) as well as proteins (three-fold) compared to non-fractionated and non-enriched samples.

The bottom-up proteomic method developed has successfully increased the identification and quantification of numerous peptides and proteins modified by Amadori products in human undepleted plasma samples. To our knowledge, this is the first glycated proteome using a combination of fractionation/boronate-based enrichment strategies coupled with comprehensive DIA-MS. This approach will now be applied to samples from different avian species in order to better understand how birds manage their usually high levels of blood glucose.

1. A. Shin *et al.*, *Adv Clin Chem* 113:101-156 (2023).
2. C. Seonghyeon *et al.*, *Reviews in Analytical Chemistry* 41: 83-97 (2022).
3. V. Demichev *et al.*, *et al. Nat Methods* 17, 41-44 (2020).

## Hunting for mouse-specific variants of histone H3 by middle-down proteomics

Julie MANESSIER<sup>a</sup>, Tina RAVNSBORG<sup>b</sup>, Hassan HIJAZI<sup>a</sup>, Marie COURÇON<sup>a</sup>, Sabine BRUGIÈRE<sup>a</sup>, Ole JENSEN<sup>b</sup>, Delphine PFLIEGER<sup>a</sup>

<sup>a</sup> Exploring the Dynamics of Proteomes (EDyP), University Grenoble Alpes, CEA, Inserm, IRIG-BGE UA13, CNRS/CEA FR2048, <sup>b</sup> Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, University of Southern Denmark

In the nucleus of eukaryotes, DNA wraps around nucleosomes, protein octamers made up of two copies of histones H2A, H2B, H3 and H4, to form a structure called chromatin. Regulation of gene expression is tightly controlled by the incorporation of specific variants into nucleosomes and by a vast array of transient post-translational modifications (PTMs). When using mouse as a model system, it may be relevant to consider additional variants of histone H3 which have been recently described (1,2) In particular, the transcripts of H3mm7 and H3mm13 appear to be of high levels in mouse testis and brain, which raises the question of considering the corresponding protein variants during proteomic analysis of histones.

Variants H3mm7 and H3mm13 are in particular distinguishable from canonical H3 and from the usually considered variant H3.3 by a few amino acids between residues 27 and 31. Mass spectrometry analysis is very well suited to discriminate between sequence variants and to characterize a range of PTMs. We have considered a classical protocol of propionylation of endogenously non-modified lysines followed by trypsin digestion, to perform the bottom-up analysis of histones. Additionally, we have assessed the possibility to identify mouse-specific variants by the middle-down analysis of the long N-terminal tail of histone H3 spanning its 50 first amino acids. We will show MS/MS results obtained by LC-MS/MS analyses using a C18-HCD and HILIC-ETD couplings, and insist on the difficulty to clearly distinguish between isobaric (variant x PTM) combinations.

### References

(1) Maehara, Kazumitsu et al. "Tissue-specific expression of histone H3 variants diversified after species separation." *Epigenetics & chromatin* vol. 8 35. 17 Sep. 2015

(2) Harada, Akihito et al. "Histone H3.3 sub-variant H3mm7 is required for normal skeletal muscle regeneration." *Nature communications* vol. 9,1 1400. 11 Apr. 2018

## Optimized MS/MS pre-processing for reliable Histone Post-Translational Modification (hPTMs) identification and localization

Hassan HIAJZI<sup>a</sup>, Marie COURÇON<sup>b</sup>, Sabine BRUGIÈRE<sup>b</sup>, Anne-Marie HESSE<sup>b</sup>, Christophe BRULEY<sup>b</sup>, Delphine PFLIEGER<sup>b</sup>

<sup>a</sup> Exploring the Dynamics of Proteomes (EDyP), University Grenoble Alpes, CEA, Inserm, IRIG-BGE UA13, CNRS/CEA FR 2048, CEA Grenoble, <sup>b</sup> Exploring the Dynamics of Proteomes (EDyP), University Grenoble Alpes, CEA, Inserm, IRIG-BGE UA13, CNRS/CEA FR 2048, CEA Grenoble, Grenoble, FRANCE

Histones are basic proteins abundant in lysine (K) and arginine (R) residues. Tetramers of different histone proteins (H2A, H2B, H3 and H4) assemble to form along with ~ 146 DNA base pairs the so-called nucleosome structures. Protruding outside the nucleosome are unstructured histone tails (~ 50 A.A.) that are prone to a wide spectrum of post-translational modifications (PTMs). Additional histone H3 variants have been discovered in mouse and differ in a few amino acids from the canonical ones [i]. Both histone PTMs and variants are of central importance in epigenetics studies.

Mass spectrometry (MS) has proven crucial for high-throughput identification and quantification of histones and histone PTMs. Nonetheless, data processing and interpretation of histone mass spectra is still a daunting task due to the absence of a decoy strategy for database search and the presence of ample isobaric peptides due to PTM and/or amino acid variations. So far, hundreds of PTMs have been discovered, many of which occur on lysine, including methylation, acetylation, lactylation and crotonylation. To perform a sequence database search, MS raw files are converted into peaklists by a preprocessing software. This helps to improve the accuracy of peptide identification by reducing noise in the data. To validate the presence of a PTM in histone peptides, it is important to look at the low mass range of the tandem mass spectra (MS/MS) which contain PTM-specific fragment ions, so-called diagnostic ions (e.g, 156.1019 and 152.1070 m/z for lactyl- and crotonyl-lysines, respectively) [ii]. However, we noticed that the diagnostic ions could either be lost or dampened in the preprocessing step depending on the selected parameters, which strips a crucial information in validating PTMs from the MS/MS spectra.

Exploring this effect, we here present Mascot database search results validated with Proline [iii] showing the variation of the number of peptide-spectrum-matches (PSMs) and number of identifications after using three different pre-processing methods. More specifically, we compare the results between Mascot Distiller and our in-house pre-processing tool, MGFboost, conjugated or not with pClean [iv]. Furthermore, we emphasize on the retention time (RT) as an important dimension to rule out false positive identifications in ambiguous spectra by showing the RT of 31 synthetic peptides of different PTM-variant histone combinations.

In all, histone PTM confident characterization requires validating multiple aspects including the fragmentation profile, the RT, and the presence of diagnostic ions.

[i] Maehara, Kazumitsu, et al. "Tissue-specific expression of histone H3 variants diversified after species separation." *Epigenetics & chromatin* 8.1 (2015): 1-17.

[ii] Hseiky, Alaa, et al. "Small mass but strong information: Diagnostic ions provide crucial clues to correctly identify histone lysine modifications." *Proteomes* 9.2 (2021): 18.

[iii] Bouyssié, David, et al. "Proline: an efficient and user-friendly software suite for large-scale proteomics." *Bioinformatics* 36.10 (2020): 3148-3155.

[iv] Deng, Yamei, et al. "pClean: an algorithm to preprocess high-resolution tandem mass spectra for database searching." *Journal of proteome research* 18.9 (2019): 3235-3244.

## Improved immunoaffinity enrichment methods for arginine methylation

Dr Maxime JACQUET<sup>a</sup>, Barry M. ZEE<sup>b</sup>, Hayley J. PECKHAM<sup>b</sup>, Charles L. FARNSWORTH<sup>b</sup>, Alissa J. NELSON<sup>b</sup>, Jian Min REN<sup>b</sup>, Michael C. PALAZZOLA<sup>b</sup>, Kathryn E. ABELL<sup>b</sup>, Matthew P. STOKES<sup>b</sup>

<sup>a</sup> Cell Signaling Technology, Inc., <sup>b</sup> Cell Signaling Technology, Inc., Danvers MA, United States

### INTRODUCTION

Immunoaffinity purification (IAP) coupled with LC-MS/MS is an established technique for the study of posttranslational modifications (PTMs). Recently we introduced an antibody-based magnetic bead conjugate IAP method that enables identification of sites of PTMs including lysine ubiquitination (KGG), lysine acetylation (AcK) and tyrosine phosphorylation (pY) with improved sensitivity and specificity over preexisting agarose bead methods. Here we present three novel magnetic bead IAP methods that enable identification of monomethylated (Me-R), symmetric dimethylated (SDMA), and asymmetric dimethylated (ADMA) arginine sites, with improvements over the respective agarose bead methods.

Arginine methylated proteins are involved in processes including transcriptional regulation and signal transduction. Arginine methylation is carried out by the arginine N-methyltransferase (PRMT) family of enzymes. Though all PRMT proteins catalyze monomethylation, Type I PRMTs such as PRMT1 add an additional methyl group to produce ADMA, while Type II PRMTs such as PRMT5 produce SDMA. Recent reports have highlighted the relevance of PRMT5, and SDMA, to cancer and stem cell self-renewal.

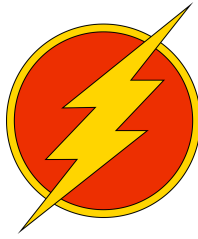
### METHODS

Tryptic mouse and human peptides were used as input material for Me-R, SDMA, and ADMA magnetic IAP beads. Peptides were incubated with beads using cold 1x IAP buffer for 2 hours at 4 °C. Beads were washed with cold 1x IAP buffer, then water, and bound peptides were eluted with trifluoroacetic acid. Eluted peptides were desalted using C18 Stage-Tips and analyzed on ThermoFisher Q-Exactive or Fusion Lumos mass spectrometers using data-dependent acquisition (DDA). Peptides were resolved using a 90 or 120min reversed phase gradient from 7.5 to 32% acetonitrile on a 25cm or 50cm C18 column. Peptides were identified by Comet with 1% of total identifications mapped to reverse sequences. Skyline software was used for MS1 peak review and quantification. For automation on Kingfisher Apex robot, input peptides were resuspended in bulk and aliquoted to deep well plates. After IAP on Kingfisher Apex robot, all enriched peptides were manually desalted as described above.

### CONCLUSIONS

The new Me-R, SDMA, and ADMA magnetic IAP beads provide 6-60% improvement in number of unique modified peptides identified and nearly 100% improvement in PTM capture specificity across a variety of different cell and tissue sample types. The ease of handling of magnetic beads facilitates manual benchwork and implementation on automation platforms.





## Deciphering the signaling mechanisms of the C-type lectin receptor Dectin-1 in myeloid cells using high throughput proteomics

Dr Anne GONZALEZ DE PEREDO<sup>a</sup>, Rim ABDERRAHIM<sup>a</sup>, Alexandre STELLA<sup>a</sup>, Karima CHAOUI<sup>a</sup>, Marie LOCARD-PAULET<sup>a</sup>, Benjamin RAYMOND<sup>a</sup>, Thibaut SANCHEZ<sup>a</sup>, Véronique LE CABEC<sup>a</sup>, Frédéric LAGARRIGUE<sup>a</sup>, Odile SCHILTZ<sup>a</sup>, Olivier NEYROLLES<sup>a</sup>, Yoann ROMBOUTS<sup>a</sup>

<sup>a</sup> CNRS, Institut de Pharmacologie et de Biologie Structurale

Dectin-1 is a Pathogen Recognition Receptor expressed on myeloid immune cells, known to play a major role in the antimicrobial immune response. It belongs to the C-type lectin Receptors (CLR) family and can bind to carbohydrate structures, such as  $\beta$ -glucans present on pathogens. This recognition triggers the phosphorylation of key tyrosine residues in the ITAM domain of Dectin-1, leading to the activation of SYK-dependent and independent signaling pathways, and finally to the secretion of pro-inflammatory cytokines and chemokines as well as to the production of reactive oxygen species and nitric oxide. Even though Dectin-1 is an important CLR known to be involved in many diseases such as fungal infections, inflammation and autoimmune diseases, its signaling mechanisms remain unclear. In this project we applied unbiased global proteomic approaches to explore Dectin-1 signaling and identify some potential medical targets. We used global phosphoproteomics to obtain a time-resolved characterization of the early signaling events taking place in the first minutes following Dectin-1 engagement, and identify key intracellular signaling proteins involved in that pathway. In addition, we established a high-throughput approach for secretome analysis, to monitor in a fast and global manner the inflammatory mediators secreted by immune cells 24 h after Dectin-1 stimulation.

To that aim, we generated ER-Hoxb8 immortalized murine cell lines from either WT or Dectin-1 KO progenitors, that were in-vitro differentiated into dendritic cells (DC). Cells were stimulated with depleted Zymosan, a canonical ligand for Dectin-1, during different time length. Phosphorylated peptides were enriched using a two-step protocol based on TiO<sub>2</sub> and phospho-Tyrosine antibody affinity purification, and resulting phosphopeptides were analyzed on an Orbitrap-Exploris instrument equipped with a FAIMS interface. To monitor the secretory phenotype of DCs, a high-throughput approach for secretome analysis was developed using the Evosep chromatographic system and DIA-PASEF analysis on a TIMS-TOF SCP mass spectrometer.

Using this analytical pipeline, we could globally detect more than 10,000 phosphosites with localization confidence  $>0.75$ , and identify after statistical analysis more than 300 phosphosites significantly regulated upon Dectin-1 stimulation, showing different kinds of kinetic profiles. These included sites from known proteins of the Dectin-1 signaling pathway, as well as novel targets that were selected for further functional validation by generating the corresponding KO cells. We could also provide a detailed characterization of the DC phenotypic response following Dectin-1 activation, with more than 500 proteins significantly over-represented in the secretome of activated cells, including many canonical inflammatory mediators such as TNF or IL6. The high-throughput secretomic analytical workflow based on the Evosep system and fast DIA-MS runs provides a quick and comprehensive read-out to monitor the phenotypic response of immune cells in many different condition. It thus allows the systematic screening of different KO cells to assess the importance of newly identified signaling proteins in on-going validation studies.



## Evaluating the benefit of dia-PASEF approaches and sample-specific database strategies for metaproteomics of very complex microbiomes

Thibaut DUMAS<sup>a</sup>, Olivier PIBLE<sup>a</sup>, Guylaine MIOTELLO<sup>a</sup>, Lucia GRENGA<sup>a</sup>, Kristina MARX<sup>b</sup>, Jonathan KRIEGER<sup>b</sup>, Pierre-Olivier SCHMIT<sup>c</sup>, Jean ARMENGAUD<sup>a</sup>

<sup>a</sup> Laboratory «Innovative technologies for Detection and Diagnostics» CEA-Marcoule, DRF/JOLIOT/DMTS/SPI/Li2D PRAE, <sup>b</sup> Bruker Daltonik GmbH & Co. KG, <sup>c</sup> Bruker France S.A.

Understanding interaction of microorganisms with their host is crucial in microbiology. The identification and quantification of proteins from a complex mixture of a large variety of organisms, known as Metaproteomics, has emerged recently as a unique tool to get functional and taxonomical insights into microbiota and is currently one of the most challenging areas in proteomics. Species-specific signals can be low, highly diverse, and the search space is extremely large. For these reasons, the increased selectivity potential of LC-IMS-MS based approaches might prove beneficial for metaproteomics analyses. In this communication, we are using a specific sample representative of gut microbiome to benchmark dia-PASEF approaches against PASEF approaches and investigate the benefit of sample-specific database strategies.

A gut microbiome tryptic digest (200 or 800 ng) was injected on a 25cm X 75µm pulled emitter column (IonOptiks). Nano-HPLC separation was performed with a 36, 66 or 100 min gradient using a nanoElute (Bruker) connected to a timsTOF HT mass spectrometer (Bruker). LCMSMS data were acquired in PASEF or dia-PASEF acquisition mode. PASEF data have been processed in real-time on PaSER (Bruker) or Mascot (MatrixScience), searching against an NCBI database. Dia-PASEF data have been searched against a reduced protein sequence database using TIMS DIA-NN on PaSER (Bruker) or Spectronaut 17 (Biognosys). The protein sequence reduced database was constructed after confidently proteotyping the most abundant organisms present in the standard sample. It comprised 893,451 protein entries from 57 taxa.

The results initially obtained from a Mascot search from all 200 ng injections showed only a slight increase of the protein group ID number while jumping from a 36 min to a 66 min gradient and from a 66 min gradient to a 100min gradient (12 and 15%, respectively). There was more effect by quadrupling the injected amount (+18% protein group ID's while jumping from 200 to 800ng injected with the 100 min gradient). In fine, 2110 protein groups could be identified from a 200ng injection and a 66min gradient. On average, 4 peptide sequences were identified for each protein Group ID's. Using the ProLucid Algorithm on PaSER while allowing to take the peptides's collisional cross section (CCS) value into account for the scoring process (TIMScore), a high stringency allowed to double the number of identified protein groups, while the number of sequences/proteins was down to 3. This suggest that the use of the TIMScore allowed to identify more of the low-intensity peptides, which also resulted into adding protein groups described by a lower number of peptides per protein. In the best conditions (800ng-100min), up to 6,537 protein group corresponding to 19,962 peptide sequence could be identified. The use of dia-PASEF with the restricted database allowed to boost the number of protein group ID's: 8,827 for 35,516 peptides for the 200ng/66 min injection (+35% protein group ID's, +78% peptide sequences). Meanwhile, the average number of peptides per protein group was 4, giving high confidence in identification. In conclusion, we have developed a novel strategy based on the construction of sample-specific database for DIA-based metaproteomics that allows quicker data treatment and superior results when used with a dia-PASEF approach. For fair comparison purposes, we will also disclose the results obtained with PASEF and the sample-specific database.

## Contribution of the Hsp90 chaperone in bacterial proteostasis

Marie CORTEGGIANI<sup>a</sup>, Manel KHELIL BERBAR<sup>b</sup>, Maya BELGHAZI<sup>b</sup>, Régine LEBRUN<sup>b</sup>, Olivier GENEST<sup>a</sup>

<sup>a</sup> BIP (UMR 7281), IMM, CNRS - Aix Marseille Université, France, <sup>b</sup> Plateforme Protéomique de l'IMM, Marseille-Protéomique, CNRS - Aix Marseille Université, France

Protein homeostasis (proteostasis) is controlled in every living cells by a complex network of chaperones and proteases that together ensure folding or degradation of proteins. Many chaperones and proteases are named “heat shock proteins” (Hsp) since their role becomes more important under stress conditions (including heat stress) known to affect protein folding. One of these proteins, Hsp90, is an ATP-dependent chaperone present from bacteria to human. In bacteria, Hsp90 is essential under heat stress in the environmental bacterium *Shewanella oneidensis*, and it is involved in the virulence of some pathogenic bacteria. However, only few bacterial proteins taken in charge by Hsp90 - called Hsp90 clients - have been identified to date.

The goal of our study is to evaluate the importance of the Hsp90 chaperone on the bacterial proteome and to identify new clients. Using *Shewanella oneidensis* as a bacterial model, we compared the proteome of a WT strain and a strain that does not produce Hsp90 ( $\Delta$  *hsp90* strain) by label free quantitative proteomics. Although the abundance of only few proteins was significantly different under non-stress conditions, several dozens of proteins varied under stress conditions. We focused on the proteins that were less abundant in the  $\Delta$  *hsp90* strain than in the WT strain. Indeed, some of these proteins could be Hsp90 clients assuming that, in the absence of Hsp90, the clients are misfolded and therefore degraded. This hypothesis was confirmed by in vivo targeted degradation tests in the absence of Hsp90, as well as protein-protein interactions assays. Altogether, these experiments allowed the identification of new *bona fide* clients of Hsp90, and will undoubtedly lead to a better global understanding of the role of Hsp90 to control bacterial proteostasis.

## Re-Routing of *Chlamydomonas reinhardtii* Mutants Metabolism demonstrated by Quantitative Proteomics.

Dr Régine LEBRUN<sup>a</sup>

<sup>a</sup> CNRS

Cassy Gérard <sup>1</sup>, **Régine Lebrun** <sup>2</sup>, Erwan Lemesle <sup>1</sup>, Luisana Avilan <sup>1</sup>, Kwang Suk Chang <sup>3</sup>, EonSeon Jin <sup>3</sup>, Frédéric Carrière <sup>1</sup>, Brigitte Gontero <sup>1</sup>, Hélène Launay <sup>1</sup>

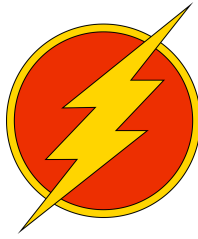
<sup>1</sup> Aix Marseille Univ, CNRS, BIP, UMR 7281, IMM, FR3479, 31 Chemin J. Aiguier, 13 402 Marseille Cedex 20, France ; cgerard@imm.cnrs.fr; carriere@imm.cnrs.fr; bmeunier@imm.cnrs.fr; helene.launay@univ-amu.fr

<sup>2</sup> **Plate-forme Protéomique, Marseille Protéomique (MaP), IMM FR 3479**, 31 Chemin Joseph Aiguier 13009 Marseille, France, [rlebrun@imm.cnrs.fr](mailto:rlebrun@imm.cnrs.fr)

<sup>3</sup> Department of Life Science, Research Institute for Natural Sciences, Hanyang University, Seoul, 04763, Republic of Korea; [kschang@hanyang.ac.kr](mailto:kschang@hanyang.ac.kr) (K.S.C.); [esjin@hanyang.ac.kr](mailto:esjin@hanyang.ac.kr) (ES.J.)

The chloroplast protein CP12 is involved in the dark/light regulation of the Calvin-Benson-Bassham cycle that is responsible for CO<sub>2</sub> assimilation. In particular CP12, under dark inhibits two enzymes: glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK), but it may have other functions related to stress. We knocked out the unique *CP12* gene to prevent its expression in *Chlamydomonas reinhardtii* (DCP12). The growth rates of both wild-type and DCP12 cells were nearly identical, as was the GAPDH protein abundance and activity in both cell lines. On the contrary, the abundance of PRK and its specific activity were significantly reduced in DCP12, as revealed by **relative quantitative proteomics**. Many proteins involved in redox homeostasis and stress responses were more abundant and the expressions of various metabolic pathways changed in the absence of CP12. These results highlight CP12 as a moonlighting protein with additional functions beyond its well-known regulatory role in carbon metabolism.

Gerard, C., **Lebrun, R.**, Lemesle, E., Avilan, L., Chang, K.S., Jin, E., Carriere, F., Gontero, B., and Launay, H. (2022). Reduction in Phosphoribulokinase Amount and Re-Routing Metabolism in *Chlamydomonas reinhardtii* CP12 Mutants. *International Journal of Molecular Sciences* 23.



## Flash MS/MS identification of microbial isolates

Madisson CHABAS<sup>a</sup>

<sup>a</sup> CEA

### Flash MS/MS identification of microbial isolates

Madisson Chabas<sup>1,2</sup>, Jean-Charles Gaillard<sup>1</sup>, Olivier Pible<sup>1</sup>, Béatrice Alpha-Bazin<sup>1</sup>, Jean Armengaud<sup>1</sup>

<sup>1</sup> CEA, ProGénoMix platform; <sup>2</sup> Montpellier University

Rapid identification of microorganisms is vital for medical diagnostics, food product control, and screening for new enzymatic potentials for biotechnological applications. High throughput analytical platforms have been established in analytical laboratories to process thousands of samples per day. Furthermore, culturomics is very much in vogue to isolate and detail the components of the microbiota. MALDI-TOF MS technology has become the reference method for the rapid identification of microorganisms in microbiology for clinical diagnostic purposes, as it is a fast and inexpensive method. However, this method has some severe limitations because it requires that the information of a reference strain close to the sample is pre-registered in the database, and thus cannot be applied for most environmental isolates. Moreover, this technology does not allow distinguishing between closely related strains or is refractory to mixtures of microorganisms.

Tandem mass spectrometry proteotyping is more efficient because the method is able to identify the taxonomic position in the tree of life whatever the sample. By deciphering the protein sequences extracted from the samples, accurate taxonomic data can be obtained. The method we have developed, "Phylopeptidomics" (1), allows to quantify the relative biomass of the identified organisms (2-3) and is highly sensitive (4). We propose here a methodology for ultra-fast identification of microorganisms, which is amenable to very high throughput, based on direct sample infusion and adapted computer processing. This pipeline will be illustrated with results obtained on numerous isolates from skin microbiota culturomics. The performance of this methodology will be detailed in terms of analysis time required per isolate identified. These results open new perspectives for the application of MS/MS proteotyping of microorganisms.

1. Pible O, Allain F, Jouffret V, Culotta K, Miotello G, Armengaud J (2020) Estimating relative biomasses of organisms in microbiota using "phylopeptidomics". *Microbiome* 8:30.
2. Hayoun K, Pible O, Petit P, Allain F, Jouffret V, Culotta K, et al. (2020) Proteotyping Environmental Microorganisms by Phylopeptidomics: Case Study Screening Water from a Radioactive Material Storage Pool. *Microorganisms* 8:1525.
3. Lozano C, Kielbasa M, Gaillard J-C, Miotello G, Pible O, Armengaud J (2022) Identification and Characterization of Marine Microorganisms by Tandem Mass Spectrometry Proteotyping. *Microorganisms* 10:719.
4. Mappa C, Alpha-Bazin B, Pible O, Armengaud J (2023) Evaluation of the limit of detection of bacteria by tandem mass spectrometry proteotyping and phylopeptidomics. *Microorganisms*, in press.

## Evaluation of the efficiency of the SPEED protein extraction protocol on plant samples

Thierry BALLIAU<sup>a</sup>, Shabnam HUSEYNOVA<sup>a</sup>, Marlène DAVANTURE<sup>a</sup>, Mélanie BLEIN-NICOLAS<sup>a</sup>

<sup>a</sup> Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, PAPPSO, 91190, Gif-sur-Yvette, France

In bottom-up proteomics, sample preparation is a critical step that strongly influences the quality of the results. In particular, protein extraction should enable efficient protein digestion for subsequent mass spectrometric analysis. Recently, Doellinger et al. [1] reported a detergent-free method called Sample Preparation by Easy Extraction and Digestion (SPEED), which consists of three mandatory steps: acidification, neutralization, and digestion. SPEED is an attractive method that holds the potential to dramatically simplify and standardize sample preparation while improving the depth of proteome coverage, especially for challenging samples. In addition, its authors report it as universal because it uses pure trifluoroacetic acid, which is supposed to extract protein from any sample by complete dissolution. However, this method has not been tested on plant samples. Yet, plants have specific compounds (e.g., polyphenols, chlorophyll, or major proteins) that can interfere with sample preparation. Here, we aim to evaluate if the SPEED method is suitable for preparing plant protein extracts. To this end, we sampled leaves, roots, and seeds from Arabidopsis Col-0 plants. We ground each tissue in liquid nitrogen and divided the resulting powders into nine aliquots that we subsequently used to extract proteins by three different protocols in three replicates: the SPEED method, the TCA-acetone-based, and the phenol-based extraction methods. We routinely use the latter two protocols to extract proteins from leaves and seeds, respectively. Mass spectrometry analysis for the whole experimental design is currently in progress. Preliminary results obtained on leaf samples indicate that the SPEED method provides good extraction yields and mass spectrometry signals.

[1] Doellinger et al. (2020) Sample Preparation by Easy Extraction and Digestion (SPEED) - A Universal, Rapid, and Detergent-free Protocol for Proteomics Based on Acid Extraction. Mol Cell Proteom.(1):209-222. doi: 10.1074/mcp.TIR119.001616.

## Multi-omics workflow for the identification of discriminant markers associated with *Trypanosoma cruzi* populations

Dr Yannick CHARRETIER<sup>a</sup>, Amy HESKETH<sup>a</sup>, Hayat HAGE<sup>a</sup>, Francisco OLMO<sup>b</sup>, Thomas RIBETTE<sup>a</sup>, Fanny ESCUDIE<sup>c</sup>, Martin TAYLOR<sup>b</sup>, Gilles COURTEMANCHE<sup>a</sup>, Derry MERCER<sup>a</sup>, Adrien SALIOU<sup>a</sup>, Eric CHATELAIN<sup>c</sup>, John KELLY<sup>b</sup>, Joséphine ABI GHANEM<sup>a</sup>

<sup>a</sup> BIOASTER, <sup>b</sup> LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE, <sup>c</sup> DNDi

### Introduction:

*Trypanosoma cruzi* (*T. cruzi*) is a protozoan parasite that causes Chagas disease. This zoonotic disease is transmitted via human by blood-sucking triatomine bugs. According to WHO, there are approximately 30 000 new cases of Chagas disease per annum and causes 13 000 deaths every year, mainly in Mexico, Central and South America. The life cycle of *T. cruzi* implies a differentiation of non-infective to infective forms within the triatomine gastrointestinal tract and an intracellular progression of *T. cruzi* infections in humans. After surviving the acute stage, patients progress in large part to an asymptomatic chronic infection. In a small proportion of cases, Chagas cardiomyopathy and a megasyndrome of the gastrointestinal tract can occur. Currently, knowledge limitations of the mechanistic processes associated with the persistence of *T. cruzi* compromise the development of improved treatments for Chagas disease.

To improve the understanding of Chagas disease, two populations (replicating epimastigotes and replicating amastigotes) have been isolated. An integrated proteome and transcriptome profiling method has been performed to identify discriminant markers/pathways associated with the stage of infection.

### Material & methods

Intracellular amastigotes can be isolated by repeated passage of infected host cells through a syringe needle. This is a rapid procedure but with safety issues. Fluorescently-tagged amastigotes can be further sorted and isolated using an Aria BD cell sorter. This technique achieves a higher purity, but with a lower yield and can be time-consuming. In this study, cells were inactivated with a nonionic detergent (Cell Disruption Buffer, PARIS kit - Thermo) and frozen until required. For the proteomic analysis, cells (100 µL) were precipitated with ice-cold acetone for two hours at -20 °C and resuspended directly into iST lysis buffer (Preomics). Protein lysates were digested according to Preomics instructions. Samples were adjusted to 0.5 µg/µL digested proteins and 1 µg of protein digested were loaded on a pre-column and separated onto 25cm\*75µm, 2µm EasySpray column (Thermo Electron, Waltham MA). Reverse phase separation was performed using an U3500 RS nano system coupled to an Exploris 480 in DIA mode (Thermo).

For the transcriptomics analysis, RNA was isolated according to the PARIS kit instructions (Thermo). Sequencing libraries were prepared with the Trio RNA-Seq™ library preparation kit (Tecan genomics) then sequenced on the Illumina NextSeq 2000 System.

### Results

A clear separation of amastigote and epimastigote populations was observed for both the transcriptome and the proteome profiles. An additional separation of the amastigote populations was observed for both techniques. This separation could be explained by the difference of infection stages between the replicates (optimization of the isolation procedure). A higher number of down regulated genes and proteins was observed in amastigotes compared to epimastigotes. An off/on pattern of protein expression was observed mainly in amastigotes (590 proteins were absent in amastigotes and present in epimastigotes).

### Conclusion

A robust method was developed to explore the transcriptomic and proteomic profiles of diverse parasite populations within a single sample. The observed distinction in transcriptomic and proteomic profiles of amastigotes and epimastigotes might be explained by their intracellular or extracellular states, respectively. Identification of differentially expressed genes and proteins between populations will further allow us to select discriminant markers/pathways. The ability to sort different amastigote sub-populations will be of interest to isolate "dormant/quiescent" parasites that may have a role in treatment failure. Low input and/or single-cell approaches are needed to further investigate these potentially relevant "dormant/quiescent" populations.

## Composition of the Rainbow trout testicular extracellular matrix using DDA-PASEF and diaPASEF.

Blandine GUÉVEL<sup>a</sup>, Nathalie CHENAIS<sup>b</sup>, Régis LAVIGNE<sup>a</sup>, Aurélien BRIONNE<sup>b</sup>, Emmanuelle COM<sup>a</sup>, Charles PINEAU<sup>a</sup>, Jean-Jacques LAREYRE<sup>b</sup>

<sup>a</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35042 Rennes cedex, France, <sup>b</sup> Laboratory of Fish Physiology and Genomics, INRAE, 35042 Rennes cedex, France

### INTRODUCTION

Testicular organoids are promising 3D cell culture models to study direct and indirect cell to cell interactions between germ stem cells (GSC) and somatic cells, as well as mechanisms leading to male infertility (disease, ecotoxicology). Testicular organoids could also allow the amplification of GSC for the conservation of genetic resources and the regeneration of the cohorts of interest. The production of organoids based on the use of heterologous (natural and synthetic) extracellular matrix (ECM) has been commonly reported in mammals, but endogenous scaffold could be more suitable. The present study used Mass spectrometry-based proteomics in order to investigate the composition of the Rainbow trout testicular extracellular matrix (tECM).

### METHOD

Testes were collected from immature Rainbow trout and then decellularized using SDS (Sodium dodecyl sulfate). The proteins were extracted from decellularized or non-decellularized testicular fragments. High-throughput mass spectrometry-based proteomics were carried out to identify and compare the proteins on these two groups of samples using DDA-PASEF or direct diaPASEF approaches.

In a second time, a spectral library was generated from 3 pools of 6 individual trout testicular fragments previously and separated by a 1D gel electrophoresis. The separated protein lanes (46 different fractions) were acquired with DDA-PASEF technology. All DDA- and diaPASEF analysis were performed on a nanoElute liquid chromatography system, and a timsTOF Pro mass spectrometer (Bruker Daltonics). Instrument's methods were adapted to perform data-independent isolation of multiple precursor windows within a single 100ms timsIMS separation. 18 samples were then successfully analyzed using a library-based workflow with the Spectronaut 16 software (Biognosys) compared to a DIA direct approach.

### RESULTS

Proteomics analyses unraveled the protein complexity of non-decellularized and decellularized testes including growth factors and components of the basal lamina, respectively. In this study, we could identify with the direct diaPASEF approach 5632 and 2784 protein groups for the non-decellularized and decellularized testes respectively, including 4212 and 1555 unique gene ID. The DDA-PASEF library resulting from the analysis of 46 peptide fractions and the analysis of 18 additional samples will allow us to deepen and confirm the results obtained with the DIA direct approach.

### CONCLUSION

Our data provide new perspectives for improving the scaffolds and media used for the production testicular organoids.

*The project received funding from the EU research project Aquaexcel 3.0 (2020-2025) and the FAPESP program (20/03569-8).*



## Study of the mode of action of miPEPs in *Arabidopsis thaliana*

Clémence GUILLON<sup>a</sup>, Joris MACQUET<sup>b</sup>, Serge PLAZA<sup>a</sup>, Jean-Philippe COMBIER<sup>a</sup>, Patrice THULEAU<sup>a</sup>, Bertrand FABRE<sup>a</sup>

<sup>a</sup> Laboratoire de Recherche en Sciences Végétales, Université Paul-Sabatier-CNRS-Toulouse INP, <sup>b</sup> Laboratoire des interactions plantes - microbes - environnement, INRAe-CNRS

MicroRNAs (miRNAs) are negative regulators of gene expression via an interfering RNA mechanism. Through this mechanism miRNAs regulate a wide range of processes in the life of a plant, such as vegetative growth, development and response to various biotic and abiotic stresses. The biosynthesis of miRNAs is well documented in the literature: a miR gene is transcribed into a precursor called pri-miRNA, which in parallel to its transcription is cleaved by the multi-protein complex DICER into a microRNA.

In 2015 it was shown that short open reading frames (ORFs) are present on pri-miRNAs and can be translated into a peptide called miPEPs (miRNA-encoded peptides). MiPEPs act as positive regulators of the expression of their own pri-miRNA.. Thus, plants can be sprayed or watered with a given miPEP leading to an increase in the expression of its corresponding pri-miRNA and miRNA. *In fine*, this results in an increased inhibition of the target mRNA, making miPEPs an interesting molecular reservoir for crop biocontrol, as they have the potential to regulate as many processes as miRNAs.

However, the mode of action of miPEPs at the molecular level still remains unclear. miPEPs do not appear to have any defined domains nor conserved sequences and vary in length and amino acid composition. Recent data suggest that miPEPs have the ability to interact with their own ORFs on pri-miRNAs, thus providing a first element of response to the specificity of action of miPEPs. However, the precise molecular mechanism by which miPEPs act as positive regulator of pri-miRNA expression remains unresolved to date.

In this context, we hypothesized that the mode of action of miPEPs is mediated through miPEP-protein interactions. Among all the proteins involved in pri-miRNA transcription, processing and maturation, based on preliminary experiments, we narrowed them down to three potential interacting candidates: HYL1, NOT2a and DDL. Thus, while exogenous application of miPEP165a to wild type *Arabidopsis thaliana* plants resulted in an increase of root growth, the same treatment applied to *hyl1*, *not2a* and *ddl* mutants showed no effect of the miPEP. These results reveal that the three proteins are important for the activity of miPEPs. In addition, using an expression system in insect cells coupled with co-immunoprecipitation experiments, we found that DDL and HYL1 interacted with miPEP165a but also with its scrambled version. These results suggest that DDL and HYL might interact with short and poorly structured amino acid sequences. The expression of pri-miR165a in response to miPEP165a treatment in the *hyl1*, *not2a* and *ddl* mutants is currently being investigated.

In parallel, we are planning to characterize the interactome of miPEPs in *A. thaliana*, with the aim of identifying all protein partners in an unbiased manner. To this end, we will develop a proximity labelling approach using a Turbo-ID biotin ligase-fused miPEP expressed in *A. thaliana*. The biotinylated partner proteins will then be purified and identified by mass spectrometry. This approach will allow us to detect weak and transient interactions between a miPEP and its associated proteins.

## Identification of partner proteins of chlordecone by drug affinity responsive target stability (DARTS)

Dr Mélanie LAGARRIGUE<sup>a</sup>, Blandine GUÉVEL<sup>b</sup>, Iwan JAN<sup>c</sup>, Simon MONNERAIS<sup>c</sup>, Tanguy FERLIER<sup>c</sup>, Régis LAVIGNE<sup>b</sup>, Charles PINEAU<sup>b</sup>

<sup>a</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35042, <sup>b</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35042 Rennes cedex, France; Univ Rennes, CNRS, Inserm, Biosit UAR 3480 US\_S 018, Protim core facility, F-35000 Rennes, France, <sup>c</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35042 Rennes cedex, France

We are continuously exposed to a multitude of exogenous molecules (pesticides, plasticizers, drugs, industrial pollutants, cosmetics, food additives, etc). Some of them present endocrine disruptor (ED) properties. The assessment of the impact of this exposure on health and the investigation of the mechanisms of action of EDs have become major issues but remain extremely difficult to address.

Several tests validated by the OECD make it possible to study the agonistic or antagonistic properties of EDs on hormone receptors and their impact on hormone synthesis or to reveal deleterious effects on sexual development, reproduction and growth. Nevertheless, this is not sufficient to elucidate the mechanisms of action leading to these outcomes, as most EDs probably disrupt a wide range of biological processes beyond hormone receptors. In this context, the characterization of interactions between EDs and molecular targets such as partner proteins constitutes essential information to better understand their mechanisms of action. Conventional biochemical approaches used for the identification of proteins interacting with a small molecule are based on its immobilization. This requires the chemical modification of the small molecule, which can be difficult with a high risk of losing its activity. The objective of this study is to evaluate the interest of recent techniques avoiding this chemical modification for the simple and rapid identification of proteins interacting with EDs.

Several approaches to identify proteins interacting with a non-immobilized small molecule are based on the principle that the protein is stabilized when bound to its ligand. This stabilization can be revealed by a greater resistance to a short non-specific proteolysis. This property is the basis of the drug affinity responsive target stability (DARTS) technique chosen for this study. Chlordecone was selected as a model molecule because of its well-characterized ED activity. A complex protein extract was incubated with chlordecone and then subjected to a short non-specific proteolysis by pronase under native conditions. The sample was subsequently digested with trypsin/LysC and analyzed by LC-MS/MS. A label-free differential analysis was performed to compare protein expression levels between the chlordecone-treated extract and the control extract treated with solvent only. Proteins whose expression levels increase in the presence of chlordecone are considered as potential partners.

Extraction and DARTS conditions were optimized on MCF7 cell pellets. Best results were obtained with a pronase:proteins ratio of 1:100 and a chlordecone concentration of 10  $\mu$ M. Label-free differential analysis was then performed on six biological replicates. Six potential protein partners were identified. The protocol is now transposed on rat prostate tissue to better reflect the physiological context. This protocol before being tested on human prostate biopsies.

In conclusion, the DARTS protocol is easily accessible for any laboratory equipped for proteomics and avoids the biases induced by the immobilization of the small molecule. However, the sources of variation in the protocol must be as controlled as possible. The best potential partners of chlordecone must be validated by an orthogonal technique such as microcalorimetry to conclude on the applicability of this protocol to the discovery of proteins interacting with EDs.

## Development of a proteomic pipeline addressing routing meat quality assessment within food industry

Miguel SENTANDREU<sup>a</sup>, Pablo FERNANDEZ<sup>a</sup>, Enrique SENTANDREU<sup>a</sup>

<sup>a</sup> Instituto de Agroquímica y Tecnología de Alimentos (IATA-CSIC), C/Agustín Escardino 7, 46980 Paterna, Valencia (Spain)

Enrique Sentandreu, Pablo Fernández and Miguel A. Sentandreu\*

*Instituto de Agroquímica y Tecnología de Alimentos (IATA-CSIC), C/Agustín Escardino 7, 46980 Paterna, Valencia (Spain)*

\*To whom correspondence should be addressed: [ciesen@iata.csic.es](mailto:ciesen@iata.csic.es) (Miguel A. Sentandreu).

There is an urgent necessity for developing reliable, sustainable and affordable analytical methods for meat quality assessment that could be easily implemented by food industry. High-output strategies based on liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) have clearly demonstrated their efficiency to study in detail quality of meat-based foodstuffs mainly through elucidation of key peptide biomarkers. However, industrial implementation of such effective MS-based solutions is normally constrained by their budget-dependent requirements, limiting R&D activities about meat quality assessment outside basic research actions. Moreover, during the last years high-resolution MS (HRMS) analysis emphasized the technological gap between basic research and meat industry activities, discouraging implementation of HRMS technology for quality evaluation of traded foodstuffs.

As an affordable alternative to untargeted HRMS approaches, targeted analyses performed by low-resolution (LRMS) exploratory devices such as three-dimensional (3D) ion-traps can be of interest for routine assessment of meat quality. Thus, in this research there was combined the untargeted and targeted capacities of these conventional analytical platforms for biomarker hunting capable to discriminate normal from defective bovine meats through a flexible gel-free proteomic pipeline. There were initially proposed different potential protein biomarkers by comparing proteomic profiles of direct extracts elucidated by data-dependent LC-MS/MS analyses (dd-MS<sup>2</sup>) carried out in both previous studies and a complementary research. There were subsequently considered only those proteotypic peptides derived from proteins exhibiting larger differences (mainly absence/presence) among meat groups assayed. Quantitative analysis of such peptides was carried out by targeted Selected Reaction Monitoring (SRM) experiments of direct protein extracts from six (Normal/Defective) biological replicates. Among 18 peptides assayed, results confirmed how L-lactate dehydrogenase A chain and Phosphoglucosmutase 1 were clear protein biomarkers characterizing normal meat. In contrast, Glycogen debrancher, ATP-dependent 6-phosphofructokinase and chaperones Heat shock proteins beta 1 and beta 6 unambiguously typified defective meats. It must be emphasized how targeted SRM analysis of these six selected peptide biomarkers can be easily performed by routine analytical protocols for early prediction of meat quality at industry level. As a result, significant economic losses may be prevented through the early detection of defective meats prior to reaching the market.

Flexibility, affordability and easiness of proteomic pipeline proposed allowed the monitoring of peptide biomarkers suggested by both literature and customized analyses featured by conventional LC-MS analytical platforms merging exploratory and targeted approaches. Consequently, this dynamic and sustainable strategy can be readily implemented by meat industry for quality assurance of resultant products.

## Chemically reversible Nanobody coupling for a more specific elution

Yéléna MARTIN<sup>a</sup>, Rachel BOURDON-ALONZEAU<sup>a</sup>, Célia CUCULIÈRE<sup>a</sup>, Pierre-François MERY<sup>a</sup>, Jean CHEMIN<sup>a</sup>, Emmanuel BOURINET<sup>a</sup>, Franck VANDERMOERE<sup>a</sup>

<sup>a</sup> INSTITUT DE GÉNOMIQUE FONCTIONNELLE, Université de Montpellier, INSERM, CNRS

Small single-chain llama antibodies, and nanobodies (LaG), are interesting molecules for drug discovery, therapeutics, and research. Their very small size (15 kDa) gives them access to epitopes inaccessible to conventional antibodies. Made of only one protein chain, they are easy to produce at large scale in bacterial expression systems and can be engineered with classical molecular biology tools. Furthermore, nanobodies are extremely stable and their affinity for antigens is in the nanomolar range.

A recent publication on anti-GFP nanobodies brings out that the dimerization of nanobodies can reach even better affinity. An incredible 36 pM Kd was described for two fused nanobodies LaG-16-LaG-2 separated by a peptide linker composed of three repetitions of GGGGS).

The goal of this project is to make use of this unprecedented nanobody affinity to purify from mice's spine and brain the interactome of a GFP-tagged Cav3.2 channel, a very low abundant calcium channel whose role in chronic pain has been discovered by the team. We produced the biparatopic nanobody (LaG-16-LaG-2) tool in the periplasm of a bacterial system, purified it by affinity and steric exclusion chromatography and covalently labeled it with biotin with a chemical linker sensitive to a reduction (thiol).

The affinity of streptavidin beads for the biotinylated nanobody (Kd 10 fM) and the affinity of the nanobody for the GFP tag being too high for classical SDS elution, we had to take advantage of the thiol group of the linker to release the nanobody by reduction. This method should also not elute proteins aspecifically bound to beads or plastic. After optimization, more than 70% of the nanobody was recovered in the elution, more than double the amount obtained by the classical SDS elution.

We are now testing our optimized protocol on brain samples from GFP-Cav3.2 knockin vs Wild type mice for APMS analysis. This protocol might be very useful to study the interactome of any low-abundant protein.

Keywords: Nanobody, affinity purification, biotinylation, chemical elution, mass spectrometry



## Going against the grain: development of denaturing mass photometry for fine-tuning of protein-protein cross-linking reactions

Hugo GIZARDIN-FREDON<sup>a</sup>, Oscar HERNANDEZ-ALBA<sup>a</sup>, Sarah CIANFERANI<sup>a</sup>

<sup>a</sup> 1 Laboratoire de Spectrométrie de Masse BioOrganique, IPHC UMR 7178, Université de Strasbourg, CNRS, 67000 Strasbourg, France ; 2 Infrastructure Nationale de Protéomique ProFI - FR2048, 67087 Strasbourg, France

Structural MS is a generic term that encompasses a series of MS-based strategies adapted for the characterization of non-covalent assemblies. Among them, cross-linking followed by mass spectrometry (XL-MS) is a covalent labelling proteomics-based technique that has drastically progressed this past decade from *in vitro* to even *in situ/in cellulo* applications<sup>1</sup>. XL-MS approaches provide snapshots of non-covalent protein-protein interactions (PPI) through determination of spatial proximities and distance constraints. The first step of chemical XL reaction itself is of utmost importance in the XL-MS workflow. Indeed, before LC-MS/MS analysis, optimal XL conditions have to be determined (optimal reagent and XL conditions) without generating extensive non-specific XL aggregates. To date, denaturing 1D SDS-PAGE is recommended to monitor and optimize XL conditions<sup>2</sup>, allowing visual observation of high mass bands on the gel corresponding to cross-linked species with concomitant vanishing of bands corresponding to individual free protein partners. Even if denaturing SDS-PAGE analysis is easy, widely available and robust, it has some limitations: i) low mass-resolution, ii) time consuming (24 hour time-scale from gel casting to protein staining), iii) not suited for high-masses or highly heterogeneous complexes which do not enter the gel.

To circumvent these limitations, we aimed at evaluating mass photometry (MP), a recently developed “single-molecule” biophysical technique<sup>3</sup> that gains interest in structural biology, for XL reaction optimization. MP is a versatile, fast and low sample-consuming technique that complements native MS for protein mass distributions measurements in native conditions. Despite its routine use to study native protein-protein interactions or multiprotein complexes oligomerization<sup>4</sup>, MP workflows are not adapted for the characterization of covalent assemblies.

We report here on the development of a suitable and efficient MP methodology in denaturing conditions (called dMP) for the straightforward characterization of covalent assemblies, with a special emphasis on its benefits for rapid XL reaction optimization. The developed workflow consists of a robust and efficient 2-step protocol that ensures 95% of irreversible denaturation within 5min in urea, just before MP measurement. After evaluating its efficiency on reference multimeric proteins (BSA, ADH, GLDH and 20S proteasome), we benchmarked dMP against the reference 1D SDS-PAGE gel migration to fine-tune protein XL reactions (impact of XL reagents, concentrations, etc.). dMP allowed a rapid screening of 20 different XL conditions including different commonly used reagents (PhoX, DSAU, DSSO, DSBU) within ~1.5 hour (with measurements in triplicates), highlighting that size and flexibility of the spacer arm are of utmost importance on XL efficiency. Finally, we successfully applied the dMP-based XL optimization within a complete XL-MS workflow that aimed to study the PPI of R2SP, a ~540 kDa complex involved in major cell functions as DNA repair, remodeling of chromatin or transcriptional regulation<sup>5</sup>.

By providing direct visualization along with relative quantification of all sub-complexes and aggregates, dMP can be uniquely positioned as the go-to method for more precise and accurate, mass-based optimization of best XL conditions. Consequently, optimal XL conditions (cross-linker, reaction time, temperature, cross-linker excess, pH value, etc.) can be rationally screened in a straightforward manner and with more accurate readouts (visualization of sub-complexes vs over-XL aggregates, relative quantification of the species, etc.). We envision dMP as a more precise and accurate alternative technique than 1D SDS-PAGE analysis for XL reaction optimization, which might be fully automatized in a near future.

(1) Piersimoni, L.; and al. *Chem. Rev.* 2021

(2) Iacobucci, C.; and al. *Anal. Chem.* 2019

(3) Asor, R.; and al. *Curr. Opin. Chem. Biol.* 2022

(4) Soltermann, F.; and al. *Angew. Chem. Int. Ed.* 2020

(5) Maurizy, C.; and al. *Nat. Commun.* 2018



## Enhancing crosslinking reactions in photosynthetic membrane systems for proteomics

Nicolas FRANCES<sup>a</sup>, Anne-Marie HESSE<sup>a</sup>, Sylvie KIEFFER-JAQUINOD<sup>a</sup>, Gilles CURIEN<sup>b</sup>, Myriam FERRO<sup>a</sup>, Giovanni FINAZZI<sup>b</sup>, Pascal ALBANESE<sup>b</sup>

<sup>a</sup> BGE-EDyP, UA13 / FR2048 UGA-CEA-Inserm-CNRS, <sup>b</sup> PCV-LPM, UMR5168/UMR1417 UGA-CNRS-CEA-INRAe

### Context:

Membrane proteins play a vital role in maintaining cell homeostasis, and some are involved in crucial reactions in the energy transducing membranes of chloroplasts and mitochondria. The required steps of solubilization and purification make structural studies challenging, and crosslinking (XL) mass spectrometry (MS) has emerged as an alternative tool for studying their structure. PhoX (1) is an innovative reagent used to enrich XL-peptides with a phosphonic moiety, however, the yield is generally low due to negative repulsion with the membrane/protein surface. In order to address this problem, we propose that Trimethylphenylammonium chloride (TMPAC) could accumulate at the membrane interface, mask negative charges, and serve as an enhancer to bring PhoX closer to the membrane surface, thereby enabling it to react with the present lysine residues.

Experiments were conducted using Spinach (*Spinacia oleracea*) thylakoid membranes. The effects of TMPAC on the integrity of the membrane and the generation of crosslinks were evaluated using photosystem II (PSII) structure and biokinetic assays. The compound was shown to be a competent enhancer, increasing the number of crosslinks identified by 81%, without modifying the protocol for XL-peptides enrichment. These findings demonstrate the potential of TMPAC in combination with PhoX as a powerful tool to elucidate the conformation of membrane proteins.

### Main results:

Concentrations lower than 100 mM of TMPAC do not affect the physiological state of the thylakoid membrane (measured as enzymatic efficiency of PSII for electron production), at any intensity light, respect the control condition for the thylakoids solution without TMPAC (equivalent to 100% of photosynthetic performance)

The physiological state of the membranes is not altered, therefore, there are no induced artifacts and false positive crosslinks.

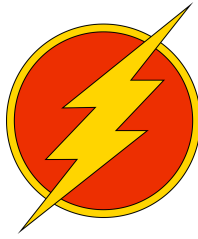
In other assays, we show that no synergistic effects between PhoX 1 mM and TMPAC solutions at the lower concentrations. Thus, all the physiological changes are inherent to the crosslinking formation.

### Conclusions:

Our research revealed that the combination of TMPAC and PhoX enhances the yield of identified crosslinks by nearly one-fold, enabling greater accessibility of the crosslinking reagent without altering the physiological state of the system. Our 3D modeling analysis suggests that TMPAC induces novel crosslinks in protein regions with negative surface charges. Moreover, it was shown to have critical importance considering loose terminal regions when modeling complete 3D structures. Overall, TMPAC is a promising tool for advancing our understanding of membrane-bound proteins. It is commercially available, cost-effective, and highly efficient, facilitating crosslinking reactions in native membranes without generating artifacts and disruptions. Also, it shows a great display in combination with low crosslinking reagent concentrations. Nonetheless, further experimentation is required to explore its full potential for in vivo studies and optimal use in diverse experimental setups.

- (1) Steigenberger, B. et al. 2019. DOI: 10.1021/acscentsci.9b00416.





## Short modified peptides or dimeric peptide isolated from animal venoms: Are LC-MS-MS, Edman sequencing enough to determine their sequences?

Yasmine BOUGHANMI<sup>a</sup>, Pascal MANSUELLE<sup>b</sup>, Regine LEBRUN<sup>b</sup>, Patrick FOURQUET<sup>c</sup>, Marielle BALZANO<sup>d</sup>, Caroline BERENQUER<sup>e</sup>, Maxime MOULARD<sup>d</sup>, Harold DE POMYERS<sup>f</sup>, Didier GIGMES<sup>g</sup>, L'houcine OUAFIK<sup>e</sup>, Kamel MABROUK<sup>g</sup>

<sup>a</sup> Aix-Marseille Univ CNRS ICR UMR 7273, Équipe CROPS, <sup>b</sup> Aix Marseille Univ, IMM Institut de Microbiologie de la Méditerranée (Marseille Protéomique), FR 3479, Marseille, France, <sup>c</sup> Aix-Marseille Univ, INSERM, CNRS, Institut Paoli-Calmettes, CRCM, 13009 Marseille Protéomique, Marseille, France, <sup>d</sup> BioCytex, Marseille, 13010, France, <sup>e</sup> Aix Marseille Univ, CNRS, Institut de Neurophysiopathologie (INP), Marseille, France, <sup>f</sup> LATOXAN SAS, 845 Avenue Pierre Brossolette, 26800 Portes-lès-Valence, France, <sup>g</sup> Aix-Marseille Univ CNRS ICR UMR 7273, Équipe CROPS, 13397 Marseille, France

The research for new therapeutic molecules is a major public health issue, more particularly in the treatment of certain diseases. In our group, venoms have been screened in several areas in order to identify new painkillers and mainly anti-tumors or antibacterial agents.

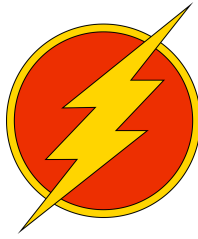
So far, the success of cancer treatments has not been satisfactory. Indeed, synthetic chemicals used in chemotherapy and radiotherapy have low solubility and sometimes extremely high cytotoxicity that can cause significant collateral damage to normal tissues. In addition, they can induce chemoresistance or radioresistance after a period of treatment. It is therefore essential to develop new molecules capable of blocking the division of cancer cells and having less toxic effects on healthy tissues. Many anti-cancer peptides (ACPs) have the ability to specifically and selectively target cancer cells.

Moreover, natural resistance of bacteria consecutive to overuse of antibiotics in human and animal therapies as well as in various industrial fields have led to the selection of pathogenic multi- drug resistant bacteria. Over the last few decades, resistance to conventional antibiotics has limited treatment options, resulting in a significant increase in mortality and morbidity in hospitals. Venoms are known to be a rich source of unique and innovative pharmacologically active substances, including ACPs or AMPs.

In our study, the antibacterial or antiproliferative properties of 200 venoms of snakes, scorpions, amphibians and insects were studied. The identification of these AMPs or ACPs was carried out by: i) tests for the inhibition activity ii) HPLC for the purification of compounds from fractions found active and iii) mass spectrometry conjugated with amino acid sequencing for sequences/structures determination.

The difficulties encountered when determining 2 sequences of bioactive peptides, one short and is modified and the other long and is dimeric will be discussed.





## Immunoglobulins expression by non immune cells

Dr Alice CAPUZ<sup>a</sup>, Soulaimane ABLOUARD<sup>a</sup>, Marco TREROTOLA<sup>a</sup>, Dasa CIZKOVA<sup>b</sup>, Etienne COUYAUD<sup>a</sup>, Firas KOBEISSY<sup>c</sup>, Franck RODET<sup>a</sup>, Michel SALZET<sup>a</sup>

<sup>a</sup> Univ. Lille, Inserm, U-1192 - Laboratoire Protéomique, Réponse Inflammatoire et Spectrométrie de Masse-PRISM, F-59000 Lille, France, <sup>b</sup> Institute of Neuroimmunology, Slovak Academy of Sciences, Dúbravská cesta 9, 845 10 Bratislava, Slovakia, <sup>c</sup> Department of Biochemistry and Molecular Genetics, Faculty of Medicine, American University of Beirut, Lebanon.

The CNS includes the brain and spinal cord, and is composed of neurons, glial cells, and immune cells. Neurons carry nerve information while glial cells maintain CNS stability. Astrocytes are the most abundant glial cells and regulate nutrient supply, CNS homeostasis, nerve information transmission, and synaptic formation. The central nervous system has long been considered to be excluded from the peripheral immune system. However, contrary to the immunology dogma stipulating that only B lymphocytes are capable of producing antibodies, we have shown that astrocytes and neurons could be a new cellular source in the central nervous system. An antibody of two identical heavy and two light chains with a paratope that binds to specific antigens. This diversity is achieved through V(D)J recombination, where the V, J, and D segments are rearranged. However, the edification of antibodies issued of these neural cells is different from the ones found in B cells and resembles to aberrant immunoglobulins. Some are pseudogenes and belongs what we named, the Ghost proteome or alternative proteins. Alternative proteins originate from the 5'- and 3'- UTR mRNA regions, frameshifts of mRNA ORFs or from non-coding RNAs. Long-time considered as non-coding, recent in-silico translation prediction methods enriched the protein databases allowing the identification of new target structures that have not been identified previously.

Then, we focused our interest on one a kappa light chain variable region. We demonstrated that this protein named "Heimdall" was expressed by astrocytes and was secreted under inflammatory conditions with lipopolysaccharides (LPS). Immunoprecipitation experiments showed that the Heimdall interactome contained proteins related to astrocyte fate keepers such as "NOTCH1, EPHA3, IPO13" as well as membrane receptor protein including "CHRNA9; TGFBR, EPHB6, and TRAM". However, when Heimdall protein was neutralized utilizing a specific antibody or its gene knocked out by CRISPR-Cas9, sprouting elongations were observed in the corresponding astrocytes. Interestingly, depolarization assays and intracellular calcium measurements confirmed a phenotype switch from astrocytes to neuron-like phenotype. Proteomic analyses performed under LPS stimulation, revealed the expression of neuronal factors, stem cell proteins, and neurogenesis of astrocyte convertor factors such as EPHA4, NOTCH2, SLIT3, SEMA3F, suggesting a role of Heimdall to regulate astrocytic fate.

Besides Heimdall, we also identified an aberrant form of the IgG2b heavy chain constant region. In fact, transcriptomic analysis showed that the messenger encoding this IgG2b constant part didn't have a variable part coding sequence at its 5' end, but a Kozak sequence located directly upstream of the exon coding the CH1 domain. Moreover, invalidation of IgG2B by CRISPR-Cas9 in astrocytes show a higher abundance of proteins characteristic of hematopoietic cells (FRG1, PIR), neural stem cells (YAP1, TEAD3, SMAD2/9, MCC, KDM2A and DDHD2) and regulation of neuritogenesis (PLXNA1, ROCK1, MAGI3, TBL1). These results suggest that this inhibition of this heavy chain led to dedifferentiation of astrocytes into neural progenitor cells via activation of the BMP-YAP1-TEAD3 pathway (Capuz et al., 2023). Conversely, overexpression of IgG2B in astrocytes induced the CRT1-CREB-BDNF signaling pathway involved in gliogenesis and instead directs astrocytes toward a pro-inflammatory profile. Furthermore, proximal labeling (Bio-ID) experiments revealed that IgG2B is N-glycosylated by the OST complex, addressed to the membranes of ATPase complex-containing vesicles, and partially behaves as SLC3A2 (CD98hc) through its association with SLC7A5 (LAT1) and its ability to recruit ILKAP that potentially regulates downstream signaling pathways (Capuz et al., 2023).

Taken together, these results support that aberrant IgG2B and Heimdall chains are involved as a gatekeeper of astrocyte identity and that its inhibition leads to the conversion to a neural stem cell phenotype.

## A fast and efficient method for peptide fractionation at basic pH in proteomic studies

Florine HALLEZ<sup>a</sup>, Sami BAYOUDH<sup>a</sup>, Michel AROTÇARENA<sup>a</sup>, Cerina CHHUON<sup>b</sup>, Chiara GUERRERA<sup>b</sup>, Kaynoush NARAGHI<sup>a</sup>

<sup>a</sup> Affinisep, <sup>b</sup> INSERM - US24 SFR Necker Proteome, Paris

Peptide fractionation plays an important role in bottom-up proteomic approaches, since it simplifies the peptide pool for easier analysis by mass spectrometry, to allow for more accurate interpretation. Indeed, full proteome characterization is often needed for complex biological matrices and, with growing interest in post-translational modifications, an extended protein sequence coverage is required.

However, efficient peptide fractionation is very challenging, especially in the case of complex samples containing peptides with a wide range of properties (charge, polarity and size). The objective of this study was therefore to develop a simplified procedure for the efficient and fast fractionation of peptides at basic pH. This fractionation method is orthogonal to reversed-phase liquid chromatography, thus contributing to the simplification of the peptide separation and analysis.

A new reversed-phase sorbent, based on small sorbent particles tightly embedded in a soft, uniform and mechanically stable monolithic membrane packed in SPE StageTips, was used for the fractionation of peptides resulting from the enzymatic proteolysis of proteins contained in HEK293 cell lysate, and the results were compared to the ones obtained with a reference commercial fractionation kit.

50µg of peptides were used for the fractionation tests on the commercial columns and on the SPE Tips. Eight fractions, with percentages of acetonitrile ranging from 5 to 50%, were performed on the commercial column, following the protocol described by the supplier. The same fractionation gradient was used on the SPE Tips, as a starting point, before being optimized. Each fraction obtained with both fractionation methods was evaporated to dryness and then resuspended in an appropriate solvent for nanoLC-MS/MS analysis.

If the total number of proteins identified and the percentage of peptides eluting in only one fraction (50%) were similar for both sorbents, with a good distribution of peptides over the eight fractions, it appeared that the fractionation on the newly developed sorbent presented several advantages compared to the reference kit. Indeed, if the commercial columns have to be stored at 4 °C in a storage buffer, the new sorbent has no storage constraints since it can be stored dry at room temperature over several years, without observing degradation. Moreover, due to the SPE StageTips format, the time required for the evaporation of each fraction is almost halved compared to the commercial columns.

Thus, the new sorbent appears as a promising solution for the fractionation of complex samples such as plasma, or the generation of spectral libraries, since it leads to an increase of more than 25% in the number of proteins identified, compared to unfractionated samples. Finally, this new sorbent offers a real flexibility of format and capacity, since it is also available as spin columns, for high amounts of peptides, or as 96 SPE well plates, for high throughput experiments, and it can be used to fractionate peptides samples, from few ng to several mg.

## Investigation of the human lysosomal proteome by targeted proteomics

Dr Sabine JOURDAIN<sup>a</sup>, Stephanie KASPAR-SCHOENEFELD<sup>a</sup>, Dhriti ARORA<sup>a</sup>, Andreas SCHMIDT<sup>a</sup>, Pierre-Olivier SCHMIT<sup>a</sup>, Dominic WINTER<sup>b</sup>

<sup>a</sup> BRUKER, <sup>b</sup> Institute for Biochemistry and Molecular Biology, Medical Faculty, Rheinische Friedrich-Wilhelms-University of Bonn, Bonn, Germany

### Introduction

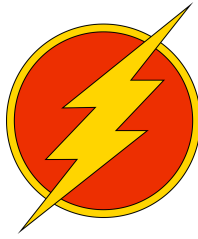
*Lysosomes, the main lytic organelles of almost all mammalian cells, fulfil crucial roles for the degradation and recycling of intra- and extracellular macromolecules, and play essential roles in metabolic signaling. These functions are facilitated by >340 lysosome-related proteins, whose loss of function can result in a variety of disorders. Proteomic analysis of such diseases is complicated by the low abundance of lysosomal proteins. An ideal method to address this is data-independent acquisition (DIA), which enables the reproducible and accurate protein identification and quantification across large sample cohorts. dia-PASEF merges the benefits of DIA with the advantages of ion mobility in proteomics experiments. Here, we use dia-PASEF to develop a large-scale targeted quantitation assay for the lysosomal proteome.*

### Methods

Human cell lines were grown in adherent cultures and lysosomes enriched using superparamagnetic iron oxide nanoparticles, yielding whole cell lysates and lysosome-enriched fractions. Proteins were precipitated by acetone, resuspended using RapiGest, and digested overnight using trypsin. After desalting of peptides, samples were analyzed by coupling a nanoElute online to a timsTOF Pro 2 (both Bruker Daltonics). Peptides were separated on a 25cm reversed-phase C18 column using a 60-minute gradient. For dia-PASEF a method consisting of sixteen TIMS ramps with 2 mass ranges in each 100ms scan was applied. Data were processed with Spectronaut v17 (Biognosys) applying directDIA using either a targeted lysosomal protein library (297 proteins) or a full human proteome fasta file for untargeted discovery proteomics.

### Preliminary Data

Using dia-PASEF we identified on average 97,528 peptides and 8,623 protein groups from whole cell lysate and 70,453 peptides and 7,912 protein groups for lysosome-enriched samples, using a 60-minute linear gradient. We observed excellent reproducibility using the dia-PASEF acquisition method, with median CV values of 5.7% and 8.8% for the whole cell lysate and lysosome-enriched samples, respectively. A common problem with respect to the analysis of lysosomal proteins is the lack of reproducible detection in complex samples. In theory, dia-PASEF results in the fragmentation of all possible precursors, thereby also including low-abundant lysosomal proteins, irrespective of the highly complex background. To investigate the coverage of lysosomal proteins in both sample types, we used targeted data processing by applying a library consisting of 297 lysosomal proteins, which were shown in previous experiments to be reproducibly detected in these samples. On average, we identified 3013 peptides and 269 protein groups from whole cell lysate and 3078 peptides and 276 protein groups from lysosome-enriched samples. Even though the total number of identified proteins between enriched samples and whole cell lysate sample doesn't differ significantly, the overall signal intensity for peptides originating from lysosomal proteins was ~30% higher in the enriched sample set. To mimic regulation of the lysosomal proteome between different states, enabling an assessment of quantification performance, a fraction of the lysosome-enriched sample was spiked-into the whole cell lysate simulating constitutive upregulation of the lysosomal proteome. In total, 165 lysosomal proteins showed significantly higher abundance (fold change  $\geq 1.2$ , p-value  $\leq 0.05$ ). This shows that dia-PASEF is well-suited for analyzing the lysosomal proteome, providing both good coverage and quantitative reproducibility of the targeted lysosomal proteins in enriched and whole cell lysate sample. In a next step we will further improve sample throughput by using shorter gradients and improved dia-PASEF methods.



## Comparison of *Bacillus thuringiensis* virulence potential on human Caco-2 intestinal cells using multi-omics investigation

Sarah ALILAT<sup>a</sup>, Rachele LANCELEUR<sup>b</sup>, Mathilde BONIS<sup>c</sup>, Olivier FIRMESSE<sup>c</sup>, Gaëlle VIAL<sup>d</sup>, Xavier SARDA<sup>d</sup>, Valérie FESSARD<sup>b</sup>, Sophie LIUU<sup>c</sup>, Thibaut LÉGER<sup>b</sup>

<sup>a</sup> Toxicology of Contaminants Unit, National Agency for Food, Environmental and Occupational Health & Safety, <sup>b</sup> Toxicology of Contaminants Unit, National Agency for Food, Environmental and Occupational Health & Safety (ANSES, Fougères), <sup>c</sup> Staphylococcus, Bacillus and Clostridium Unit, Food Safety Laboratory, National Agency for Food, Environmental and Occupational Health & Safety (ANSES, Maisons-Alfort), <sup>d</sup> Direction of Regulated Product Evaluation (DEPR), Residue and Food Safety Unit (URSA), National Agency for Food, Environmental and Occupational Health & Safety (ANSES, Maisons-Alfort)

The soil bacterium *Bacillus thuringiensis* (Bt) can produce spores and crystals composed of toxins with insecticidal activity. Due to these properties, some Bt strains have been developed as biopesticides for pest control, with a high value to restrict the use of chemical insecticides. However, Bt strains belong to the group of *Bacillus cereus* (Bc) and share some virulence genes encoding diarrheal enterotoxins such as hemolysin BL (Hbl), as well as non-hemolytic enterotoxin (Nhe) and cytotoxin K (CytK2), known to be involved in foodborne outbreaks (FBOs) following Bc contamination. For this reason, it has been suggested that Bt may also provoke FBOs and it is therefore crucial to evaluate their virulence potential and to characterize their effects on humans.

For this purpose, differentiated human intestinal Caco-2 cells have been treated with serial dilutions of the culture supernatant of five different Bt strains used as biopesticides. The cytotoxicity (MTT assay) and inflammatory response (interleukin-8 release) were first measured after a 24 h treatment. No cytotoxicity was induced with all the Bt supernatants up to a 25% dilution in cell medium, but the secretion of interleukin 8 was significantly induced.

To investigate the molecular effects of Bt supernatants on Caco-2 cells, untargeted proteomics and metabolomics were performed on both cell media and lysates after a 24h treatment with 3 selected non-cytotoxic concentrations. Peptides and metabolites were detected using a Q Exactive Plus mass spectrometer coupled to a Vanquish micro-UHPLC system (Thermo Fisher Scientific) and data were analyzed using Maxquant and Compound Discoverer tools. The pathways affected will be compared for all the Bt strains.

In conclusion, this project will better characterize the toxic effects induced by a panel of Bt strains on the human intestinal barrier. The relationship between the production of toxins and the pathways affected will be further investigated.

## Pipetting-free single cell analysis with the label-free proteoCHIP and the Evtip adapter for high sensitivity proteomics on the timsTOF SCP.

Christophe SIROIT<sup>a</sup>, Pierre-Olivier SCHMIT<sup>b</sup>, Christoph KRISP<sup>c</sup>, Dorte B. BEKKER-JENSEN<sup>d</sup>, David HARTLMAYR<sup>e</sup>, Andreia ALMEIDA<sup>f</sup>, Gary KRUPPA<sup>g</sup>, Anjali SETH<sup>e</sup>, Moritz HEUSEL<sup>d</sup>, Magnus HUUSFELDT<sup>d</sup>, Thorsten LEDERTHEIL<sup>c</sup>, Jean-Francois GREISCH<sup>c</sup>, Jarrod SANDOW<sup>f</sup>, Guilhem TOURNIAIRE<sup>e</sup>, Nicolai BACHE<sup>d</sup>, Markus LUBECK<sup>c</sup>

<sup>a</sup> Bruker France S.A, Wissembourg, <sup>b</sup> Bruker France S.A, Wissembourg, France, <sup>c</sup> Bruker Daltonics GmbH & Co. KG, Bremen, Germany, <sup>d</sup> Evosep Biosystems, Odense, Denmark, <sup>e</sup> Cellenion, Lyon, France, <sup>f</sup> Ionopticks, Melbourne, Australia, <sup>g</sup> Bruker s.r.o., Brno, Czech Republic

### Introduction

Single cell proteomics is a rapidly developing field with the potential to make important contributions to the understanding of cellular heterogeneity. Single cell protein extraction, minimal exposure of samples to surfaces and optimal storage and transfer conditions are crucial for loss-less single cell proteome analyses. Recent enhancements in trapped ion mobility spectrometry (TIMS) coupled to fast and sensitive mass spectrometry established in the timsTOF SCP paired with automated single cell sorting and sample preparation realized with the cellenONE® platform allows for sensitive proteome analyses at the single cell level.

### Methods

Single HeLa cells were isolated into the label-free proteoCHIP®, directly lysed, and proteins digested at 50 °C with high humidity on deck using the cellenONE platform. The label-free proteoCHIP with tryptic peptides was placed into the nanoElute2 autosampler, peptides injected onto a 25 cm Aurora C18 column (IonOpticks) and eluted into a timsTOF SCP. cellenONE isolated and prepared single HeLa cells in a Evtip adapter for pipetting-free single cell transfer by centrifugation were loaded onto Evtip Pure tips and analyzed with an Aurora Elite column using Whisper 40 SPD and a timsTOF SCP. dia-PASEF data were processed with Spectronaut 17 (Biognosys) using directDIA for project-specific library generation. For quantitative analyses, the quantification results output, with quan-data filter applied, was used.

### Preliminary Data

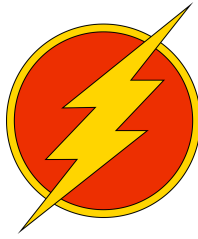
Sample pick-up directly from the label-free proteoCHIP was assessed with HeLa lysate digests (Pierce) showing excellent reproducibility at various concentrations. Injections of 250 pg of HeLa peptides on column (1 µL in well) resulted in 15,000 peptides from 2,600 proteins which was matched by 250 pg HeLa peptides injected from a vial (250 pg/µL). We then analyzed single HeLa cells which were directly isolated and prepared in the label-free proteoCHIP (48 cells) and identified in average more than 3400 proteins per single cell with good reproducibility. 3200 proteins were quantified in at least half of the analyzed samples.

Further, we assessed the Evtip adapter by first comparing manual transfer versus transfer by centrifugation. Half of the prepared cells in the Evtip adapter plate were manually pipetted onto the Evtips whereas the remaining half were transferred by centrifugation. Data analysis revealed increased precursor identification reproducibility in cells transferred by centrifugation. In a second experiment, HeLa cells were isolated according to their cell size and binned into either 18 - 20 µm, 21 - 23 µm, 24 - 26 µm or 27 - 30 µm cell size groups. Protein group identification rates across the different cell size groups were comparable with all being close to 4000 protein groups. However, the reproducibility in the smaller cells (18 - 23 µm) was better than in the larger cells (24 - 30 µm) with CVs of 4.69 % (18 - 20 µm), 2.68% (21 - 23 µm), 7.65% (24 - 26 µm), and 13.05% (27 - 30 µm). Quantitative comparison between the protein groups identified in the different cell size groups was performed. 3922 proteins were present in at least 48 single cells of the 96 single HeLa cells. Principal component analysis of the 2018 proteins identified in all 96 single cell showed clear difference between the small cells and the larger cells.

### Novel aspect

Label-free analysis workflow that reproducibly identifies up to 4000 proteins from single HeLa cells using cellenONE platform and timsTOF SCP





## LC-MS/MS analysis of penguin tissues and customization of protein database to identify novel antimicrobial peptides

Bastien ARNAUD<sup>a</sup>, Adrien BROWN<sup>b</sup>, Philippe BULET<sup>c</sup>, Fabrice BERTILE<sup>b</sup>

<sup>a</sup> Inovarian, Paris, France, <sup>b</sup> BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC UMR 7178, University of Strasbourg, CNRS, Strasbourg, France, <sup>c</sup> Institute for Advanced BioSciences, IAB UMR 5309, CNRS, INSERM, Université Grenoble Alpes, La Tronche, France

### Introduction

According to the WHO, the clinical pipeline and recently approved antibiotics are insufficient to tackle the challenge of increasing emergence and spread of antimicrobial resistance. The identification and improvement/development of natural antimicrobial peptides (AMPs) is a promising alternative to conventional antibiotics, due to their low probability of inducing resistance. In the early 2000s, we have isolated and fully characterized two isoforms (AvBD103a, AvBD103b) of an AMP in the King penguin (*A. patagonicus*), namely the avian  $\beta$ -defensin spheniscin<sup>1</sup>. We showed that AvBD103b is active against micro-organisms (G+, G- bacteria, yeast and filamentous fungi) even at high salt concentration<sup>2</sup>. This is a considerable advantage to fight the pathogens that develop in salt-rich environments (e.g. during cystic fibrosis or ocular infections). Therefore, our ultimate goal is to decipher the mechanisms of action of spheniscin. Here we seek to analyze the structural diversity of AMP members of the spheniscin family by refining the annotation of the King penguin genome analyzing and several penguin tissue samples by mass spectrometry.

### Methods

Penguin samples selected based on their septic and saline environment (bone, lung, kidney, mouth floor, tongue, stomach, esophagus, duodenum and jejunum) were cryogrinded using a ball mill, and proteins were extracted in urea/thiourea/CHAPS or Laemmli buffer. After protein assay, 50  $\mu$ g of proteins were loaded onto glycine SDS-PAGE and tricine SDS-PAGE gels and electrophoresed. Eight to ten protein bands of 2mm width were excised in the 1-10kDa range, and trypsin *in-gel* digestion was performed after reduction and alkylation. Samples were analyzed using a nanoElite chromatography system coupled to a timsTof Pro mass spectrometer (Bruker). MS raw data were processed using Mascot (v2.8.1) and ProlineStudio (v2.1.2) with an FDR of 1% at PSM and protein levels. Spectra were searched against a home-made protein database containing protein sequences available from three different annotations of the *A. patagonicus* genome. We also added to the database protein sequences derived from our own annotation of the *A. patagonicus* genome and comparison (blast) with sequences of all known beta-defensins in specialized resources (APD and DRAMP). In addition, a de novo sequencing strategy was considered using Novor (v1.1; min peptide score 50) and blastp v2.12.0.

### Results

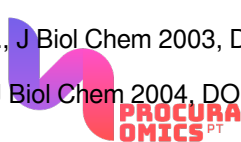
We successfully built a protein database enriched with 29 putative penguin beta-defensins in addition to the 8 already annotated. From there, sequence clustering and alignment allowed us to eliminate redundancies, and we finally retained a database version with 25 penguin defensin sequences. Strikingly, retained sequences were rich in well-aligned cysteine motifs, which supports their actual defending status. Mascot searches allowed 10 penguin defensins to be identified, and the de novo sequencing strategy allowed detection of an additional one. Defensin sequence coverage was determined in the 9-47% range, which is unreasonably low and will be corrected, due to the fact that the sequences we predicted were artificially elongated to avoid missing the true 'start' and 'stop' of penguin defensins. Tissue specificity could be noticed, the bone, lung and tongue already bringing the identification 10 of the 11 defensins.

### Conclusion/perspectives

The combination of bioinformatics manipulation of protein sequence databases, mass spectrometry, classical Mascot analysis, and de novo sequencing proved relevant and successful. Further analyses should increase the coverage of penguin defensin sequences. These results will be used to better understand the structure/function relationship of penguin defensins. In parallel, we will conduct a bio-guided strategy to evaluate which of the 11 defensins we identified are actually active against microorganisms under high salinity conditions.

<sup>1</sup>Thouzeau et al., J Biol Chem 2003, DOI:10.1074/jbc.M306839200

<sup>2</sup>Landon et al., J Biol Chem 2004, DOI:10.1074/jbc.M401338200



## Comparative proteomic analysis of the terminal erythroid differentiation of two cell lines with primary erythroblasts

Maryam ABOUTIMAN<sup>a</sup>, Sylvain METRAL<sup>b</sup>, Laetitia CLAER<sup>b</sup>, Marjorie LEDUC<sup>c</sup>, François GUILLONNEAU<sup>d</sup>, Léa KUZNICKI<sup>b</sup>, Yves COLIN<sup>b</sup>, Olivier HERMINE<sup>e</sup>, Isabelle MOURO-CHANTELOUP<sup>b</sup>, Morgane LE GALL<sup>c</sup>, Emilie-Fleur GAUTIER<sup>c</sup>, Benoit GAMAIN<sup>b</sup>, Sandrine LAURANCE<sup>b</sup>

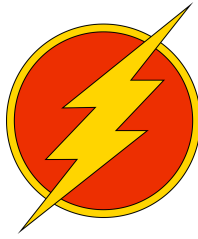
<sup>a</sup> PROTEOM'IC, INSTITUT COCHIN, Université Paris Cité, INSERM U1016, CNRS UMR8104, Team « Molecular mechanisms of hematological disorders and therapeutic implications », IMAGINE Institute, INSERM U1163, Laboratoire d'Excellence du Globule Rouge - Gr-Ex, <sup>b</sup> INSERM UMR\_S 1134, Université Paris Cité, Laboratoire d'Excellence du Globule Rouge - Gr-Ex, <sup>c</sup> PROTEOM'IC, INSTITUT COCHIN, Université Paris Cité, INSERM U1016, CNRS UMR8104, Laboratoire d'Excellence du Globule Rouge - Gr-Ex, <sup>d</sup> PROTEOM'IC, INSTITUT COCHIN, Université Paris Cité, INSERM U1016, CNRS UMR8104, <sup>e</sup> INSERM UMR\_S 1134, Université Paris Cité, Team « Molecular mechanisms of hematological disorders and therapeutic implications », IMAGINE Institute, INSERM U1163, Laboratoire d'Excellence du Globule Rouge - Gr-Ex

The use of *in vitro* cell lines has been extremely useful, throughout the years, to study diverse biological phenomenon, and elucidate normal and pathological mechanisms. In the case of erythropoiesis, several cell lines have been established in this regard <sup>1</sup>. We performed a comparative analysis of protein expression and evolution all along erythroid differentiation of two cell lines and primary cells in the aim to evaluate the similarities and differences. Although these cell lines have been characterized at a cellular level, they were never characterized at a proteomic level. However, a comparative proteomic functional analysis of the proteome of a cell line with primary cell counterpart is highly informative to know which proteins and functions can be studied or not in this cell line.

For each cell lines, four different *in vitro* erythroid differentiations were performed and purified erythroblasts populations were obtained for each stage of terminal erythroid differentiation by cell sorting. Each of these populations were analyzed using a label free quantification proteomic approach. Protein were digested by Filter aided sample preparation (FASP) and peptides fractioned by Strong Cation eXchange (SCX) in five fractions. Each fraction was analyzed by nLC-MS/MS mass spectrometry with a three-hour gradient. Proteins were identified and quantified using Maxquant Software. Using histones as a reference, we calculated the absolute quantification of proteins in copy number per cell. Data obtained were compared to already published proteomic data on peripheral blood CD34+ and cord blood CD34+ cell differentiation <sup>4,5</sup>. Furthermore, a functional analysis has been performed on differentially expressed to characterize the functions that are not correctly regulated in each of these two cell lines.

In both cell lines, over 7,000 identified proteins and over 4,000 quantified proteins have been found. Over 4,000 quantified proteins in the two cell lines have been aligned with those of the two CD34+ differentiated erythroblasts, with a good reproducibility between biological replicates. By this approach, we have established that the proteome of these cell lines is highly correlated with those of peripheral and cord blood CD34+, all along the differentiation. We observed high similarity in protein expression pattern for a large part of the proteins for the two cell lines. As a consequence, these results suggest that these cell lines could be used as a reliable model for terminal erythroid differentiation, at a proteomic level. Moreover, future analysis on specific proteins involved in erythropoietin response and enucleation could be done, in order to further comprehend those mechanisms.





## Optimization of sample preparation and nanoLC-MS/MS methods for single cell proteomic analysis

Pauline PERDU-ALLOY<sup>a</sup>, Laurence SABATIER<sup>a</sup>, Christine CARAPITO<sup>a</sup>

<sup>a</sup> Laboratoire de Spectrométrie de Masse BioOrganique, IPHC (UMR 7178), CNRS, 67000 Strasbourg, France

Understanding the various mechanisms in which proteins are involved is crucial for comprehending biochemical phenomena, including the onset and development of diseases<sup>1</sup>. Conducting qualitative and quantitative analysis of the entire proteome can provide insights into the response of an organ or a population of cells to treatments or diseases. However, analysis of proteins from an average cell population often neglects cellular heterogeneity and does not allow to distinguish healthy cells from infected cells, which cohabit in the same tissue<sup>2</sup>. The rise of single-cell proteomics is therefore meaningful because it allows the identification of the proteome expressed in individual cells with increasing depth at a given time<sup>3</sup>, with applications ranging from disease diagnosis to the response of a chemical treatment. However, despite its importance, the proteome analysis at the single-cell level remains a major analytical challenge due to the very low abundance of protein material available within a cell (<ng). Thus, every step of the workflow needs to be revisited to minimize losses during sample preparation and achieve highest sensitivity in liquid chromatography tandem mass spectrometry analysis (LC-MS/MS).

In recent years, the development of trapped ion mobility spectrometry (TIMS) coupled to a high sensitivity mass spectrometer (TimsQTOF) as well as the development of automated methods for sample preparation and sophisticated data software constitute an anchor for the analysis of protein material at the single cell scale<sup>4</sup>.

On the one hand, Hela cell digests (Pierce, Ref. 88329) were used to prepare a dilution series from 0.1 to 10 ng of injected material. On the other hand, HeLa cells were sorted, lysed and digested using the CellenONE instrument at Cellenion (Lyon). Varying cell numbers, from 1 to 20 cells were sorted, lysed, digested prior to nanoLC-MS/MS analysis.

Peptide mixtures were separated on an Aurora C18-RP 120Å (75µm x 250mm, 1.9µm) (IonOpticks) column using a nanoElute (Bruker Daltonics) coupled to a timsTOFpro2 (Bruker Daltonics) mass spectrometer. Separation was performed on a gradient from 2 to 30% ACN in 0.1% formic acid over 30min at 0.2µL/min. Data Dependent Acquisition - Parallel Accumulation and Serial Fragmentation (DDA-PASEF) and Data Independent Acquisition - PASEF (DIA-PASEF) methods were used. A comparative analysis was performed to assess the performance of two data analysis tools, MaxQuant and Proline for DDA data analysis, and Spectronaut and DIA-NN for DIA data interpretation.

In DDA-PASEF mode, around 300 proteins were identified in the 0.1 to 0.3ng region, while more than 1000 proteins could be identified in the same region in DIA-PASEF. These first results were encouraging to further explore the instrument's performances for very low amounts as these protein amounts are close to single cell protein contents (approx. 0.3ng). In isolated cells, an average of 540 proteins were identified in single cells in DDA-PASEF mode, while more than 2400 proteins were identified using a DIA-PASEF method. This remarkable outcome highlights the great capacities of DIA-PASEF for single-cell analysis.

Investigating the data processing aspect of DIA, comparing the two data processing software (Spectronaut and DIA-NN) revealed close average numbers of quantified proteins for single cells and up to 5 isolated cells.

While those initial results are promising, further optimizations are still required from sample preparation to nanoLC-MS/MS acquisition methods optimization and finally data processing to be able to unlock deeper insights into cellular heterogeneity and disease mechanisms.

<sup>1</sup> Minakshi, *Single-Cell Proteomics: Technology & Applications*. (2019).

<sup>2</sup> Marino, Zito, Accardo, Ronchi, Cozzolino, Morgillo, Rossi, Franco, *Molecular heterogeneity in lung cancer*. (2019).

<sup>3</sup> Templer, Ces, *New frontiers in single-cell analysis*. (2008).

<sup>4</sup> Meier, Park, Mann, *Trapped Ion Mobility Spectrometry and Parallel Accumulation-Serial Fragmentation in Proteomics*. (2021).

## Development and optimization of a method for automated peptide desalting on the DigestPro MSi robot using AttractSPE® C18 tips

Célia JARDIN<sup>a</sup>, Soumia HAMADA<sup>b</sup>, Cédric PIONNEAU<sup>b</sup>, Martin TECHNAU<sup>c</sup>, Solenne CHARDONNET<sup>b</sup>

<sup>a</sup> Sorbonne University, Inserm, UMS PASS, Plateforme Post-génomique de la Pitié-Salpêtrière, P3S, <sup>b</sup> Sorbonne University, Inserm, UMS PASS, Plateforme Post-génomique de la Pitié-Salpêtrière, P3S, Paris, FRANCE, <sup>c</sup> CEM GmbH, Carl-Friedrich-Gauß-Straße 9, D-47475 Kamp-Lintfort, Germany

Bottom-up approaches are commonly used for proteomics analysis in LC-MS/MS. However, the salts present in the digestion buffers can generate ion-suppression that can limit the detection of peptides and greatly impact proteins identification. Therefore, peptide desalting is an indispensable step. The P3S proteomics facility is equipped with a nanoElute - timsTOF Pro (Bruker) LC-MS/MS system and has chosen to perform the desalting step prior to LC-MS/MS analysis with direct injection on the analytical column.

With the aim to automate peptide desalting, P3S has asked CEM (previously Intavis) to create a custom needle to run home-made StageTips on the DigestPro MSi robot. The program was adapted from manual StageTips protocol and further optimized by playing with liquid aspirate and dispense speeds and volumes. Moreover, P3S has evaluated AttractSPE® C18 Tips (Affinisept) compared to home-made StageTips (with Empore C18 SPE disks) from 2 µg and 100 ng of protein digest. The AttractSPE® Tips are packed with small sorbent beads, embedded in a thin and mechanically stable membrane that combines high capacity and small dead volume, and are adapted for centrifugation or positive pressure assays. The comparison showed little difference between the two kinds of tips when working with 2 µg of peptides but showed a 10% increase in the number of identified proteins with the AttractSPE® C18 Tips when working with 100 ng of peptides. Moreover, the peptide intensities were higher with the AttractSPE® C18 Tips. In both cases, 2 µg and 100 ng starting material, the AttractSPE® C18 Tips performed better in the hydrophilic range.

Our results show that AttractSPE® C18 sorbent offers a wider spectrum of interactions with a broadest range of peptides, from the most hydrophilic to the most hydrophobic ones, compared to home-made StageTips. In addition, the P3S team has optimized the program on the DigestPro MSi robot to reduce the time from 40 min down to 10 min per sample.

# Sponsor Advertisements





MASS SPECTROMETRY

## timsTOF Series

### Next generation ion mobility platforms for high speed, high sensitivity 4D-Multiomics

Bruker's timsTOF series introduces the next generation of ion mobility mass spectrometry. Utilizing trapped ion mobility spectrometry (TIMS) unlocks an additional dimension of separation and delivers revolutionary improvements in enhanced specificity and robustness.

- **timsTOF Pro 2** - the platform of choice for 4D-Multiomics applications with its proven robustness, high sensitivity and fast MS/MS acquisition
- **timsTOF HT** - unprecedented depth and high throughput for 4D-Proteomics™
- **timsTOF SCP** - ultra-high sensitivity to unlock deep quantitative single cell 4D-Proteomics™
- **timsTOF fleX** - with **MALDI-2** and **microGRID** towards enhanced sensitivity at high spatial resolution
- **timsTOF MALDI PharmaPulse®** - taking label-free high-throughput screening to the next level



For Research Use Only. Not for use in clinical diagnostic procedures.

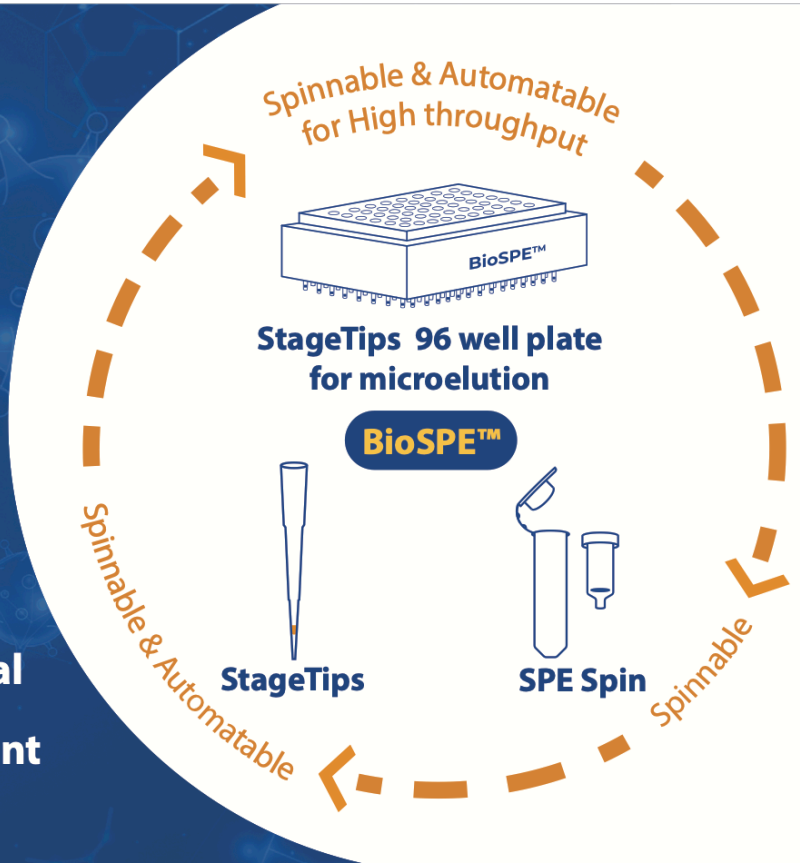
For more information please visit [www.bruker.com](http://www.bruker.com)

Innovation with Integrity

“

Every **PROTEOMIC** challenge has its **BioSPE™** solution

- Peptide desalting
- Peptide fractionation
- Protein purification
- Magnetic beads removal
- Glycopeptide enrichment



[www.affinisep.com](http://www.affinisep.com)



**Covaris**

## Your LC-MS Samples Deserve The Best Treatment

Secure Your Data by Investing in Sample Preparation!



Visit the Covaris booth and our presentation on “Scalable Sample Prep Methods for LC-MS Assays” by Nicolas Autret, Ph.D.

Tuesday, June 20<sup>th</sup> at 17:00

For more information, please contact:  
Nicolas Autret, Ph.D.  
([nautret@covaris.com](mailto:nautret@covaris.com))

Data on file. Information subject to change without notice. For research use only. Not for use in diagnostic procedures.

USA: Tel: +1 781 662 3939 | Email: [sales@covaris.com](mailto:sales@covaris.com) | Europe: Tel: +44 (0)1835 873 0133 | Email: [emsa.sales@covaris.com](mailto:emsa.sales@covaris.com) | APAC: +66 131 627 8714  
Email: [APAC@covaris.com](mailto:APAC@covaris.com) | Web: [www.covaris.com](http://www.covaris.com) | Solutions: [application@covaris.com](mailto:application@covaris.com) | Service and Support: [techsupport@covaris.com](mailto:techsupport@covaris.com) | May 2015 | 2023© Covaris, LLC

Stay Connected!





**CELLENION** >>

A BICO COMPANY

Automate and miniaturize  
your entire single cell  
proteomics sample  
preparation workflows with  
cellenONE and proteoCHIP.



scan here  
to learn more  
[cellenion.com](https://cellenion.com)



Mass spectrometry

**ThermoFisher**  
SCIENTIFIC

# Rethink what is possible

Orbitrap Astral  
mass spectrometer

Learn more at  
[thermofisher.com/OrbitrapAstral](https://thermofisher.com/OrbitrapAstral)

© 2023 Thermo Fisher Scientific Inc.  
All rights reserved. AD001724-EN 0423M

thermo scientific



## Are you fully satisfied with your LC-MS service?

Are you looking for an alternative service and support supplier for your ThermoFisher, Sciex or Waters LC-MS instruments? As the only independent service provider with own hardware R&D, MS Vision has the knowledge to provide high quality service for your instruments.

- Flexible, cost effective and transparent service solutions
- Instrument qualifications
- Relocations
- Trainings
- Instrument upgrades (High mass, ECD, Omnitrap)
- Refurbished LC-MS instruments installed with warranty



**MS Vision** – your trusted LC-MS partner for almost 20 years. ISO 9001:2015 accredited.  
[www.msvision.com](http://www.msvision.com) | [info@msvision.com](mailto:info@msvision.com) | +44 8453 732 493



## Imagerie MS et analyses HRMS **iMScope QT**

*Fusion de la chromatographie, microscopie et HRMS*



Informations qualitatives & quantitatives sur les composés d'intérêt et sur leur distribution dans un tissu biologique



Analyse rapide des résultats d'imagerie sur un logiciel unique



Passage d'une technique à l'autre en moins de 20 min



[www.shimadzu.fr](http://www.shimadzu.fr)

## ENRICH-iST kit for plasma and serum sample preparation

- Enrichment of low-abundant proteins reducing the proteomes dynamic range
- From plasma to ready-to-measure peptides in 5h, up to 96 samples in parallel
- High reproducibility  $R^2 > 0.95$
- Highly flexible in starting material (species independent and compatible with serum)
- Fully automatable on liquid handling platforms suitable for magnetic bead-based workflows



[www.preomics.com](http://www.preomics.com)

## Built for standardization

The six standard Evosep One methods cover the range from ultra-high throughput single-shot analysis to full proteome analysis with fractionation strategies.

	Gradient Length (min)	Cycle Time (min)	Flow Rate (µl/min)
High throughput	2.2	2.9	4.0
	3.2	4.8	4.0
	5.6	7.2	2.0
	11.5	14.4	1.5
Proteome coverage	21.0	24.0	1.0
	44.0	48.0	0.5



Evosep One methods are designed and tested to provide the best chromatography with the lowest possible system overhead every time, all the time.

**EVOSEP**



### Are you ready for a step-change in data independent

Introducing  
**Zeno SWATH**  
data independent acquisition (DIA)  
exclusively on the  
**ZenoTOF 7600 system**

Enhance speed, sensitivity and accuracy of your workflows using Zeno SWATH data independent acquisition (DIA) for MS-based proteomics.

Harnessing the power of the Zeno trap for Zeno SWATH DIA allows you to:

- Confidently discover and translate significant biomarkers
- Identify and quantify up to twice as many proteins as with traditional SWATH DIA approaches
- Shortened run times with minimal compromise in proteome coverage

Seeing is believing at:  
[SCIEX.com/zenoswathdia](http://SCIEX.com/zenoswathdia)

The SCIEX clinical diagnostic portfolio is For In Vitro Diagnostic Use. In-Only Product(s) not available in all countries. For information or availability, please contact your local sales representative or refer to [www.scix.com/diagnostics](http://www.scix.com/diagnostics). All other products are For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein, including associated logos, are the property of AB Sciex Pte. Ltd. or their respective owners in the United States and/or certain other countries. See [www.scix.com/trademarks](http://www.scix.com/trademarks). © 2015 SCIEX, Inc. All rights reserved. RUJ-MKT-07-1N808-A



Venez découvrir notre générateur d'azote et d'air aux avantages uniques:

- **Intelligent** : contrôle du débit le plus efficace.
- **Silencieux et sans vibrations**
- **Écologique** : Consomme jusqu'à 30 % d'énergie en moins.
- **Entretien réduit**: première maintenance compresseur à 20000 heures de fonctionnement
- **Capacité d'accès à distance** : diagnostique par nos experts

Pour TOUS les instruments LC-MS  
LNI Swissgas



Because stable isotope MS methods are bringing scientists closer to curing human diseases through quantitative biomarker analysis and quantitative proteomic analysis.

Eurisotop is pleased to offer a variety of stable isotope reagents for labeling any proteome for quantitative MS analysis.

Metabolic Labeling

QC & Qtt Kits

Chemical Tagging

Amino Acids

Glycans

Cell Growth Media

 **Eurisotop**  
www.eurisotop.com

### PTMscan HS : Simplified PTM Workflow solution



Improved sensitivity with less starting sample

**Detect thousands of post-translationally modified (PTM) peptides with the enhanced methods in PTMscan HS kits:**

- PTMscan HS Phospho-Tyrosine (P-Tyr-1000) Kit #38722
- PTMscan Acetyl-Lysine motif (AC-K) Kit # 46784
- PTMscan Ubiquitin/SUMO Remnant Motif (K-e-GG) kit # 59322
- PTMscan Succinyl-Lysine Motif ( Succ-K) kit # 60724

And many more on [cellsignal.com/applications/proteomics](https://cellsignal.com/applications/proteomics)

