24-25 April, 2023
Utrecht (NL)

11th Symposium on Structural Proteomics

About The Event

11th Symposium on Structural Proteomics.

The symposium was held on Monday 24 & Tuesday 25 of April, 2023 in Gasthuis Leeuwenbergh, Servaasbolwerk 1a, Utrecht.

Two days of interesting keynotes and lectures about topics like crosslinking mass spectrometry, HDX, and top-down/native mass spectrometry.

WHERE

Gasthuis Leeuwenbergh
Servaasbolwerk 1a
3512 NK Utrecht

WHEN

24-25 April, 2023 Monday and Tuesday

About The Event Organizers

The symposium was organized by:

Dr. Richard Scheltema of Utrecht University. The overarching theme of his research is the development and application of mass spectrometry approaches for the detailed characterization of proteins in the context of their structure and their interactions to other proteins and DNA/RNA.

Dr. Charlotte Uetrecht of CSSB Hamburg. She focuses on uncovering the structural dynamics in the assembly of norovirus particles and coronaviral replication complexes by applying state-of-the-art structural mass spectrometry.

Dr. Pascal Albanese currently holds a MSCA individual fellowship to work at the CEA in Grenoble (France). He’s committed toward establishing a seamless connection between cryo-EM and structural Mass Spectrometry to study light-dependent protein dynamics in photosynthesis regulation.

About the speakers

Prof. Dr. David Schriemer – cryo-proteomics

Prof. Dr. Konstantinos Thalassinos – cyclic ion mobility

Dr. Aneika Leney – native mass spectrometry

Dr. Anton Calabrese – integrative mass spectrometry

Dr. Francis O’Reilly – crosslinking mass spectrometry


Dr. Malvina Papanastasiou – HDX


Dr. Markku Varjosalo – Interactomics


Dr. Panagiotis Kastritis – integrative mass spectrometry


Prof. Dr. Alexandre Bonvin – Structural Modeling

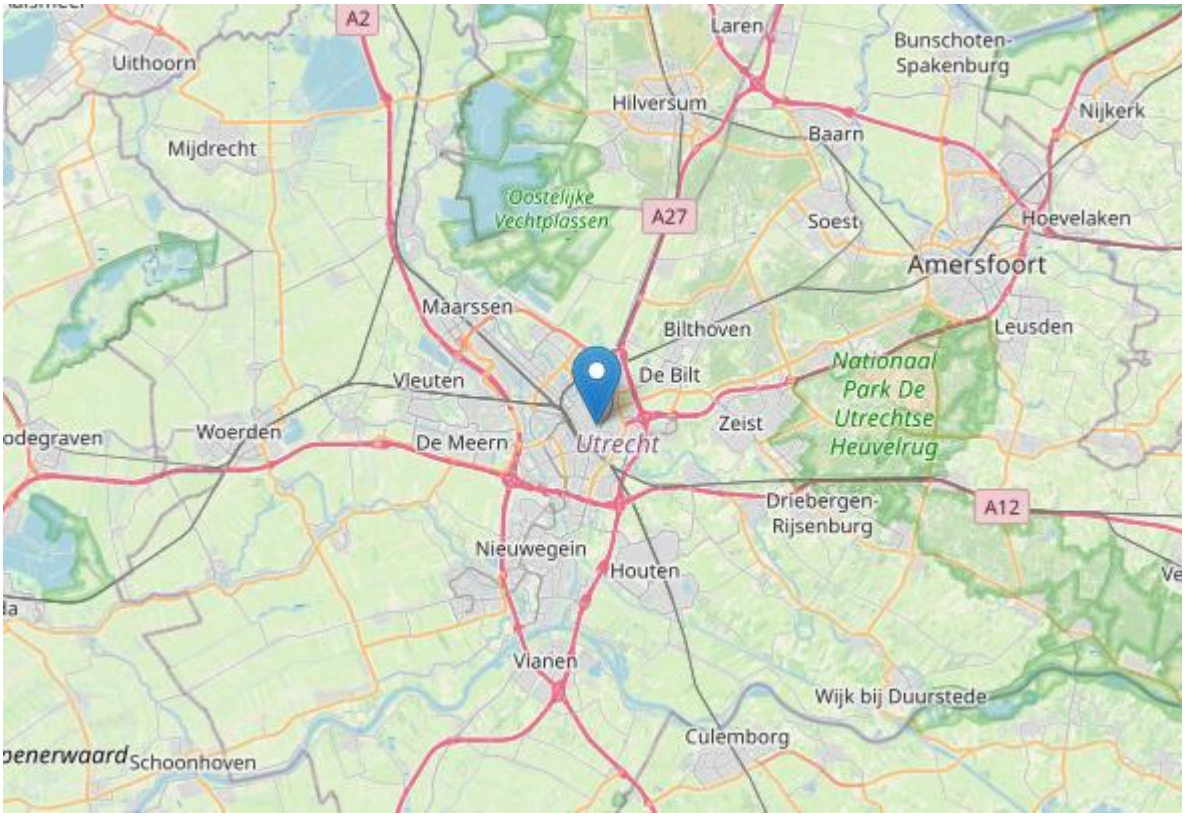

Prof. Dr. Kelly K Lee – HDX


Prof. Dr. Dina Schneidman – Structural Modeling

Venue

Gasthuis Leeuwenbergh
Servaasbolwerk 1a
3512 NK Utrecht

The symposium was held in the beautiful Gasthuis Leeuwenbergh. A former church, the building plays a central role in organizing gatherings around music and science.



Schedule

Monday	Tuesday
Aneika Leney	Malvina Papanastasiou
9:00 am - 9:20 am	9:00 am - 9:20 am
Angelique Sanchez Dafun	Kelly K Lee
9:40 am - 9:55 am	9:20 am - 9:40 am
Thermo - Albert Konijnenberg	Frantisek Filandr
9:55 am - 10:10 am	9:40 am - 9:55 am
Kostas Thalassinos	Nadine Hellmold
10:10 am - 10:30 am	9:55 am - 10:10 am
Lars Thiede	Tatiana Shamorkina
11:00 am - 11:15 am	10:10 am - 10:25 am
Albert Heck	Joey Sheff
11:15 am - 11:35 am	11:00 am - 11:15 am
David Schriemer	Catarina Malta
11:35 am - 11:55 am	11:15 am - 11:30 am
Catherine Gilbert	Markku Varjosalo
11:55 am - 12:10 pm	11:30 am - 11:50 am
Alexander Leitner	Chloé Van Leene
12:10 pm - 12:25 pm	11:50 am - 12:05 pm
Henning Urlaub	Cathy Marulli
12:25 pm - 12:40 pm	12:05 pm - 12:20 pm
Francis O'Reilly	Sebastian Steiner
2:00 pm - 2:20 pm	12:20 pm - 12:35 pm
Tara Bartolec	Anton Calabrese
2:20 pm - 2:35 pm	2:00 pm - 2:20 pm
Andrea Graziadei	Markus Räschle
2:35 pm - 2:50 pm	2:20 pm - 2:35 pm
Alexandre Bonvin	Bruker
3:30 pm - 3:50 pm	2:35 pm - 2:50 pm
Dina Schneidman	Panagiotis L. Kastritis
3:50 pm - 4:10 pm	3:30 pm - 3:50 pm
Anton Feenstra	MSVision - Steven Daly
4:10 pm - 4:25 pm	3:50 pm - 4:05 pm
Hadeer Elhabashy	Franz Herzog
4:25 pm - 4:40 pm	4:05 pm - 4:20 pm
Arunima Singh	Awital Bar Barroeta
4:40 pm - 5:10 pm	4:20 pm - 4:35 pm
	Dina Schuster
	4:35 pm - 5:00 pm

Abstracts

Monday

Using native mass spectrometry to determine the strength of molecular glues

Aneika Leney
School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

Approximately 300,000 protein-protein interactions (PPI) occur in human cells. The disruption of these PPIs into their individual protein constituents contributes to many human diseases including cancer and neurodegeneration. Thus, there is a growing interest in the development of drugs that 'glue' protein interactions back together. However, the discovery of novel glues has been hindered due to the lack of tools available to monitor PPI stabilisation.

The work presented will focus on the development of native mass spectrometry to aid in the search for novel PPI stabilisers. We focus on the eukaryotic regulatory protein 14-3-3 and its binding partners estrogen receptor ERα, and a peptidyl prolyl isomerase Pin1 and monitor their binding equilibria upon the addition of two different stabilisers, FusA and A28. Both these molecular glues act by different modes of stabilisation that we readily differentiated by tandem mass spectrometry experiments. Overall, the data highlights how native MS is making advances in the discovering of novel molecular glues and the mechanism behind how these glues stabilise protein-protein interactions.

Establishing proteoform footprints of 20S proteasomes from biological and patient samples using top-down mass spectrometry

Angelique Sanchez Dafun[1]; Dušan Živković[1]; Stephen Adonai Leon Icaza[1]; Sophie Möller[2]; Carine Froment[1]; Delphine Bonnet[3,4]; Adriana Almeida De Jesus[5]; Laurent Alric[4]; Muriel Quaranta-Nicaise[3]; Audrey Ferrand[3]; Céline Cougoule[1]; Etienne Meunier[1]; Odile Burlet-Schiltz[1]; Frédéric Ebstein[2,†]; Raphaela Goldbach-Mansky[5]; Elke Krüger[2]; Marie-Pierre Bousquet[1]; Julien Marcoux[1]
1 Institut de Pharmacologie et de Biologie Structurale (IPBS), Université de Toulouse, CNRS, Université Toulouse III—Paul Sabatier (UPS), 31077 Toulouse, France
2 Institute of Medical Biochemistry and Molecular Biology, University Medicine Greifswald, 17475 Greifswald, Germany
3 IRSD, Université de Toulouse, INSERM, INRAE, ENVT, Université de Toulouse III—Paul Sabatier (UPS), 31300 Toulouse, France
4 Internal Medicine Department of Digestive Disease, Rangueil Hospital, Université de Toulouse III—Paul Sabatier (UPS), 31400 Toulouse, France
5 Translational Autoinflammatory Diseases Section, LCIM, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA
* Correspondence: marie.pierre-bousquet@ipbs.fr (M.-P.B.); julien.marcoux@ipbs.fr (J.M.) † Present address: L'institut du Thorax, Nantes Université, Inserm UMR 1087/CNRS UMR 6291, 44000 Nantes, France.

Introduction: Proteasome (20S) is the main responsible for protein degradation within the cells, making it a good therapeutic target for some diseases. Its activity is altered by post-translational modifications (PTM), genetic variations, and replacement of its catalytic subunits. Due to its complexity and heterogeneity, these informations are difficult to obtain through immunochemical methods and classical peptide-based approaches. To address this, we developed a miniaturized workflow combining top-down MS (TD-MS) and bottom-up MS (BU-MS) of immunopurified 20S providing their assembly status and proteoform footprint, revealing PTMs, mutation, single-nucleotide-polymorphisms (SNP) and induction of alternative subunits in different samples: organoids, biopsies, B-lymphoblastoid cell lines (BLCL) from Proteasome-Associated-Autoinflammatory-Syndrome (PRAAS) patients. This technique enables better description of 20S, which can later lead to more targeted therapies.
Methods: Proteasome complexes were immunopurified from cell lysates and analyzed by online nano-LC (UltiMate-3000) coupled with an Orbitrap Fusion Tribrid (Thermo). Each sample (1-2 pmol) was injected onto a reverse-phase C4-precolumn (Thermo) and separated on an in-house analytical C4 nanocolumn. MS and MSMS (EThcD) scans were acquired in positive and intact protein modes, at 7,500 and 60,000 resolution, respectively. MS spectra were deconvoluted using RoWinPro and visualized with VisioProtMS. MSMS data was processed in Proteome Discoverer 2.3 (Thermo) using a three-tier search to identify the proteoforms from a custom database. Label-free quantification of 20S subunits was done with deconvoluted results from UniChrom by applying correction using relative

ionization yield factors obtained beforehand, from pure constitutive (c20S) and immunoproteasome (i20S).
Preliminary Data: We demonstrated the benefits and current limitations of 20S immunopurification followed by TD-MS to unravel the proteoforms of 20S subunits and to gain information on their PTMs, PRAAS-related mutations, common SNPs and immuno subunits on IFNγ-induced Caco2 cells, lung organoid and intestinal crypts. We identified a novel triphosphorylated and oxidized proteoform of α7 subunit in different biological samples that is independent of 20S subtype and is highly unstable, which likely explains why conventional BU-MS and immunoblotting techniques were unable to detect it so far. With our approach combining TD-MS and BU-MS, we were able to confirm mutations in BLCLs produced from PRAAS patients, which provided insights about the potential impact of these mutations on 20S assembly. In addition, we were able to track the i20S subunits in samples with increasing levels of complexity, including cell lines, organoids, and biopsies. We noticed that 1-3 mg of total protein were sufficient as a starting point to establish these 20S proteoform footprints, suggesting that this strategy can be used with rare biological samples such as disease-related biopsies. We also investigated the potential of TD-MS for intra-acquisition relative quantification of different subunits to establish the proportion of c20S vs. i20S in biological samples. We confirmed that the relative ionization yields (RIY) of the different subunits were quite different (30% to 160% relative to the average of all the non-catalytic subunits), but reproducible. This allowed us to semi-quantify the relative amount of catalytic subunits after induction of the i20S in Caco2 cells, colorectal crypts from patients, and BLCLs produced from PRAAS patients. Even though the standard deviations were not negligible, the obtained values were in good agreement with the expected ratios from label-free BU-MS, showing that the correction of TD-MS abundances with RIY can be used to determine the relative abundance of c20S vs. i20S proteasomes.
Novel Aspect: Characterization of new proteasome proteoforms; Analysis of proteasome from disease-related biopsies; Label-free quantification using relative ionization yields

Direct single molecule imaging on a modified Q Exactive UHMR with electron holography capability
Albert Konijnenberg[1], Josh Gilbert[2], Kyle Fort[3], Alexander Makarov[3], Alan Bahm[2]
1 Thermo Fisher Scientific, Achtseweg Noord 5, 5651 GG Eindhoven, the Netherlands
2 Thermo Fisher Scientific, 5350 NE Dawson Creek Drive, Hillsboro, OR97124, USA
3 Thermo Fisher Scientific, Hannah-Kunath-StraBe 11, 28199 Bremen, Germany
Introduction: In order to understand protein function, it is critical to have access to a protein structure. Most of the currently available experimental structural information comes from either X-ray crystallography (XRD) or cryo-electron microscopy (Cryo-EM). However, neither of these techniques allow the user to directly correlate proteoform information such as post translational modifications to the protein conformation, which can be critical to fully understand protein function. The ability of mass spectrometry to provide unaveraged data allows for detection of multiple coexisting proteoforms or conformations. However, structural mass spectrometry techniques such as Ion Mobility or Hydrogen Deuterium Exchange MS provide inferred structure at best, and often still requires a detailed 3D protein structure in order to interpret the data.
Methods: Recently, low energy electron holography (LEEH) was shown to be able to image nanoscale particles (1). Critical to this breakthrough was the ability to prepare an ultraclean substrate, which was done through *in vacuo* in beam deposition using a mass spectrometer, after which the samples were transferred using a vacuum suitcase to a different experimental setup for imaging. Using a mass spectrometer to prepare samples also characterization and even selection of different species (proteoform, conformation or even one protein from a complex matrix) for subsequent imaging. Furthermore, since LEEH makes use of low energy electrons it is particularly well suited to image proteins, as there is little to no damage when these proteins are exposed to the electron beam.
Preliminary data: Here we present an integrated mass spec-based system that can image individual single particles with high contrast and over extended periods of time. Our system combines ion beam deposition of native proteins on a modified Q Exactive UHMR mass spectrometer with direct protein imaging capabilities using single particle LEEH within a single instrument. The instrument workflow consists of 3 consecutive steps: acquiring a native MS spectrum to select the protein or proteoform of interest, energy-controlled deposition of the mass selected protein on an ultraclean graphene monolayer at ultrahigh vacuum and image the deposited proteins on the substrate using LEEH. In order to validate the instrument and assess its performance we will show how our setup imaged various types of proteins over a broad size range. We will demonstrate how the resulting holograms

can be reconstructed to generate 2D protein images and how due to the non-damaging nature of the low energy electrons, even videos can be acquired following the behavior and dynamics of the particles on the freestanding graphene. Finally, we show how single particle LEEH can image the structure and dynamics for a FAB-antigen complex (~70 kDa) that which was too small and dynamic for Cryo-EM studies, and where the conformational dynamics were absent in the crystal structure.
Novel aspect: Structural dynamics observed through low energy electron holography on proteins after ion beam deposition mass spectrometry on graphene supports.

Probing the Structure and Dynamics of Proteins by Means of Cyclic Ion Mobility Spectrometry and Electron Capture Dissociation

Kostas Thalassinos
University College London

Modern protein prediction methods using deep neural networks can now achieve extremely accurate results when predicting protein structures from the protein's sequence. Despite this impressive achievement these approaches still struggle to predict the structure of flexible proteins, and those where a single folded state does not represent the majority conformation.
By contrast, structural mass spectrometry approaches, and in particular ion mobility-mass spectrometry, can probe the dynamic conformational landscape of proteins and proteins in complex with other molecules.
Here we describe results using a travelling wave (TW) cyclic IM (cIM) device. The geometry of this instrument enables complex tandem IM experiments to be performed which were used to obtain more detailed collision induced unfolding pathways for the protein studied. In addition, we combined conformer selection with electron capture dissociation (ECD) in a single experiment to obtain a more in-depth characterisation of the different conformers. We used our multi-layered experimental approach to study proteins exhibiting multiple conformations, such as calmodulin, as well as proteins prone to misfolding and aggregation such as human islet amyloid polypeptide (hIAPP).

Probing viral capsid assembly by native MS

Lars Thiede[1,2], Ronja Pogan[1,2], Jürgen Müller-Guhl[1,3], Charlotte Uetrecht[1,2,3]
1 Centre for Structural Systems Biology (CSSB), Deutsches Elektronen-Synchroton DESY, Leibniz Institute of Virology (LIV), Hamburg, Germany.
2 Faculty V: School of Sciences, University of Siegen, Siegen, Germany.
3 Bernhard Nocht Institute for Tropical Medicine (BNITM), Hamburg, Germany

The assembly of viral capsids is a complex and still not fully deciphered process, that constitutes a vital step in the viral lifecycle. This is due to the morphological diverse nature of viral particles and the dynamics of the formation process, warranting more insight into the mechanisms of capsid assembly. This is even more important when the objective is the application of norovirus capsid proteins to a biotechnological context. Human norovirus (HuNoV) is the largest driver of viral gastroenteritis globally, threatening mostly children, the elderly and immunocompromised people. A vaccine has yet to be developed and the few available cell culture systems are still difficult to manage and unreliable. Hence, virus-like proteins (VLPs) have become the prime method for investigation.
These VLPs are produced by expressing the major capsid protein VP1 in insect cell culture and mimic the original virions in spherical form and icosahedral symmetry. We utilize native mass spectrometry as our main tool for probing the weight and thus state of assembly of these VLPs. Previous work from our lab determined that the capsid assembly is highly strain dependent, which is mirrored in their distinct stability profiles. Further probing by Pogan and colleagues revealed an expression batch specific N-terminal truncation. This truncation, revealed with charge detection mass spectrometry, gas phase electrophoretic mobility analysis and proteomics, seemingly coincided with an exclusive formation of smaller virus particles.This phenomenon was observed across different strains, pointing to a conserved mechanism.

Exploring AAV subspecies using single molecule mass detection

Eduard H.T.M. Ebberink[1], Alisa Ruisinger[2], Markus Neubel[2], Marco Thomann[2], **Albert J.R. Heck[1]**
1 Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Padualaan 8, 3584 CH Utrecht, the Netherlands
2 Roche Diagnostics, Penzberg, Germany

Introduction: Adeno-associated viruses (AAV) have proven to be the most promising vehicles for gene therapy. Because of their stability, low immunogenicity and non-pathogenicity, AAVs have ideal attributes as gene vector and delivery system. Unfortunately, disparity in AAV sample preparations (*e.g.* in capsid protein composition, DNA packaging and impurities) following large-scale production gives rise to AAV heterogeneity, with as of yet an unknown effect on gene delivery. Therefore, accurate analysis of AAV composition is important. Moreover, in a therapeutic context, to have complete control of AAV structure and its genetic payload a robust, quick and low material consuming method would be desirable.
Methods: Here we explore novel, analytical methods to accurately measure AAV molecular weight and unravel capsid structure, DNA packaging and sample purity. Single-particle techniques such as charge-detection mass spectrometry (CDMS) and mass photometry (MP) were performed side-by-side to allow for direct comparison. Two different serotypes (AAV8 and AAV2) were evaluated, either lacking DNA content or with a CMV-GFP transgene. In addition, we were able to compare AAVs from different vendors that utilize different production and purification workflows.
Preliminary data: Subtle differences between the AAV sets could be characterized by MP and CD-MS. We detected different molecular weights of the viral capsids of the same serotype but from different suppliers. Moreover, once a transgene was introduced, next to the expected empty or single genome packed capsids, a third population was measured yielding extra mass (potentially a double packed genome). Strikingly, the CD-MS measurements showed good agreement with the MP data. This work demonstrates that both techniques can be valuable tools in studying AAVs for gene therapy.
Novel aspect: Side-by-side comparison of two single molecule mass analysis techniques to assess structural integrity of adeno-associated virus gene-delivery vehicles.

Fixing cells prior to XL-MS for improved sampling of the spatial proteome
David Schriemer[1,2]
1 Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, Alberta T2N-4N1, Canada
2 Department of Chemistry, University of Calgary, Calgary, Alberta T2N-4N1, Canada
Introduction: XL-MS can be used to map protein-protein interactions (PPIs) *in situ*, ideally capturing them in their native environment and reflecting the organizational properties of the proteome. However, the nature of the cellular medium hinders the efficiency of the crosslinking experiment and limits our ability to deeply sample PPI space. The dynamics of the proteome during long crosslinker incubation times promotes reagent scavenging and distortion of the spatial proteome. Ideally, the proteome should be stabilized prior to crosslinking. Fixation strategies have been used successfully in microscopy for many years. We present two concepts for integrating cell fixation with XL-MS, one based on cryogenic methods and the other on classical formalin fixation, both designed to improve the utility of *in situ* XL-MS.
Methods: Our apparatus for cryogenic crosslinking includes a spray-freezer of cell suspensions, a specially designed unit for desiccating cells with -80°C acetone, and an adapted autosampler/pump system for freeze-substituting protein-reactive compounds at ultralow temperatures. The apparatus also includes a programmable-temperature chamber to control the rates of chemical reaction. For chemical fixation, cells are treated with 0-4% formaldehyde, washed, and then treated with protein-reactive reagents. Cells are imaged by wide-field fluorescence microscopy, using fluorescent markers for various cell structures. Standard 1D and 2D proteomics methods were used after labeling, with an Eclipse Mass Spectrometer fronted by a Vanquish Neo nanoLC C18 separation system. Data were analyzed with MSFragger for proteome coverage and monolink detection, and by CRIMP 2.0 for crosslink detection.
Preliminary data: The cryofixation process was optimized to ensure that freeze-substitution would not disturb cellular ultrastructure. For example, we show the expected phalloidin staining patterns of the actin cytoskeleton, which demands preservation of actin's 3D structure, and the expected DAPI staining patterns of an undistorted nucleus. We then surveyed an extensive list of amine-selective chemistries (in "monovalent" form) to determine the foundation for an effective crosslinker for low-temperature labeling. Ortho-phthaldehydes demonstrate the best balance of selectivity, water resistance (for ease of handling) and reactivity at low temperatures. For example, we show that greater than 15% of all cellular lysines could be labeled at -40°C with simple "monovalent" ortho-phthaldehydes. We then synthesized a series of di-ortho-phthalaldehydes (DOPAs), designed for high solubility in cold acetone and effective interaction sampling, and tested them on frozen cells. Data analysis of "DOPA4" through an initial shotgun LC-MS/MS experiment from E. coli cells demonstrated

extensive crosslinking. Parallel control XL-MS experiments using the DSS crosslinker in a conventional *in situ* manner demonstrated fewer CSMs. Digest fractionation and experiments with enrichable DOPAs are ongoing. An alternative fixation procedure involves formaldehyde pre-treatment, long demonstrated to preserve cell structure. Surprisingly, we show that formaldehyde-fixed cells support extensive monovalent chemical labeling of lysines. Labeling yields actually increase, plateauing between 1 and 4% formaldehyde pretreatment (a conventional fixation recipe). Amine-based crosslinkers work equally well when applied to formalin-fixed cells, generating high quality CSMs. In the talk, we will demonstrate how fixation allows for alternative crosslinker reaction schemes, including membrane permeabilization to introduce normally non-permeable reagents, and multiple applications of reagent to boost signal, to achieve full control over the crosslinking reaction.
Novel aspect: Prefixing cells preserves cell structure, restoring organizational fidelity to XL-MS experiments and expanding the potential of the technique.

Isolation and characterisation of naturally occurring cross-linked collagen peptides using mass spectrometry

Catherine Gilbert[1], Katell Bathany[1], Benoit Bertrand[2], Valery Hedouin[2], Caroline Tokarski[1]
1 CBMN, UMR CNRS 5248, Proteome Platform, University of Bordeaux, F-33000 Bordeaux, France
2 Univ. Lille, CHU Lille, ULR 7367 – UTML&A – Forensic Taphonomy & Anatomy, F-59000 Lille, France

Collagen, the most abundant extra-cellular matrix protein in the body, significantly contributes to the high mechanical strength of many tissues, due to enzymatically controlled cross-linking of the collagen network. This can be a key biomarker in many biological processes, such as cancer, gestation, and osteogenesis imperfecta. We are interested in the analysis of collagen cross-links to identify biomarkers to determine the post-mortem interval, focusing on the structural analysis of collagen cross-linking in forensic bones using optimized proteomic workflows.
To characterise cross-linked collagen peptides with MS, there are significant challenges, primarily due to the low abundance and structural heterogeneity of naturally occurring cross-links, but also due to the lack of bioinformatic support for the proteomic analysis of the mature trivalent cross-links in collagen. We present a multi-angle analysis strategy to characterise the different stages of collagen cross-linking. This include the analysis of post-translational modifications (PTMs) that are known to be associated with cross-linking, the analysis of divalent cross-links using currently available bioinformatic tools, and the targeted, multi-stage enrichment of trivalent cross-links.
For the optimal extraction of cross-linked collagen peptides, an optimised protocol has been developed efficiently solubilising collagenous tissue, through a multi-stage extraction protocol targeting insoluble collagen using chemical digestion.
For the analysis of divalent, immature cross-links (eg. Ketonorleucines), a targeted reduction was performed in order to stabilise the cross-link structure before proteomic sample preparation, followed by fractionation of the resulting digest. These fractions were then analysed using nanoLC-MS/MS (Orbitrap, Thermo Fisher Scientific), data were searched for cross-links using XlinkX (Thermo Fischer Scientific), and cross-link identifications were manually confirmed and correlated with PTM identification to identify potential cross-link sites.
We focus on the structural analysis of trivalent, mature cross-links (eg. pyridinolines); as these are more likely to persist in forensic samples, in comparison to the chemically unstable immature cross-links. To isolate naturally occurring, trivalent, cross-linked peptides, a multi-chromatographic separation was used to enrich cross-linked collagen peptides based on fluorescence activity. These fractions were analysed using nanoLC-MS/MS (Orbitrap), with targeted middle-down proteomics methodologies combined with MS3 fragmentation methods to fully characterise isolated peptides. With the low concentration of naturally-occurring cross-linked peptides in biological tissues, as well as the structural heterogeneity of the cross-links and of the highly modified collagen protein itself, a multi-level enrichment procedure was necessary to confidently identify and characterise such structures using targeted MS analysis.

Integrative structural biology of protein-RNA complexes using cross-linking mass spectrometry

Giacomo Padroni[1], Maria Bikaki[2], Mihajlo Novakovic[1], Antje C. Wolter[1], Simon H. Rüdisser[3], Alvar D. Gossert[3], Frederic H.-T. Allain[1], **Alexander Leitner[2]**

1 Department of Biology, Institute of Biochemistry, ETH Zürich, Zurich, Switzerland

2 Department of Biology, Institute of Molecular Systems Biology, ETH Zürich, Zurich, Switzerland

3 Biomolecular NMR Spectroscopy Platform, ETH Zürich, Zurich, Switzerland

Introduction: Cross-linking approaches that connect reactive sites in protein complexes, organelles and even entire cells constitute some of the central techniques in structural proteomics. The expansion of cross-linking to probe interactions of different classes of biomolecules (for example, of proteins with oligonucleotides, lipids and carbohydrates) may be considered as the next frontier. Using stable isotope labeling of RNA, we have previously shown that protein-RNA binding interfaces can be characterized at up to single amino acid and nucleotide resolution. In this work, we show how such information can be combined with other experimental and computational methods to unravel interactions between the nucleocapsid protein of SARS-CoV-2 and part of its viral genome, the s2m element.

Methods: Protein-RNA cross-linking was induced by irradiation with 254 nm UV light. Cross-linked complexes were processed using standard workflows, including digestion with RNases A and T1 and trypsin as the protease. Peptide-RNA conjugates were enriched by titanium dioxide affinity chromatography and enriched samples were analyzed by data-dependent LC-MS/MS on a Thermo Orbitrap Fusion Lumos instrument. Data was analyzed by xQuest/RNxQuest. NMR spectroscopy was performed on Bruker instruments of different field strengths equipped with cryoprobes. Docking was performed using HADDOCK 2.4.

Preliminary data: The s2m element has previously been shown to interact with the nucleocapsid (N) protein of SARS-CoV-1 and -2. It is presumed that this interaction plays important roles in viral packaging and replication. To obtain a model of the N-s2m interaction, we followed an approach that combines NMR spectroscopy and CLIR-MS (cross-linking of stable isotope labeled RNA coupled to mass spectrometry). NMR and CLIR-MS data were generated from complexes of the two RNA-binding domains of the N protein, the N-terminal and the C-terminal domain (NTD and CTD, respectively). Restraints obtained from NMR (through chemical shifts) and MS (through cross-links) proved highly complementary and mutually supportive and allowed to build models for the two complexes using computational docking. In a second step, we used the obtained models to find small molecule scaffolds that can bind to either of the complex members in their interaction regions as determined from the integrative models (fragment screening). Using a small, fluorine-labeled compound library, low-affinity binders were discovered for all targets (N-NTD, N-CTD, s2m) using ^{19}F-NMR. Such compounds could be further optimized using lead optimization strategies, although this was not attempted here. Our integrative workflow highlights the combination of different experimental techniques, including cross-linking MS, for the structural characterization of protein-RNA complexes. Beyond the illustrated example with a potential application in drug discovery, we are currently also exploring the value of this approach for different applications in biophysics and systems biology.

Novel aspect: Combination of protein-RNA cross-linking data with other experimental and computational methods in a drug discovery context

RNA/DNA-protein crosslinking mass spectrometry – Applications of UV and chemical crosslinking and data analysis with NuXL

Luisa M. Welp[1,2], Alexander Chernev[1], Alexander Wulf[1], Fanni Bazso[1], Timo Sachsenberg[3,4], Yehor Horokhovskyi[5], Monika Raabe[1], Arslan Siraj[3], Aditi Sharma[6], Yi He[7], Bernard Delanghe[6], Rosa Viner[7], Juliane Liepe[5], Oliver Kohlbacher[3,4,8], **Henning Urlaub[1,2]**

1 Bioanalytical Mass Spectrometry Group, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany

2 Bioanalytics Group, Institute of Clinical Chemistry, University Medical Center Göttingen, 37075 Göttingen, Germany

3 Institute of Bioinformatics and Medical Informatics, University of Tübingen, 72076 Tübingen, Germany

4 Applied Bioinformatics, Dept. of Computer Science, University of Tübingen, 72076 Tübingen, Germany

5 Quantitative and Systems Biology; Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany

6 Thermo Fisher Scientific (Bremen) GmbH, 28199 Bremen, Germany
7 Thermo Fisher Scientific, San Jose, USA 8 Institute of Translational Bioinformatics, University Hospital Tübingen, 72076 Tübingen, Germany
Introduction: Crosslinking mass spectrometry (XL-MS) can resolve sites of interaction in protein-RNA and DNA complexes at single amino acid and nucleic acid resolution. Processing of (large) XL-MS data, reliable assignment, and visualization of crosslinking results – including annotation of spectra – remain key challenges in the database search of crosslinked species. We have developed NuXL, a database search engine, which robustly identifies UV and chemically induced crosslink sites within peptide sequences from MS2 spectra, and significantly improves data-processing and identification of protein–RNA/DNA crosslinks, irrespective of source (i.e., isolated complexes or cellular entities). NuXL not only supports all types of UV crosslinking with non-substituted and substituted nucleotides but also the use of various chemical crosslinking reagents including formaldehyde.
Methods: We used UV and chemical XL-MS to investigate protein-RNA/DNA interaction sites/regions in E. coli and analyzed the data with NuXL. For this, crosslinked E. coli cells were processed by nuclease and endoproteinase digestions and crosslinked peptides were enriched using TiO2 or SILICA-based chromatography. Finally, enriched crosslinked peptide-RNA/DNA (oligo)nucleotides were separated by off-line chromatography followed by LC-MS analysis. MS data were analyzed with NuXL, using different presets for the applied crosslinkers. Crosslink spectrum matches (CSMs) were rescored by percolator and filtered for 1% FDR. Resulting crosslink sites were mapped to available protein-DNA/RNA complex 3D structures and located to conserved domains within the crosslinked proteins.
Results: We developed NuXL, a dedicated software package for the analysis of XL-MS data obtained from UV and chemically crosslinked protein–RNA/DNA samples. NuXL is available in the OpenMS platform and as a node in Proteome Discoverer (PD). It allows for reliable, FDR-controlled assignment of protein–nucleic acid crosslinking sites from samples treated with UV light or chemical crosslinkers and offers user-friendly matched spectra visualization including ion annotations. We demonstrate the robust and sensitive performance of NuXL on MS data acquired from UV and chemically crosslinked cells. With respect to the different crosslinking approaches, i.e. UV and chemical crosslinking, we show that C, G, F, H, K, Y and C, K, H, M are the most prominent crosslinked amino acids, respectively, and that chemical XL-MS almost exclusively identifies crosslinks to purines, whereas UV XL-MS identifies pyrimidine crosslinks in both RNA and DNA-protein crosslinking experiments. In E. coli cells, we identified in total more than 4800 unique crosslinked peptides of more than 1490 proteins crosslinked to RNA. Proteins found in all crosslinking approaches are enriched for RNA-binding proteins, while proteins that are specifically identified through chemical crosslinking are enriched for proteins with oxidoreductase activity, small molecule binding proteins, and proteins that have been mainly categorized to participate in protein-protein interactions. Importantly, the crosslinked peptides/amino acids that we identified through NuXL are located within or in close neighborhood to the conserved domains/sites of these proteins. In summary, NuXL provides rapid, robust, sensitive, and reliable data annotation of protein-RNA and DNA crosslinking sites irrespectively of sample complexity.
Novel aspect: NuXL search tool for XL-MS data annotation form protein-RNA and DNA crosslinking with UV and chemicals applicable on in vivo crosslinking.

Protein Complex Discovery in *Bacillus Subtilis* by AI-assisted Structural Proteomics
Francis O'Reilly
National Cancer Institute
Experimental data enable a candidate-based approach to systematically model novel protein assemblies using AlphaFold-Multimer. We use a combination of in-cell crosslinking mass spectrometry and co-fractionation mass spectrometry (CoFrac-MS) to identify protein–protein interactions in the model Gram-positive bacterium Bacillus subtilis. We show that crosslinking interactions prior to cell lysis reveals protein interactions that are often lost upon cell lysis. We predict the structures of these protein interactions and others in the SubtiWiki database with AlphaFold-Multimer and, after controlling for the false-positive rate of the predictions, we propose novel structural models of 153 dimeric and 14 trimeric protein assemblies. Crosslinking MS data independently validates the AlphaFold predictions and scoring. 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.

Large scale cross-linking to discover, evaluate, and validate the human cell structural proteome and interactome

Tara K. Bartolec[1,2], Xabier Vázquez-Campos[1], Alexander Norman[3], Clement Luong[4], Marcus A. Johnson[4], Richard J. Payne[3,5], Marc R. Wilkins[1], Joel P. Mackay[4], Jason K. K. Low[4]
1 Systems Biology Initiative, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Randwick, NSW, Australia
2 Genome Biology Unit, European Molecular Biology Laboratory, 69117, Heidelberg, Germany
3 School of Chemistry, University of Sydney, Sydney, NSW 2006, Australia
4 School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia
5 Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, NSW 2006, Australia

Introduction: Recent advances in structural biology have expanded our ability to create experimental structures for proteins and complexes. However, many proteins remain refractory to these approaches or have not yet been analysed. Machine-learning based structure predictors have enabled access to highly accurate protein structure models for entire proteomes. These modellers are trained on experimental structures available in the Protein Data Bank (PDB), which constitute a relatively small subset of proteins, with many structures solved using non-native conditions or sequences. Therefore, a critical question is whether predicted (and PDB) structures reflect the *bona fide* structures and complexes formed by proteins in their native environment. We investigate this question using a large-scale cross-linking mass spectrometry (XL-MS) resource generated for the human cell.

Methods: To generate a high density and high depth XL-MS dataset for human HEK293 cells, we utilised a multipronged approach. Briefly, we cross-linked four subcellular fractions (nucleus, endoplasmic reticulum, mitochondria and cytosol) using three different cross-linkers with orthogonal chemistries (DHSO, DSSO, DMTMM). Then, we enriched cross-linked peptides using offline size-exclusion chromatography followed by further fractionation by high pH reverse phase HPLC. Mass spectrometry was performed on concatentaed fractions using hybrid-MS2-MS3, or MS/MS with EThcD or HCD, fragmentation strategies. Cross-linked peptides were identified using XlinkX 2.3 or pLink2, using stringent search parameters and post-hoc filtering to control the false discovery rate to <2% at the unique residue pair (URP) or PPI levels.

Preliminary data: Our study has generated the most comprehensive XL-MS dataset reported to date for any species, with 28,910 URPs representing 4,084 unique proteins and 2,110 unique putative PPIs. The use of a subcellular fractionation strategy before cross-linking resulted in significantly improved proteome coverage, whilst orthogonal reactivities (D/E-D/E, K-K and K-D/E) improved the density of cross-linking per protein, especially in intra-protein links. We demonstrate that our resource of URPs confirm and rediscover existing experimental structures, capturing proteoforms and complexes within their approximate subcellular niches and range of conformations. Remarkably, our intra-molecular URPs also largely corroborate thousands of new structures predicted by next-generation modeller AlphaFold2, including those involving proteins (or regions of proteins) without existing resolution, and those lacking any structural precedent. Furthermore, our inter-protein crosslinks recapture the topology of well-described complexes and PPIs, whilst supporting or discovering poorly characterised PPIs. Critically, the inter-protein crosslinks also help localise PPI interfaces, and we use this information to assess quaternary protein structures modelled in AlphaFold-Multimer.

Novel aspect: We have experimentally corroborated thousands of experimental (*in vitro*) or predicted structures for proteins and PPIs in the human cell.

Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning: AlphaLink

Andrea Graziadei[1,3]*, Kolja Stahl[2]*, Therese Dau[3], Oliver Brock[2,5], Juri Rappsilber[3,6]
1 Fondazione Human Technopole, 20157 Milan, Italy
2 Technische Universität Berlin, Robotics and Biology Laboratory, 10623 Berlin, Germany
3 Fritz Lipmann Institute, Leibniz Institute on Aging, 07745 Jena, Germany
4 Technische Universität Berlin, Chair of Bioanalytics, 10623 Berlin, Germany
5 Si-M/"Der Simulierte Mensch", a Science Framework of Technische Universität Berlin and Charité – Universitätsmedizin Berlin, Berlin, Germany.
6 Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3BF, UK
*These authors contributed equally.

While AlphaFold2 can predict accurate protein structures from the primary sequence, challenges remain for proteins that undergo conformational changes or for which few homologous sequences are known. Here, we introduce AlphaLink, a modified version of the AlphaFold2 algorithm that incorporates experimental distance restraint information from crosslinking mass spectrometry into its network architecture. By employing sparse experimental contacts as anchor points, AlphaLink improves on the performance of AlphaFold2 in predicting challenging targets. We confirm this experimentally by using the noncanonical amino acid Photo-Leucine to obtain information on residue-residue contacts inside cells by crosslinking mass spectrometry. The program can predict distinct conformations of proteins based on the distance restraints provided, demonstrating the value of experimental data in driving protein structure prediction. The noise-tolerant framework for integrating data in protein structure prediction presented here opens a path to accurate characterization of protein structures from in-cell data.

Integrative Structure Modeling in the Age of Deep Learning

Dina Schneidman
The Hebrew University of Jerusalem, Israel

Deep learning models like AlphaFold2 and RoseTTAFold enable high-accuracy protein structure prediction. However, large protein complexes are still challenging to predict due to their size and the complexity of interactions between multiple subunits. Integrative structure modeling is often used to characterize structures and dynamics of large macromolecular assemblies by combining various types of input information, such as available protein structures and models, cross-linking mass spectrometry, cryo-electron microscopy, and small-angle x-ray scattering. Recent progress in protein folding enabled by deep learning has improved structural coverage for domains, and even protein-protein interactions, which are essential inputs for integrative structure modeling. I will present CombFold, a hierarchical and combinatorial assembly algorithm for the prediction of structures of large protein complexes utilizing pairwise interactions between subunits predicted by AlphaFold2. We test the method on a benchmark of large heteromeric assemblies (up to 30 chains and 18,000 amino acids) and obtain a success rate of ~70%. Distance restraints, based on crosslinking mass spectrometry, can guide the assembly. Moreover, we design a deep learning model for predicting the optimal distance range for a crosslinked residue pair based on the structures of their neighborhoods. These tools will be useful in expanding structural coverage beyond monomeric proteins.

PIPENN: protein interface prediction from sequence with an ensemble of neural nets

Bas Stringer, Hans de Ferrante, Sanne Abeln, Jaap Heringa, **K. Anton Feenstra**, and Reza Haydarlou

Introduction: Understanding protein function is greatly helped by understanding how and where a protein interacts with other molecules, and predicted binding interfaces allow experiments to be focused on most promising regions. Here, we tackle the problem of predicting from sequence the binding interfaces of proteins for interacting with other proteins (PPI), DNA/RNA (nucleic acids) or small molecules. Many computational and machine learning approaches have been developed over the years to predict such interface residues from sequence [1-4]. However, the effectiveness of different deep learning architectures and learning strategies for protein interface prediction, has not yet been investigated in great detail. Therefore, we here develop an extensive dataset for Deep Learning (DL), dubbed BioDL, and explore six DL architectures and various learning strategies with sequence-derived input features for the prediction of protein interfaces.

Methods: We introduce a new and extensive dataset 'BioDL' for benchmarking, which comprises protein–protein interactions from the PDB and DNA/RNA and small molecule interactions from the BioLip database, yielding BioDL_P, BioDL_N and BioDL_S, respectively, as well as the generic BioDL_A which encompasses all three interactions. We use BioDL to train and test our deep learning predictors,and assess significance of differences observed throughout so solid conclusions may be drawn. In addition, we use previously developed homo/heterometic (hhc) datasets from our group [1, 2] and the ZK448 from Zhang & Kurgan. [4] We apply six major deep learning architectures and an ensemble predictor, dubbed PIPENN, which we implemented as an additional layer on top of the other six architectures.

Preliminary data: We first explored a number of variations in building blocks for the deep learning models. The best performing combination in the dnet architecture yields an AUC-ROC of 0.733, and uses HeUniform kernel intialization, CrossEntropy loss function, 1-dimensional spatial form, PRELU activation function, no MaxPooling, and 1-Hot sequence encoding. Changing kernel initialization to GlorotNormal, and loss functions to MeanSquaredError does not significantly impact accuracy (<0.005

drop in AUC). Removing batch normalization, padding or dropout have a significant and large impact (>0.02 drop in AUC). The six individial architectures yield AUC-ROC prediction accuracies from 0.717 to 0.730, when trained and tested on the BioDL_P dataset. Excitingly, the PIPENN ensemble predictor for PPI significantly outperforms each of, reaching an AUC-ROC of 0.755 for protein-protein interface prediction. Analysis of feature importance using SHAP values, show length as highest as we observed previously, [1-3] as well as amino acid type (AA) and profile scores (pssm). We trained a generic predictor on the BioDL dataset for each of the three specific interaction types: PPI (p), protein–small molecule (s), and protein–DNA/RNA (n), as well as a generic predictors for any of the three types (a). In all cases, the specific predictors perform better than the generic predictor. For PPI, the specific predictor reaches AUC-ROC of 0.755, while the generic obtains 0.733. For small molecules this was respectively 0.864 and 0.826, and for nucleic acids 0.894 and 0.835. Using separate independent test sets for PPI prediction, hhc and ZK448, we show that PIPENN significantly outperforms all currently published methods for sequence-based prediction of the binding interface of proteins with other proteins, with AUC-ROC of 0.769 for hhc and 0.729 for ZK448. Other methods tested on ZK448 do not reach beyond 0.715.
Novel aspect: Accurately predict protein interaction interfaces from protein sequence, needs neither experimental nor structural information, thus easily applied to any protein.

XLEC: Large-scale prediction and modeling of protein-protein interaction using sequence co-evolution and cross-linking data
Hadeer Elhabashy[a,b,c*], Eugen Netz[b,c], Oliver Kohlbacher[b,c,d,*]
a Department of Protein Evolution, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, 72076 Tübingen, Germany
b Applied Bioinformatics, Department of Computer Science, University of Tübingen, Sand 14, 72076 Tübingen, Germany
c Institute for Bioinformatics and Medical Informatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany
d Institute for Translational Bioinformatics, University Hospital Tübingen, Hoppe-Seyler-Str. 9, 72076 Tübingen, Germany
Prediction and structural modeling of protein-protein interactions (PPIs) are essential for understanding biological processes. Most large-scale experimental and computational approaches that predict PPIs do not provide structural information. We present a novel approach, XLEC, combining cross-linking mass spectrometry (XL-MS) and evolutionary coupling (EC) data for efficient proteome-wide prediction and modeling of PPIs. While EC derived from multiple sequence alignments primarily yield information on direct contacts between proteins across the interface, XL-MS data preferentially captures longer-range interactions, hence these methods contain complementary information. XLEC integrates information from both approaches in a machine learning-based model and subsequent restraint-based modeling of the complex structure. We applied XLEC to data from the murine mitochondrial proteome and compared its performance to those of XL-MS and EC separately. Our assessment suggests that the XLEC predictor outperforms those of XL-MS, and EC for PPI prediction (ROC-AUC: 0.73, 0.64, and 0.68, respectively). Furthermore, XLEC-based modeling of PPIs achieved excellent L-RMSD (<10 Å) for 16.4% of the benchmark dataset (EC only: 7%, XL-MS only: 2.9%;). Using XLEC, we generated more than 400 *de novo* PPI models revealing novel insights into the mitochondrial interactome including the interaction between the alanine aminotransferase 2 (ALAT2) and the hexokinase-1 (HXK1) and the interaction between the DNA topoisomerase I (TOP1M) and the electrogenic aspartate/glutamate antiporter (SLC25A13).

Nature Methods
Arunima Singh studied biochemistry as an undergraduate, but a keen interest in programming led her to pursue a master's degree in bioinformatics at the University of Pune, India. She went on to obtain her Ph.D. from the University of Georgia, USA, where she used molecular modeling to analyze protein-carbohydrate interactions and developed a web tool for generating 3D structures of glycosaminoglycans, while working with Robert Woods. She did postdoctoral research at New York University on computational characterization of protein-protein interfaces and peptide mimetic-based inhibitor design. She handles computational biology, glycobiology and structural biology content for the journal. Arunima joined *Nature Methods* in June 2019.

Tuesday

HX-MS provides structural insights into the roles of mutations in structurally disordered regions of cancer-relevant proteins

Malvina Papanastasiou[1], Amanda L. Waterbury[2], Hui Si Kwok[2], Ceejay Lee[2],Allyson M. Freedy[2], Cindy Su[2], Will Hawkins[1], Andrew Reiter[1], Samuel M. Hoenig[2], Michael E. Vinyard[2], Brian B. Liau[1,2] and Steven A. Carr[1]
1 Broad Institute, Cambridge, MA;
2 Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA

Introduction: The histone demethylase LSD1 has emerged as a promising target for the treatment of cancers, with numerous LSD1 inhibitors currently in clinical trials, particularly for AML. Recently, the Liau group employed CRISPR-suppressor scanning to identify single-guide RNAs (sgRNAs) that target LSD1 and can promote drug resistance to LSD1 inhibitors (Vinyard, Liau, 2019). As anticipated, the most highly enriched drug-enriched sgRNAs were found to target the active site. Unexpectedly, multiple sgRNAs were discovered to target the N-terminus IDR. Recent experiments using LSD1 mutants have demonstrated that deletions in the LSD1 IDR inhibit AML differentiation and further promote in vitro condensate formation. To provide structural insights into the role of LSD1 IDR that eludes from high resolution structures, we employed HX-MS.

Methods: For HX-MS experiments, proteins were diluted in deuterium (5-fold) for various labeling periods and temperatures and were subsequently quenched with ice chilled guanidine hydrochloride. Samples were prepared either manually or using a Trajan Parallel Extended autosampler, with lipid filtration capabilities, coupled to a Thermo Q-Exactive HF Orbitrap MS. Samples were digested using a combination of proteases (Pepsin:Protease XIII, NovaBioassays). Generated peptides were desalted using a C18 trap (Acclaim PepMap, Thermo) and separated on a Hypersil Gold C18 analytical column (Thermo) employing a 18 min gradient. Deuterium content at the peptide level was measured in HDExaminer (Sierra Analytics) and bimodal distributions were analyzed using HXExpress v3.

Preliminary data: For HX-MS experiments we utilized full-length wild-type LSD1/CoREST, two mutants with short amino acid deletions in their IDR (NΔ1, NΔ2), and a mutant lacking the N-terminus IDR NΔ(1-170). More than 800 distinct peptides were identified for LSD1, including 35 peptides spanning the N-terminus IDR, and more than 300 peptides for CoREST, resulting in complete sequence coverage for both proteins. We performed HX-MS at two temperatures (0 ^{o}C and 30 ^{o}C) to expand the sampling window across many orders of magnitude and capture potential differences between wild-type and mutants in rapidly exchanging amide hydrogens in the N-terminus IDR and in slowly exchanging amide hydrogens in the folded SWIRM, amine-oxidase (AOD), and Tower domains. Differences in D-uptake between wild-type and mutants were identified in peptides encompassing the entire IDR sequence, the AOD, and the intrinsically disordered C-terminus, only at time scales faster than 10 sec and 0 ^{o}C. Along the entire length of the IDR, extensive bimodal distributions were observed, pointing to an ensemble of interconverting conformers that transiently populate folded states. Our findings demonstrate that N-terminus deletions enhance the dynamics and accessibility of LSD1 by shifting the IDR conformational landscape towards an ‘open state’. Our results corroborate recent findings demonstrating that these mutants promote in vitro condensate formation, paradoxically, not via IDR-IDR interactions, but likely as a result of increased interactions with newly exposed epitopes in folded regions of the LSD1/CoREST.

Novel aspect: HX-MS offers novel insights into the role of LSD1 IDR, which eludes high resolution structures, in condensate formation in vitro.

An HIV-1 Broadly Neutralizing Antibody Overcomes Structural and Dynamic Variation through Highly Focused Epitope Targeting

Edgar A. Hodge[1], Ananya Chatterjee[3], Chengbo Chen[1,2], Meghan McGrath[1], Gajendra S. Naika[1], Mint Laohajaratsang[1], David Montefiori[4], Vidya Mangala Prasad[1,3], **Kelly K. Lee[1,2*]**
1. Department of Medicinal Chemistry, University of Washington, Seattle, WA 98195, USA.
2. Biological Physics, Structure and Design Graduate Program, University of Washington, Seattle, WA 98195, USA
3. Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka 560012, India.
4. Laboratory for AIDS Vaccine Research & Development; Department of Surgery; Duke University Medical Center, Durham, NC USA

Introduction: The HIV-1 envelope glycoprotein (Env) mediates receptor binding and membrane fusion to initiate infection of a host cell. Env is the sole target for neutralizing antibodies and is under intense

immune pressures, causing it to evolve and diversify into the most highly variable part of the virus. While recent studies have reported high-resolution structures for Env, however they represent static, platonic ideals, and structural and functional implications of Env variation remain poorly understood. Indeed, under native conditions, Env dynamically flickers between antigenically and functionally distinct conformational states. Probing the impact of these dynamic traits on Env-antibody recognition is essential for understanding the underlying basis for differences in antigenic and neutralization profiles among HIV-1 isolates.
Methods: We used hydrogen/deuterium-exchange mass spectrometry (HDX-MS) and quantitative measurements of antibody binding kinetics by biolayer interferometry to investigate HIV Env structural dynamic variation and to measure the effect of epitope dynamics on antibody binding ability. Single-particle cryo-EM was also used to image the 3-dimensional structure of a broadly cross-reactive antibody that can neutralize diverse HIV-1 isolates (bnAb) in complex with two highly divergent Env variants.
Preliminary data: The existence of remarkable bnAbs against HIV-1 has been appreciated for more than a decade. Understanding how these antibodies grapple with variability in their targets across diverse viral isolates however lags, as this requires first characterizing the extent of structural and antigenic variation embodied in Env and then identifying how a bnAb overcomes that variation at a structural level. HDX-MS and BLI experiments show that variation in structural ordering and local dynamics in the V1/V2 apex of Env across a globally representative panel of HIV-1 isolates has a marked effect on antibody association rates and affinities. We also report cryo-EM structures of an apex-targeting bnAb bound to two divergent Env that exhibit different degrees of structural dynamics throughout their trimer structures. Parallel HDX-MS experiments demonstrate that this bnAb has an exquisitely focused binding footprint at the trimer apex where binding did not result in allosteric changes throughout the rest of the structure. These results demonstrate that structural dynamics are a cryptic determinant of antigenicity, and mature antibodies that have achieved breadth and potency in some cases achieve broad cross-reactivity by “threading the needle” and binding in a highly focused fashion.
Novel aspect: Structural mass spectrometry combined with cryo-EM elucidates how a bnAb overcomes antigenic variation in HIV-1, providing guidance for vaccine design.

HX-MS^2 : Automating Data Validation and Expanding the Scope of Deuteration Analysis
Frantisek Filandr, Vladimir Sarpe, Alex Crowder, Morgan Khan, David Schriemer
University of Calgary, Calgary, AB, Canada
Introduction: Hydrogen/deuterium exchange mass spectrometry (HX-MS) has been performed using MS1-based measurements since the 1990’s, which restricts it to the detection of high quality and well-isolated isotopic envelopes. There are two problems with this conventional mode of operation. First, the conventional mode restricts applications to protein systems of limited complexity as peptide attrition rates scale with protein size. Second, manual validation of peptide selections is very time consuming and prone to operator error. It is possible to obtain deuteration values from fully scrambled and mass[1]resolved CID fragment ions collected in MS2 mode. However, a robust implementation of this finding is still not available. Here, we show that data independent acquisition (DIA) modes can create an automated “2D” HX-MS method.
Methods: A microflow based HX system and an HDX-PAL autosampler were connected to Sciex Triple TOF 6600 and Thermo Scientific Q Exactive Plus to analyze deuterated phosphorylase B in DIA mode, using standard digestion and quench methods. To analyze the deuteration data resulting from DIA-based acquisitions, a new module was developed for the Mass Spec Studio 2.0. Undeuterated samples were first analyzed in Data Dependent Acquisition (DDA) mode to generate a library of peptide IDs and retention times and used in the analysis of deuterated samples analyzed in DIA mode. Here, we collected kinetics data on the native protein and replicate differential HX data to compare the native and perturbed states. Optimized workflows were developed and applied to several other protein systems.
Preliminary data: We completed a build of AutoHX, a new software package that measures deuterium incorporation from MS1 and MS2 domain data, collected using DIA-based data acquisition systems. The software is built as a plug-in to the Mass Spec Studio and requires the input of peptide features. AutoHX aligns chromatograms for peptide and fragment ions, selects CID fragments for deuteration calculation based on computer vision algorithms, and chooses between MS1 or MS2 derived deuteration values informed by the quality of the underlying identification. We collected HX-DIA data on phosphorylase b, a ~100kDa protein, for the two platforms. Both generated many CID

fragments for the majority of peptides detected in the digest. The deuteration precision derived from the fragment data was found to be a function of the number of fragments used in the calculation. We obtain excellent sequence coverage, redundancy, and low deuteration error even with the selection of as few as three fragments. Deuterium uptake curves from a kinetics experiment show that deuteration values from “clean” MS1 signals are virtually identical to MS2-derived values. When differences exist between MS1 and MS2 analyses (due to spectral overlap for example) we show that autovalidated MS2 fragment data can be used to override MS1 data and boost sequence coverage. We also show that deuteration differences between different protein states can be faithfully represented by MS2 domain data. The impact on automation is considerable. We show that full multifile projects can be processed with little to no user intervention, at the click of a button. Any intervention is based solely on selecting filters for the underlying data, creating a verifiable link between data quality metrics and the final deuteration report. DIA-based HX-MS2 is the routine data collection mode in our lab now, and other examples of projects will be shown for illustration.
Novel aspect: A fully automated “2D” HX-MS method for complex mixture analysis based on DIA technology.

Insights Into the Proton Translocation Pathway of the Organohalide Respiratory Complex of Dehalococcoides mccartyi Strain CBDB1
Nadine Hellmold[1], Marie Eberwein[1], Darja Deobald[1], Lorenz Adrian[1,2]
1 Helmholtz Centre for Environmental Research – UFZ, Permoserstraße 15, Leipzig, Germany
2 Institute of Biotechnology, Chair of Geobiotechnology, Technische Universität Berlin, Berlin, Germany
Introduction: Dehalococcoides mccartyi CBDB1, a strictly anaerobic bacterium of the Chloroflexi phylum uses hydrogen as sole electron donor and different organohalides as terminal electron acceptors in a process called organohalide respiration (OHR). Dehalogenation is catalyzed by a corrinoid-dependent reductive dehalogenase (RdhA) that is part of a membrane-bound OHR protein complex, composed of at least seven proteins: RdhA and its anchor protein RdhB, hydrogenase subunits HupL and HupS, ferredoxin-like protein HupX, OmeA, and the membrane-bound subunit OmeB. Notably, the OHR complex lacks quinones, cytochromes and proton pumps. Here, we developed an in-vitro enzyme activity assay using deuterium-labeled water (D2O) and methyl viologen as artificial electron donor allowing to spatially track the incorporation of protons into the dehalogenation product.
Methods: *In-vitro* dehalogenase activity assays were conducted anaerobically with CBDB1 cells cultivated in H2O- or D2O-containing medium, and reaction mixes with D2O or H2O, to have either D2O outside and H2O inside the cells, or vice versa. We optimized the activity assay and termination procedure to ensure accurate quantification of dehalogenation products by GC-MS before diffusion equilibrium across the membrane was reached. Then, the ‘deuterium degree’ was determined describing the relative proportion of deuterated products. To predict the proton path across the membrane, we performed a multiple sequence alignment of OmeB and homologous NrfD-like proteins from various bacteria using the MUSCLE algorithm in MEGA11 and calculated the structure of the OmeA/OmeB/HupX submodule and of RdhA using AlphaFold2 ColabFold.
Preliminary results: During the reductive dehalogenation reaction, the substrate is protonated. To analyze if the incorporated proton originates from the inside (cytoplasm) or outside (exoplasm) of the cell, we determined the deuterium degree in the resulting product. Our experimental data with D2O initially outside and H2O inside the cells, show an increase of the deuterium degree towards an equilibrium over time, indicating the initial incorporation of protons. The converse experiment with D2O inside and H2O outside the cells shows a decreasing deuterium degree over time. Although this second data set is less pronounced, it indicates initial incorporation of deuterium ions. Both experiments therefore suggest that protons are passed through the membrane directly onto the substrate. To identify a potential proton pathway through the protein complex, we analyzed the core structures of the complex. Structural and sequential analysis revealed that OmeB, a member of the NrfD-like protein group, contains several conserved, mostly charged amino acids. The protein structure of OmeB also features a prominent quinone binding site (Q-site), covered by the lower part of HupX, while OmeA represents the uppermost part of the subcomplex. The Fe-S clusters and molybdenum cofactors in HupX and OmeA are arranged linearly within the subcomplex, converging at the putative Q-site in OmeB. We manually docked RdhA to the OmeA/OmeB/HupX subcomplex, with its Fe-S cluster less than 10 Å distant from the membrane-nearest Fe-S cluster of HupX, in close proximity to the putative Q-site of OmeB. In the absence of quinones, RdhA and its substrate appear to take over

the quinone's function. Thus, the dehalogenation reaction contributes directly to the proton motive force by transferring protons from the cytoplasm to the substrate. Charged amino acids distributed in the membrane helices of OmeB may be involved in proton translocation, as suggested for other NrfD-like proteins
Novel aspect: Combining experimental and structural data we propose how quinone-independent protein-based organohalide respiration could couple electron flux with proton translocation.

Antigenic structure of the human coronavirus OC43 spike reveals exposed and occluded neutralizing epitopes
Chuyan Wang[1], Emma L. Hesketh[2], **Tatiana M. Shamorkina[3]**, Wentao Li[1,*], Peter J. Franken[1], Dubravka Drabek4,[5], Rien van Haperen4,[5], Sarah Townend[2], Frank J.M. van Kuppeveld[1], Frank Grosveld[4,5], Neil A. Ranson[2], Joost Snijder[3], Raooul J. de Groot[1], Daniel L. Hurdiss[1,#], Berend-Jan Bosch[1,#]
1. Virology Section, Infectious Diseases and Immunology Division, Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands
2. Astbury Centre Structural Molecular Biology, School Molecular and Cellular Biology, Faculty Biological Sciences, University of Leeds, Leeds, UK
3. Biomolecular Mass Spectrometry & Proteomics, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Utrecht, The Netherlands
4. Department of Cell Biology, Erasmus Medical Center, Rotterdam, The Netherlands
Harbour BioMed, Rotterdam, The Netherlands
Corresponding author
Human coronavirus OC43 is a globally circulating common cold virus sustained by recurrent reinfections. How it persists in the population and defies existing herd immunity is unknown. Here we focus on viral glycoprotein S, the target for neutralizing antibodies. In this work we studied the antigenic structure of the human coronavirus OC43 spike protein by characterizing the neutralizing capacity, binding sites and binding breadth of a panel of anti-OC43-S monoclonal antibodies using structural (cryo-EM and HDX-MS) and functional approaches.
We defined multiple vulnerable sites on the OC43 spike protein recognized by neutralizing monoclonal antibodies. Neutralizing antibodies are directed to the sialoglycan-receptor binding site in $S1_A$ domain, but, remarkably, also to $S1_B$. The latter block infection yet do not prevent sialoglycan binding. While two distinct neutralizing $S1_B$ epitopes are readily accessible in the prefusion S trimer, other sites are occluded such that their accessibility must be subject to conformational changes in S during cell-entry. While non-neutralizing antibodies were broadly reactive against a collection of natural OC43 variants, neutralizing antibodies generally displayed restricted binding breadth. Collectively, these results provide a structural basis for understanding humoral immunity and adaptive evolution for this endemic human coronavirus.

HDX-MS and SARS-CoV-2 spike protein: From epitope mapping to informing vaccine design
Joey Sheff, M.A. Rossotti, J. Tanha, J.F. Kelly, Y. Durocher, M. Stuible, J. Schrag
Human Health Therapeutics Research Centre, Life Sciences Division, National Research Council Canada, Ottawa, ON, Canada
Introduction: Declared a pandemic in March 2020 by the WHO, COVID-19, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), remains a severe global health and economic burden. The toll on public health has been exacerbated with the emergence of SARS-CoV-2 variants of concern (VoCs). Key to SARS-CoV-2 infection is its surface-displayed spike glycoprotein, a homotrimer where each protomer consists of S1 and S2 domains. S1 is further delineated by an N-terminal domain (NTD), a receptor-binding domain (RBD) and subdomains SD1 and SD2. The spike protein remains the antigen of choice for SARS-CoV-2 vaccines and biotherapeutics. We describe how hydrogen-exchange mass spectrometry (HDX-MS) platform was used to characterize behaviour of therapies emerging from NRC laboratories invoking the spike trimer.
Methods: A versatile bottom-up HDX-MS workflow, enabled by online electrochemical reduction and careful protease selection, was developed to probe the conformational behaviour of the spike protein. This enabled the analysis of an array of spike protein samples, including truncated RBD (76% sequence coverage), S1 (71%) and S2 (72%) domains, as well as the full length spike homotrimer ectodomain (75%). All data was collected using an HDx-3 Pal coupled to a UPLC-MS (Waters Synapt G2-Si) system. Deuteration was assigned with MS Studio, and in all cases significant changes were

assigned based on a pooled two-state student T-Test performed for each state and timepoint (ΔD > 3 x SD, 1-p value > 0.98).
Preliminary data: First, the conformational behaviour of pre-fusion stabilized soluble spike constructs produced in CHO cells to support the development of COVID-19 vaccines and treatments as well as diagnostic and serological assays was probed. Thorough biophysical characterization identified a mixture of trimer conformations (trimer 1 and trimer 2) existing in solution and highlighted that this behaviour is driven by VoC-specific mutations, formulation buffer pH and temperature, and the choice of trimerization domain. HDX-MS was used to compare, where it was determined that trimer 2 populations are significantly stabilized relative to a trimer 1 reference state. Notably, regions where HDX stabilization is observed for trimer 2 preparations are concentrated in the S2 domain which hosts a significant portion of the inter-protomer interface buried within the core of the trimer. Trimer 1 showed no HDX differences relative a monomeric protomer, suggesting that trimer 1 is composed of loosely associated individual protomers, with minimal inter-protomer contact, while trimer 2 is a more stable, compact conformation. Further, we identified antibodies that preferentially bind trimer 2. These results have implications for the development of cross-protective spike-based vaccine antigens and epitope accessibility informed by conformational data. Next, we investigated the epitopes of a collection of 24 single-domain antibodies raised against the spike protein. These have broad domain specificity, epitopic and mechanistic diversity, cross-reactivity across many VoCs, and high *in vitro* neutralization potencies. After coarse epitope binning was assigned by surface plasmon resonance and sandwich ELISA, the HDX-MS workflow was deployed to probe the conformational nature of the epitope bins at a peptide-level resolution. Unique binding modes emerged between bins, allowing for an improved understanding of the underlying binding and neutralization mechanisms. This granularity assists in understanding and predicting neutralization potency as novel variants emerge, and will guide the development of various therapeutic cocktails and multi-paratopic formats.
Novel aspect: We study demonstrate the unique versatility of HDX-MS in informing the development of therapies targeting the Covid-19 spike protein.

HDX-MS as a tool for the characterization of low-affinity fragments binding to human Cyclophilin D

Catarina F. Malta[1,2*], Tiago M. Bandeiras[1,2], Joerg Bomke[3], Diana O. Silva[1,2], Ulrich Graedler[4], Pedro M. Matias[1,2], Pedro M. F. Sousa[1], Daniel Schwarz[3], Alessio Bortoluzzi[1*]
1. iBET, Instituto de Biologia Experimental e Tecnológica, Av. da República, 2781-901 Oeiras, Portugal.
2. ITQB, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.
3. Merck Healthcare KGaA – Discovery Pharmacology, Frankfurter Strasse 250, 64293 Darmstadt, Germany.
4. Merck Healthcare KGaA – Molecular Interaction and Biophysics, Frankfurter Strasse 250, 64293 Darmstadt, Germany.
* To whom correspondence should be addressed.
Hydrogen deuterium exchange mass spectrometry (HDX-MS) is emerging as a powerful biophysical technique for probing protein interactions, structure, and conformational dynamics. While HDX-MS is well-established for the characterization of potent compounds binding to their target proteins there are only few examples in the literature regarding the application of HDX-MS to the study of low-affinity fragments and none describing the usage of HDX-MS for molecules with affinities in the mM range.
Fragment-based drug design (FBDD) has emerged as an effective alternative to high throughput screening for the identification of potent compounds in drug discovery. Fragments are small molecules with a molecular weight typically below 300 Da. The fragment-based drug design starts with library screens that generally result in fragment hits with weak affinities to their protein target. Then, the most promising hits are grown to more potent and selective compounds in a process that is often challenging and laborious. In this process structural information on the fragments binding mode is key for a successful optimization. HDX-MS is an attractive option to obtain this information, especially in cases where other high-resolution techniques, such as x-ray crystallography or NMR, do not work or are not an option.
In our work, we use human Cyclophilin D (CypD) as system to explore the possibility of using HDX-MS to characterize the binding of fragments with mM binding affinities. CypD is the mitochondrial isoform of Cyclophilins which plays an important role in the execution of cell death by regulating the mitochondrial permeability transition pore. Mitochondrial dysfunction has been implicated in a cascade

of cellular processes related to several diseases as multiple sclerosis and cardiovascular disease, making CypD an interesting target for therapeutic intervention. Colleagues at Merck Healthcare KGaA synthesized fragments targeting the different binding sites of CypD: the aniline, the proline, and the pyrrole pockets. In our work we will show that HDX-MS can be used to monitor the response of CypD to these three different series of fragments with millimolar (mM) affinities. Initially, fragments with crystal structure information available were tested in order to validate the HDX-MS method. Then, fragments with no structural information were tested. These experiments revealed a reduction in deuterium uptake in the different regions of the protein, allowing the mapping of the different binding sites. Altogether our results show that HDX-MS could be a valuable tool in fragment-based drug design projects also when the initial hits present weak affinities. Additionally, we will provide technical information to help the design of HDX experiments with this class of challenging ligands.

Interaction Proteomics – Master and Commander in Biomedical Research

Markku Varjosalo
University of Helsininki

Proteins are essential macromolecules participating in a wide range of biological processes, including signal transduction, transcription regulation, immune response, and enzymatic catalysis.. Protein-protein interactions (PPIs) are crucial for these activities and over 80% of proteins interact with others to carry out their functions (PMID: 17640003). PPIs play a pivotal role in the organization and function of the proteome, with perturbations often linked to various diseases such as cancer, neurodegeneration, and infectious diseases. PPIs can be stable or transient, with stable interactions forming protein complexes, while transient interactions predominantly occur in signal transduction. Many PPIs are part of larger PPI networks in the cell, and understanding the interactome is critical for comprehending the regulation of biological networks.

Recent advances in mass spectrometry (MS)-based protein interactomics have significantly expanded our understanding of PPIs in cells, with methods that continue to improve in terms of sensitivity, specificity, and resolution. For systematic PPI examination, experimental techniques such as two-hybrid systems, affinity purification, MS, protein chip technology, and computational modeling have been developed. The affinity purification approach uses a tagged protein of interest (POI) as “bait” to bind any interacting proteins (preys), which can be obtained from sources like cell and tissue lysates. The general proteomic workflow using MS is then applied for protein identification. Furthermore, affinity purification can be coupled with crosslinking mass spectrometry and Cryo-EM to obtain a structural proteomics view on protein complexes, providing valuable insights into their architecture and function. In our research, we have extensively applied interaction proteomics to investigate the dynamic interplay of key cellular signaling molecules, such as protein kinases, phosphatases, and transcription factors, in the context of tens of studies. Our primary focus has been on understanding how these interactions contribute to maintaining functional cells and how their perturbation may lead to various diseases, including cancer and immunological disorders. By utilizing advanced experimental techniques like affinity purification, crosslinking mass spectrometry, and Cryo-EM, we have been able to obtain comprehensive view of protein complexes, shedding light on their architecture and function. This structural proteomics approach has allowed us to gain valuable insights into the molecular mechanisms governing protein interactions and their roles in cellular processes.

One particular focus of our research has been the Commander complex, a 16-protein subunit assembly that plays multiple roles in various intracellular events, including regulation of cell homeostasis, cell cycle, and immune response. The complex is composed of COMMD1-10, CCDC22, CCDC93, DENND10, VPS26C, VPS29, and VPS35L. These proteins are expressed ubiquitously in the human body and have been linked to diseases including Wilson’s disease, atherosclerosis, and several cancers. Despite its importance, the structure and molecular functions of the Commander complex are poorly understood. Through our investigations, we have uncovered the structure and key interactions of the endogenous human Commander complex using cryogenic electron microscopy (cryo-EM) and mass spectrometry-based proteomics. Our results show that the complex is asymmetric, consisting of a stable core of a pseudo-symmetric ring of COMMD proteins 1–10 and a mobile effector consisting of DENND10 and the Retriever sub-complex, constituted by VPS35L, VPS29, and VPS26C. The two halves are scaffolded together by CCDC22 and CCDC93. This study directly confirms the cellular composition of Commander and identifies major interaction interfaces, defining the structure and interaction landscape of the complex.

These findings offer new insights into the known roles, and uncover strong association with cilium, centrosome and centriole functions.

A new tool to map the conformational proteome in cells to reveal the structural plasticity of the glucocorticoid receptor

Chloé Van Leene[1,2], Annick Verhee[1,2], Tijs Merckaert[1,2], Sven Eyckerman[1,2], Karolien De Bosscher[1,2], Kris Gevaert[1,2]
1 VIB-UGent Center for Medical Biotechnology, Ghent, Belgium
2 UGent Department of Biomolecular Medicine, Ghent, Belgium

Introduction: The glucocorticoid receptor (GR) is a multi-domain transcription factor with high conformational plasticity. GR is triggered by glucocorticoids, which are steroidal hormones involved in the regulation of cellular stress, immune responses, cellular differentiation and metabolism. Synthetic variants of these glucocorticoids have been derived to combat inflammation and cancer. However, chronic usage hereof leads to side effects and therapy resistance. Structural insights on GR are currently only derived from *in vitro* analysis of either its ligand- and/or DNA-binding domain. Studying GR's conformational flexibility in a cellular environment is essential to better understand its function and to identify targeted strategies that favour GR transcriptional programs with beneficial therapeutic effects.

Methods: Limited proteolysis coupled to mass spectrometry (LiP-MS) detects changes in protein conformation. The current LiP-MS technology works on cell or tissue lysates that undergo a first short incubation with proteinase K to introduce cuts in surface-accessible regions. Further and complete digestion with trypsin, comparison with a trypsin-only treated sample and mass spectrometric analysis leads to the identification of protein conformation-specific cuts via semi-tryptic peptides, termed conformotypic peptides. This workflow comes with several limitations such as not working in a native endogenous environment and experimental artefacts. Hence, we aim to develop a tightly controllable LiP workflow in living cells based on a rapamycin-controlled split-protease system.

Preliminary data: Three proteases were chosen as LiP proteases and their respective split fragments were generated. Full protease activity was evaluated in an activity assay. Importantly, no cell death was observed after 48 hours of LiP protease expression, allowing a comfortably large time window to set up the *in cellulo* LiP system. The reassembly of the protease fragments was validated *in vitro* and *in cellulo.* Spontaneous assembly of some split couples indicates the need for an iterative process to identify those split fragments that only reassemble and thereby form an active protease upon adding rapamycin. In future experiments, the activity of the *in cellulo* reconstituted protease will be tested. Once a controllable LiP protease is established, analysis of *in cellulo* ligand-induced GR conformotypic peptides via LC-MS/MS will follow. Such peptide patterns will be integrated with corresponding data of GR ligand-specific gene regulatory profiles aimed at finding those structural fingerprints resulting in a desired signalling pathway.

Novel aspect: Studying differences in GR conformation directly in live cells in a reproducible and quantifiable manner.

Combining XL-MS and LiP-MS to gain detailed molecular understanding about protein structural rearrangements upon perturbations

Cathy Marulli, Alexander Leitner, Paola Picotti
Department of Biology, Institute of Molecular Systems Biology, ETH Zürich, Zurich, Switzerland

Introduction: Cells quickly adapt to changing environments by modulating protein structure resulting in altered activity or function. Protein structure is regulated by numerous mechanisms such as post-translational modifications, proteolytic cleavage or binding to other proteins, DNA, RNA, and small molecules. Structural proteomics methods such as limited proteolysis (LiP) and cross-linking (XL) coupled to mass spectrometry (MS) capture those events on a proteome-wide scale. While LiP-MS only reports on surface accessibility changes between conditions, XL-MS carries information about distance between residues. We combine for the first time LiP-MS and XL-MS to the glycolytic to gluconeogenic shift in *Saccharomyces cerevisiae* to gain detailed molecular understanding of how the proteome structurally and functionally adapts to this perturbation.

Methods: *Saccharomyces cerevisiae* BY4716 was grown in triplicates in glucose and ethanol minimal medium. For LiP-MS, proteinase K (100:1 substrate:enzyme) was added to lysates (1 mg/ml) for 5 min, followed by standard workflows using trypsin. For XL-MS, lysates were fractionated by serial ultrafiltration in decreasing order of molecular weight cutoffs (100 kDa, 50kDa, 30 kDa, 10 kDa). Each fraction was cross-linked at 1 mg/ml protein concentration with 2.25 mM BS3 followed by standard workflows using trypsin. Crosslinks were enriched by size exclusion chromatography (5 fractions). Samples were analyzed by data-dependent and/or data-independent LC-MS/MS on a Thermo Orbitrap

Eclipse Tribrid or Fusion Lumos instrument (120 min gradients). LiP-MS data was analyzed in Spectronaut, XL-MS data was searched in xiSearch and quantified in Spectronaut.
Preliminary data: To generate a quantitative XL-MS dataset with high coverage and good reproducibility, we fractionated lysates by serial ultrafiltration prior to cross-linking and show that the number of intra-protein links identified in biological triplicates increases from 1258 to 4292 coming from 306 and 665 proteins, respectively. We use data-dependent acquisition to identify cross-links. From the latter, we build a spectral library and use data-independent acquisition to quantify the cross-links which considerably decreases the variability of cross-link quantifications between replicates and thus increases the statistical power to identify structural changes. Applying the XL-MS workflow in parallel with LiP-MS on yeast grown under glycolytic and gluconeogenic conditions in biological triplicates generates highly informative structural fingerprints for hundreds of proteins that report on protein abundance, intra- and inter-protein-link, mono-link and LiP changes. Those fingerprints allow to classify proteins as being regulated solely by structure, abundance, or both. Parallel identification of structural changes by LiP-MS and XL-MS gives us high confidence and detailed complementary information about the structural rearrangements and allows to generate hypotheses on potential regulatory events. In the case of phosphoglycerate kinase, our data suggests that the closed conformation of the enzyme is more occupied under gluconeogenic conditions and hypothesizes a potential site for allosteric regulation. Thus, we envision that pushing the boundaries of structural proteomics workflows is crucial to produce high throughput data on how protein structure relates to function and therefore essential for a detailed understanding of biological systems.
Novel aspect: Integration of a LiP-MS and a quantitative XL-MS dataset to gain biological understanding of protein structural rearrangements upon perturbation

Decoding functional surfaceome protein-protein interactions using light-induced cross-linking sites and HyPhoX
Sebastian N. Steiner[1,2] , Clinton Yu[3] , André Zwicky[4] , Alexander Leitner[5] , Jeffrey W. Bode[4] , Lan Huang[3], Bernd Wollscheid[1,2]
1 Institute of Translational Medicine (ITM), Department of Health Sciences and Technology (D-HEST), ETH
Zürich, Zürich, Switzerland
2 Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland
3 Department of Physiology & Biophysics, University of California, Irvine, California, USA
4 Laboratory for Organic Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zürich, Switzerland
5 Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, 8093 Zürich, Switzerland
Introduction: Not single proteins, but protein communities and protein-protein interactions (PPIs) within the surfaceome of human cells enable complex cellular signaling in response to environmental stimuli. Recent proximity-based protein tagging technology advancements enable the discovery of surfaceome-residing protein communities. Cross-linking of surfaceome-residing complexes can provide orthogonal (functional) information about the surfaceome architecture beyond the general information about proximal protein neighborhoods. Typical in vivo cross-linking approaches utilize chemistries addressing cross-linking sites already available within proteins and generally lack subcellular specificity. Here we developed a new cross-linking strategy, which is utilizing newly light-induced cross-linkable sites on cell surface proteins in combination with a novel tri-functional cross-linker molecule, termed HyPhoX. This Cell Surface Interaction MS (CSI-MS) technology enables the detection of localized surfaceome signaling architecture.
Methods: CSI-MS combines the utilization of light-induced cross-linking sites on cell surface receptors, chemical cross-linking, and mass spectrometry. In the CSI-MS strategy, we generate so-called kodecytes by coating the cell surface of human B lymphocytes with lipids carrying Singlet Oxygen Generators (SOGs). The SOGs enable the light-controlled photo-oxidation of proximal cell surface proteins. Alternatively, SOGs can also be targeted to a specific cell surface receptor using antibodies/ligands, as shown with the LUX-MS strategy, enabling targeted cross-linking of receptor communities. Subsequently, a short illumination of SOGs leads to the oxidation of Histidines within cell surface proteins which can then be used as “induced cross-linkable sites” for selective protein-protein cross-linking of receptor complexes and receptor interaction communities.
Preliminary data: First, we assessed the efficiency of our light-mediated cross-linkable site induction on B lymphocytes. Using imaging and flow cytometry, we observed consistent SOG coating of cells and, consequently global induction of cross-linkable sites on cell surface receptors within 30 seconds

of illumination. Upon MS analysis, 205 cell surface proteins were identified bearing light-induced cross-linkable sites. Knowing that we can utilize controlled illumination to induce new chemical moieties at the cell surface that can be used for subsequent chemical cross-linking of cell surface proteins, we set out to develop a new cross-linker molecule targeting these specific sites. This novel tri-functional cross-linker molecule has an IMAC-enrichable PhoX-based affinity handle (Steigenberger et al., 2019) and two reactive hydrazide moieties for targeting the light-introduced Histidine modifications in cell surface proteins. We further incorporated in the design of the cross-linker an MS-cleavable feature that generates signature ions during MS acquisition simplifying identification of spectra containing cross-links and targeted data analysis. Using model peptides and proteins, we confirmed the reactivity of our cross-linker molecule, MS-cleavability, MS cross-linker fragmentation mechanisms, and the detection of signature ions. The fragmentation pattern was used to establish a cross-linker specific search strategy which improves search efficiency for surfaceome wide cross-linking studies. The CSI-MS strategy is currently being tested in pilot experiments to map surfaceome PPIs on B lymphocytes and receptor CRISPR'ed B cells to provide functional insights into surfaceome signaling architecture and residing protein signaling complexes.
Novel aspect: CSI-MS is a strategy enabling the identification of stable cell surface protein communities/signaling complexes/PPIs utilizing tri-functional HyPhoX cross-linker and light-inducable cross-linking sites.

A bipartite GINS binding mode of TopBP1 activates the replicative helicase MCM
Matthew Day[*2], Bilal Tetik[*1], Milena Parlak[*1], Yasser Almeida-Hernández[3], Elsa Sanchez-Garcia[3], Farnusch Kaschani[4], Markus Kaiser[4], **Markus Räschle**[5], Heike Siegert[1], Anika Marko[1], Isabel A. Barker[2], Laurence H. Pearl[2+], Antony W. Oliver[2+], Dominik Boos[1+]
1 Molecular Genetics II, Vertebrate DNA Replication, Centre of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany
2 Cancer Research UK DNA Repair Enzymes Group, Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9RQ, UK
3 Computational Biochemistry, Centre of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany
4 Analytics Core Facility Essen, Centre of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany
5 Molecular Genetics, Technical University Kaiserslautern, Paul-Ehrlich Straße 24, 67663 Kaiserslautern, Germany
* equal contribution
+ corresponding author
In dividing cells genome duplication is tightly coordinated with the cell cycle to ensure that each daughter cell receives a full copy of the genome. In all eukaryotes DNA replication initiates at replication origins in a two-step process; replication licensing and firing. During G1, licensing factors load the replicative MCM helicase onto DNA. In these pre-Replication complexes (pre-RCs), MCM binds DNA as an inactive symmetrical double hexamer encircling both strands of the double helix. As cells enter S phase, the rising activities of the CDK and DDK cell cycle kinases trigger multiple reconfigurations at origins to activate the MCM helicase and to form two mature bidirectional replisomes, one for each MCM hexamer in the pre-RCs.
While replication licensing is understood well at the molecular level, relatively little is known about the steps that turn MCM into an active replisome. In this process, the "firing factor" TopBP1 plays an essential role by facilitating recruitment of the GINS complex to the MCM helicase, which is key to its activation. To gain insight into TopBP1 function, we solved a cryo-EM structure of a TopBP1 fragment bound to the GINS complex. The structure reveals two separate GINS binding sites within TopBP1, which we validate by mutational analysis, fluorescence anisotropy measurements and MS crosslinking experiments. Using an *in vitro* replication system based on *Xenopus* egg extracts, we show that both GINS binding sites in TopBP1 contribute to DNA replication and must cooperate to support full assembly of replisomes from pre-RCs, as shown by comprehensive MS analysis of the chromatin bound replication complexes.
Our studies provide novel molecular insight into replisome. They further offer a basis for understanding how TopBP1 performs it essential function in replication origin firing and how this is coordinated with TopBP1's plethora of functions in the DNA damage response and chromosome maintenance.

Exploring cellular organization in native cell extracts with high-resolution cryo-EM

Panagiotis L. Kastritis[1,2]
1 Institute of Biochemistry and Biotechnology; Biozentrum; and Interdisciplinary Research Center HALOmem, Martin-Luther University Halle-Wittenberg, Halle (Saale), Germany;
2 Institute of Chemical Biology, National Hellenic Research Foundation, Athens, Greece

Advances in electron microscopy have provided unprecedented access to the structural characterization of large, flexible, and heterogeneous complexes. Until recently, cryo-electron microscopy (cryo-EM) has been applied to understand molecular organization in either highly purified, isolated biomolecules or *in situ*. An emerging field is developing, bridging the gap between the two approaches, studying molecular organization in native cellular fractions. Due to the biochemical nature of the cell extract, functional assays, mass spectrometry, and single-particle cryo-EM analysis, combined with artificial intelligence, can provide multi-scale structure-function understanding. Here, I will present structural insights into previously elusive megadalton protein complexes uncovered *via* integrative analysis of cellular fractions. Overall, our methods and data provide a framework for further understanding cellular organization by systematically mapping the structure of endogenous biomolecular assemblies and their interactions within cellular fractions.

Hyphenating Ion Mobility and Action Spectroscopy in a Synapt G2 to Probe the Structure and Kinetics of Aggregating Peptides

Steven Daly[1], Sjors Bakels[2], Agathe Depraz[2], Iuliia Stroganova[2], Jan Commandeur[1], Anouk M. Rijs[2]
1 MS Vision,Televisieweg 40, 1322AM Almere, The Netherlands
2 Division of BioAnalytical Chemistry, AIMMS Amsterdam Institute of Molecular and Life Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HV, Amsterdam, The Netherlands.

Nature provides us with an impressive array of examples of both function and malfunction arising from extreme complexity. An important example of a highly complex molecular network is that of aggregating peptides and proteins. Aggregation, the transition from soluble functioning proteins into insoluble amyloid aggregates, is directly related to neurodegenerative diseases including Alzheimer's and Parkinson's disease. To control and prevent protein aggregation, a full understanding of the initial, neurotoxic steps of the aggregation process is vital. However, the complex, heterogenic and dynamic nature of this process, makes this a huge experimental challenge. To tackle this, we have developed a novel, multidimensional spectroscopy- and mass spectrometry-based method, which allows us to probe the structure and kinetics of the initial steps of aggregation in a single measurement.

MS Vision and Anouk Rijs's MS-LaserLab group (VU Amsterdam) have collaborated to modify a Synapt G2 such that we can hyphenate droplets-based microfluidics, electrospray ionization, mass spectrometry, ion mobility spectrometry and ion spectroscopy in a single experiment (the Photo-Synapt). To be able to measure mass-and mobility-selected UVPD and IRMPD spectra, the Synapt G2 was modified with two additional hexapoles between mobility and TOF stages. The use of pin traps allows the storing, manipulation and irradiation with UV or IR photons of mass- and mobility selected ions.

In this presentation, we will focus on the design and implementation of the Photo-Synapt, and characterization of its different operational modes. Ion optical simulations were used to guide the design the hexapole length and geometry of the pin traps, and the operation of the different trapping modes.

The pin traps are three sets of six pins positioned between the hexapole rods. Two sets of the pins are form potential well within the hexapole to trap the ions. The first set of pins is used to eject ions from the trap and into the second hexapole for trapping and irradiation or to the TOF. We show that the addition of the additional hexapoles does not affect the standard IM-MS and MS^n performance, that we can trap in both pin traps, and transfer the ions between the two traps and to the TOF. The influence of the trapping time, gas pressure and pin trap voltages are explored to find optimal trapping conditions. Successful UVPD experiments of iodotyrosine demonstrate that we can trap and induce photodissociation. We are currently being extended to optimize conditions to perform IRMPD measurements.

Finally, the ability to perform ion mobility slicing is demonstrated. This is achieved by pulsing the IMS exit lens. The trapping and irradiation of a mass and mobility selected population of ions probed by IR or UV photons is the final goal of the characterization experiments. Initial experiments have shown this is possible, although optimisation of the protocol is still required.

Quantitative crosslinking and mass spectrometry indicates kinetochore complex stabilization

Goetz Hagemann[1], Victor Solis-Mezarino[1], Sylvia Singh[1], Mia Potocnjak[1], Chandni Kumar[1], Benjamin Neuditschko[2] and **Franz Herzog**[1,2,*]

1 Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany

2 Institute Krems Bioanalytics, IMC University of Applied Scienes Krems, 3500 Krems, Austria

Introduction: Distance restraints derived from the mass spectrometric identification of crosslinked amino acids (XLMS) are widely applied in integrative approaches to determine protein connectivity and to model the topology of protein complexes. Besides structure, the critical determinant of the molecular mechanism of a complex is the interaction strength of its subunit contacts, which can be modulated through cofactors or post-translational modifications. We reasoned that crosslink intensities provide a quantitative measure for the formed complex and the free subunits at the equilibrium state. Thus, we investigated whether crosslink intensities facilitate the simultaneous estimation of individual protein-protein affinities within kinetochore multi[1]subunit complexes.

Methods: Protein complexes were crosslinked by modifying the α-amino groups with the isotopically labeled BS2G-d0/d6 reagent and crosslinked peptide fractions were analyzed by liquid chromatography coupled to tandem mass spectrometry. The raw files were processed by the xQuest/xProphet software to identify the crosslinked peptides, their precursor ion masses and retention times. This information was subsequently used for the extraction of ion chromatograms by the OpenMS software tool. We further applied inter- and intra-protein protein crosslink intensities to estimate the concentrations of the formed complex and the free subunits according to the steady state equilibrium in solution. To assess whether crosslink intensities supported the estimation of binding affinities and interfaces we titrated recombinant kinetochore subunits for complex formation over a range of molar ratios.

Preliminary data: We found a dependence between crosslink distances and intensities and applied a quantitative workflow to estimate binding affinities and aid interface prediction of kinetochore subunit contacts which link chromosomes to spindle microtubules. Titrating the assembly of 11 subunits showed that phosphorylation induces a high-affinity link to the centromeric nucleosome required for transmitting forces of depolymerizing microtubules. Phosphorylation of Mif2CENP-C putatively by Cdc5PLK1 induced binding to Ame1/Okp1CENP-U/Q in vitro and mediated their cooperative stabilization at the Cse4CENP-A nucleosomes which was enhanced by binding of the MTW1MIS12 complex phosphorylated by Ipl1Aurora-B. Together, both phosphorylation events decreased the KD-values of Cse4CENP-A interactions by ~300-fold and are essential for cell viability. This work demonstrates the potential of quantitative XLMS for characterizing the stabilization of the kinetochore at the centromeric nucleosome through phosphorylation of the outer and inner kinetochore in mitosis.

Novel aspect: 1. Crosslink intensities as quantitative measures for the formed complex and the free subunits at equilibrium allow estimation of binding affinities in multi-subunit complexes. 2. Correlation between increasing crosslink intensity and decreasing distance to protein contact facilitates narrowing down binding interfaces

Thrombin activation of the factor XI dimer is a multi-staged process for each subunit

Awital Bar Barroeta[1], Pascal Albanese[2,3], J. Arnoud Marquart[1], Joost C.M. Meijers[1,4], Richard A. Scheltema[2,3]

1 Department of Molecular Hematology, Sanquin, Amsterdam, the Netherlands

2 Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

3 Netherlands Proteomics Centre, Utrecht, The Netherlands

4 Department of Experimental Vascular Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands

Introduction: Factor XI (FXI), a protein in the intrinsic coagulation pathway, can be activated by two enzymes. In hemostasis, FXI is activated by thrombin, while FXIIa-mediated activation of FXI is thought to be prothrombotic. The interactions of these enzymes with FXI are transient in nature and therefore difficult to study in a structural context.

Methods: Crosslinking mass spectrometry (XLMS) was used to construct the FXI homodimer and to localize the binding interface of thrombin on FXI. Molecular dynamics simulations were then applied to investigate conformational changes after binding. Additionally, the binding site of nanobody 1C10 – previously shown to inhibit thrombin-mediated activation of FXI – was investigated with hydrogen-deuterium exchange mass spectrometry (HDX MS).

Preliminary data: Our investigations show that the activation of FXI is a multi-staged procedure. We identified that thrombin initially interacts with the light chain of FXI by binding Pro520. Following this initial interaction, FXI undergoes conformational changes driven by binding of thrombin to allow migration towards the FXI cleavage site by first engaging the apple 1 domain and, finally, Arg378. We validated the results with known mutation sites on FXI and additionally found that Pro520 is conserved in PK. Through this site, thrombin can bind PK even though it cannot activate PK. Moreover, the proposed trajectory of thrombin-mediated FXI activation was supported by elucidation of the 1C10 binding site on the apple 1 domain. This detailed analysis of the interaction between thrombin and FXI points a way for future interventions for bleeding or thrombosis.
Novel aspect: Elucidation of the transient enzymatic thrombin-FXI interaction and the corresponding activation mechanism.

Flexible domains on membrane proteins and how to study them – combining cryo-EM, limited proteolysis-coupled and crosslinking mass spectrometry

Diane C.A. Barret[1] *, **Dina Schuster**[1, 2, 3] *, Matthew J. Rodrigues[1], Alexander Leitner[2], Paola Picotti[2], Gebhard F.X. Schertler[1], U. Benjamin Kaupp[4, 5], Volodymyr M. Korkhov[1, 3] and Jacopo Marino[1]
1 Laboratory of Biomolecular Research, Paul Scherrer Institute, Villigen, Switzerland.
2 Institute of Molecular Systems Biology, Department of Biology, ETH Zürich, Zürich, Switzerland.
3 Institute of Molecular Biology and Biophysics, Department of Biology, ETH Zürich, Zürich, Switzerland.
4 Life and Medical Sciences Institute LIMES, University of Bonn, Bonn, Germany.
5 Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany
* Contributed equally
Introduction: Membrane proteins, due to their low expression levels and high hydrophobicity, are difficult to study and underrepresented in common bottom-up proteomics experiments. They often contain large stretches of flexible, partially disordered domains, which are not amenable to conventional structural biology techniques, such as cryo-EM and X-ray crystallography. We use a combination of cryo-EM, limited proteolysis-coupled mass spectrometry (LiP-MS) and crosslinking mass spectrometry (XL-MS) to obtain a comprehensive understanding of how calmodulin (CaM) interacts with the cyclic nucleotide-gated ion channel (CNG) of retinal rods. We apply LiP-MS on native retinal membrane suspensions and XL-MS on purified protein to gain insight in structural features that cryo-EM could not resolve.
Methods: For LiP-MS experiments, suspensions of bovine retinal membranes were titrated with increasing concentrations of CaM ranging from 0-3 µg, followed by pulse proteolysis with proteinase K and a tryptic digest under denaturing conditions. Data was acquired in DIA mode on a Thermo Scientific Orbitrap Exploris 480 mass spectrometer and analyzed using Spectronaut (Biognosys). The purified CNG channel was crosslinked with CaM using primary amine crosslinking with disuccinimidyl suberate. Data was acquired on a Thermo Scientific Fusion Lumos Tribrid mass spectrometer and analyzed using xQuest. For single-particle cryo-EM, proteins were solubilized from membrane preparations and purified by affinity- and size-exclusion chromatography. Samples were blotted onto carbon grids, plunge-frozen, and data was collected on a Thermo Scientific Krios G4 electron microscope.
Preliminary data: Our data shows that LiP-MS and XL-MS provide complementary information about the interaction of the CNG channel with CaM. In particular, these methods give structural information on flexible domains that are invisible to cryo-EM. LiP-MS, a method based on the use of an unspecific protease (proteinase K), which cleaves accessible protein regions, shows changes in protease accessibility upon interaction. XL-MS provides proximity information through covalent linkage of primary amines. LiP-MS enables structural and interaction studies *in situ* without the need for protein purification and due to the high sensitivity of mass spectrometry requires only minute amounts of the sample. The method is advantageous for structural studies of membrane proteins as these proteins can be studied directly in their native lipid environment, avoiding the use of detergent and laborious purification procedures. The additional use of XL-MS together with LiP-MS helps distinguish direct interaction sites from conformational changes, thereby overcoming a limitation of LiP-MS. We show that the CNG channel behaves similar in detergent and in its native lipid environment. We identify that, in the presence of saturating Ca^{2+} concentrations, CaM binds to two different sites on the CNG channel, with one of them being close to the coiled-coil region of the protein. Cryo-EM studies in presence and absence of CaM show that the channel adopts a more compact structure upon interaction. LiP-MS identifies the interaction sites of CaM and additional accessibility changes on

flexible domains of the protein. XL-MS helps interpret both the cryo-EM and LiP-MS results by identifying the specific binding site of CaM visible in the cryo-EM structure but unresolved due to low resolution. Taken together, the combination of cryo-EM, LiP-MS, and XL-MS can help better understand the function of membrane proteins and the role of their flexible domains.
Novel aspect: LiP-MS and XL-MS help study flexible domains on membrane proteins and features that cryo-EM cannot resolve.

Our Great Sponsors

https://msvision.com/

https://www.thermofisher.com

https://epic-xs.eu/

https://www.mswil.com/

Creative Molecules Inc.

http://www.creativemolecules.com/

https://sciex.com/

https://www.glsciences.com/

https://www.refeyn.com/

https://www.bruker.com

https://affipro.cz/

CF Plus

Shimadzu

Waters