

The emerging role of mass spectrometry in structural biology

Elisabetta Boeri Erba

▶ **To cite this version:**

Elisabetta Boeri Erba. The emerging role of mass spectrometry in structural biology. Structural Biology [q-bio.BM]. Communauté Université Grenoble-Alpes, 2018. tel-02025567

HAL Id: tel-02025567

https://hal.univ-grenoble-alpes.fr/tel-02025567

Submitted on 19 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

“The emerging role of mass spectrometry in structural biology”

HABILITATION À DIRIGER DES RECHERCHES

présentée devant

l’École doctorale Chimie et Sciences du vivant (EDCSV), Université Grenoble Alpes

par

Elisabetta Boeri Erba

Chercheur à l’Institut de Biologie Structurale (UMR5075)

Jury

Président

Eva Pebay-Peyroula, Professor, Université Grenoble Alpes

Rapporteurs

Martine Cadene, Directrice de recherche, Cnrs, Orleans

Claudio Medana, Professor, University of Turin (Italy)

Christophe Masselon, Chercheur, CEA, Grenoble

Examinateurs

Véronique Rossi, Maître de Conférences, Université Grenoble Alpes

Philippe Dugourd, Directeur de l'Institut Lumière Matière, Université Claude Bernard Lyon 1

Date de présentation des travaux: 20.12.2018

Lieu: Institut de Biologie Structurale, 71 Avenue des Martyrs, 38044 Grenoble

Table of Contents

Summary of the thesis 3

Part I: summary of research career 4

Curriculum vitae 4

Publications 8

Part II: previous research (2001-2018) 11

Brief introduction about biological MS: the different approaches I used during my career 11

The emerging role of MS in structural biology 17

University of Turin (Italy) and University of Southern Denmark (2001-2006) 17

Study of the epidermal growth factor receptor by bottom-up MS-based proteomics 17

Identification and characterisation of the oncosuppressor p140Cap by proteomics 19

University of Cambridge (UK) and Swiss Federal Institute of Technology (ETH), Zurich (CH) (2006-2011) 22

Investigation of intact protein complexes by native MS 22

Institute of Structural Biology (IBS), Grenoble (2011-2018) 29

Structural studies of biomedically-relevant protein complexes 29

Methodological development for studying intact proteins and macromolecular complexes by MS 35

Part III: current and future research at the IBS 39

Characterisation of therapeutic antibodies using MS 39

Native MS to study biomedically- and biotechnologically-relevant complexes 46

Improving MS to investigate large macromolecular complexes 48

References 52

Summary of the thesis

Biological mass spectrometry (MS) is an important analytical approach to assess the mass of biomolecules with high accuracy, sensitivity and speed of analysis. It has wide applications and it is of fundamental importance to answer many diverse biological questions. In my thesis I briefly introduce the four types of biological MS I used during my scientific career. I illustrate my work during my PhD studies (2001-2005) and as postdoctoral fellow (2006-2011) working in England and in Switzerland. Finally, I describe my past, current and future MS investigations I carry out at the Institute of Structural Biology (IBS) in Grenoble. To conclude, MS represents an emerging tool in structural biology because it provides novel insights on the two dimensional map of macromolecular assemblies with high sensitivity, complementing traditional structural biology approaches.

This thesis is dedicated to the memory of Prof. Guido Tarone (1951–2015), who helped me discover mass spectrometry and my talent as scientist.

Questa tesi e' dedicata al Prof. Guido Tarone (1951–2015), che mi ha aiutato a scoprire la spettrometria di massa e la mia vocazione per la scienza.

Part I: summary of research career

Curriculum vitae

Current position

2015-present: CEA permanent researcher at the *Institut de Biologie Structurale (IBS), Grenoble (France)*

2018-present: Head of the mass spectrometry (MS) laboratory of the IBS

2011-2018: Team leader at the IBS

Research topic: **Investigation of macromolecular complexes by MS**

Degrees

2005: Ph.D. degree in **Human Biology**, *University of Turin (Italy)* in collaboration with *University of Southern Denmark*

2000: M.Sc. degree in **Pharmaceutical Chemistry**, Pharmacy Faculty, *University of Turin,* final mark 106/110

Research experience

2010-2011: Post-Doctoral Fellow at the *Department of Chemistry and Applied Biosciences, ETH Zurich (Switzerland)*

Mass spectrometric investigation of macromolecular complexes

Supervisor: Prof. Renato Zenobi

2006-2010: Post-Doctoral Fellow at the *Chemistry Department, University of Cambridge (UK)*

Mass spectrometry of macromolecular complexes

Supervisor: Prof. Carol V Robinson

2006: Post-Doctoral Fellow at the *Department of Biochemistry & Molecular Biology, University of Southern Denmark (Odense)*

Development of mass spectrometry-based proteomic approaches to characterise protein phosphorylation

Supervisor: Prof. Ole Noerregaard Jensen

Training and education

2001-2005: Ph.D. student at the *Dept. of Genetics, Biology and Biochemistry, Medical School, University of Turin (Italy)* and in the *Protein Research Group* at the *Dept. of Biochemistry and Molecular Biology*, *University of Southern Denmark*.

Thesis title: "**The use of mass spectrometry to identify new proteins involved in signal transduction and to characterise post-translational modifications**"
Supervisors: Prof. Ole Noerregaard Jensen and Prof. Paola Defilippi

Scientific prizes and awards

2005-2006: *Lundbeck Foundation* fellowship
2004-2005: *CIRIUS* studentship (Danish Information and Documentation Centre for Internationalisation of Education)
2004: *CIB* studentship (Italian Inter-University Biotechnology Consortium)
2002-2003: *Marie Curie training* studentship
2001: medical School studentship, University of Turin, Supervisor: Prof. Paola Defilippi

Funding

2018: main responsible of CEA DRF-Impulsion grant " Investigation of phage T5 using multi-scale mass spectrometry: from intact virions to individual proteins".
2018: co-applicant on CEA Transversal Competency Program "Instrumentation and detection" "Nano-electro-mechanical-system mass spectrometry for the standardization of large scale virion production"
2018: funding from GIS IBiSA (Infrastructures in Biology, Health and Agronomy) to purchase a new Matrix Assisted Laser Desorption Ionisation mass spectrometer (280 k€)
2017 co-applicant on ANR research grant on chloroplast-encoded RNA polymerase complex
2010 co-applicant on a research grant application to the Swiss Research Council to purchase a new electrospray mass spectrometer

Teaching experience

i) Student and visitor (co)supervision

2018: Edison Zhamungui Sánchez, Master M1 student, *Université Grenoble Alpes*
Rita Puglisi, EMBO short-term visitor, *King's College London (UK)*
2017: Y. Pauvert, Master M1 student, *Université Grenoble Alpes*
2016: B. Arragain Master M1 student, *Université Grenoble Alpes*
I. Ben Mabrouk, L3 student, *Université Grenoble Alpes*
2015: A. Ben Chaaben, Master M2 student, *Université de Tunis, Tunis (Tunisie)*
A. Ruppin Master M1 student, *Université Grenoble Alpes*
2014: P. A. Klein, Master M2 student, *Université Joseph Fourier Grenoble*
A. Hanif, Brevet de Technicien Supérieur student, *Louise Michel Lycée, Grenoble*
2013-2017: Mizar Oliva, PhD student, *Université Grenoble Alpes*
Alycia Yee, PhD student, *Institut Laue Langevin, University College London*

2013: S. Camus, L1 student, *Université Joseph Fourier Grenoble*

S. Lefebvre, L3 student, *Université Joseph Fourier Grenoble*

2010-2011: K. Barylyuk, PhD student, *ETH Zurich*

S. Mädler, PhD student, *ETH Zurich*

Y. Yang, Master M2 student, *ETH Zurich*

2009: R. Salbo, PhD student from *University of Southern Denmark*

J. Bolsen, PhD student from *Technical University of Denmark, Copenhagen*

ii) University teaching responsibilities

2019-2015: Lectures at the Hercules European school, *Université Grenoble Alpes*

2018-2012: Lectures at the course: "**Proteomics**", *University of Turin,* course coordinator: Prof. Paola Defilippi

2005: Lectures at the course: "**Advanced Mass spectrometry course in protein chemistry and proteomics: Electrospray ionization mass spectrometry**", *University of Southern Denmark*

2004-2005: Lectures at the course: "**Advanced Methods in Protein Chemistry**", *University of Southern Denmark*

2004: Lectures at the course: "**Mass spectrometry in protein chemistry and proteomics: MALDI ionization mass spectrometry**", *University of Southern Denmark*

Additional information

Engagement as a reviewer

2014-present: the Flanders Organisation for Scientific Research (FWO)

2012-present: the Netherlands Organisation for Scientific Research (NWO)

2010-present: Journals: Journal of the American Chemical Society", "Protein science", "Methods", "Analytical chemistry", "Proteomics", "Scientific Reports", "International Journal of Mass Spectrometry", "Toxin" and "Rapid Communications in Mass Spectrometry".

Memberships

2011-2016: Member of the French Society of Mass Spectrometry

2011-2012: Member of the American Biophysical Society

2009-2010: Fellow at Wolfson College, Cambridge (UK)

2007-2010: Member of the American Society for Mass Spectrometry

2004, 2005, 2007, 2009, 2011: Member of the Protein Society

2008: Member of the Protein Society

2004-2005: Member of the Italian Society of Biochemistry and Molecular Biology

Organised symposium

2014: Native Mass Spectrometry and its Application in Structural Biology, *IBS*, Grenoble

Invited talks at French and International conferences

2018: Using native MS in structural biology. *23th RCJSM*, Montpellier

From intact proteins to macromolecular complexes: the emerging role of mass spectrometry in structural biology. *Pharma Network 2018*, Rome (Italy)

2018, 2016, 2014, 2012, 2010: Native mass spectrometry of protein complexes. *EMBO practical course: Structural Characterization of Macromolecular Complexes*, Grenoble

2017: The use of isotopic labelling and time-resolved native mass spectrometry to monitor the assembly of protein complexes. *AILM 2017 workshop,* Grenoble

2016: The key role of native MS in characterizing the structure and dynamics of macromolecular complexes. *Journées Françaises de Spectrométrie de masse*, Bordeaux

2015: The emerging role of native MS. *Workshop: Molecular Interaction: the complementarity between biophysical methods,* Grenoble

2014: When mass spectrometry meets structural biology: new answers to old questions.

4thSymposium on Structural Proteomics 2014, Antwerp (Belgium)

Molecular self organisation in life processes, 24th French biophysical society meeting, Guéthary, France

Monitoring the assembly pathway of large nanomachines using native mass spectrometry *IBBI 14 - Conference on Isolated Biomolecules and Biomolecular Interactions*, Porquerolles, (France)

2012: Investigating a large human protein assembly by native mass spectrometry: the SAGA complex. *29th French Society for Mass Spectrometry Conference (JFSM)*, Orléans (France)

Determination of intra-subunits interactions within noncovalent complexes by ESI-mass spectrometry *Congress of Mass Spectrometry Applied to Proteomics (SMAP 2011)*, Avignon (France)

2011: Quantitative determination of protein-protein interactions within noncovalent complexes by ESI-mass spectrometry. *59th American Society for Mass Spectrometry (ASMS) Conference*, Denver (Colorado)

Assembly and disassembly of protein complexes monitored by electrospray mass spectrometry. *Anakon*, Zürich (Switzerland)

2009: Homo- and hetero-oligomeric complexes in the gas phase investigated by ion mobility and tandem mass spectrometry. *57th ASMS Conference*, Philadelphia (Pennsylvania)

Publications

*: corresponding author

1. **Boeri Erba E***, Signor L, Oliva M F, Hans F, Petosa C. Characterising intact macromolecular complexes using native mass spectrometry. Methods Mol Biol. 2018;1764:133-151. doi: 10.1007/978-1-4939-7759-8_9

2. Pellegrini E, Signor L, Singh S, **Boeri Erba E**, Cusack S. Structures of the inactive and active states of RIP2 kinase inform on the mechanism of activation. PLoS One. 2017; 12(5):e0177161. doi: 10.1371/journal.pone.0177161

3. Alfieri A, Sorokina O, Adrait A, Angelini C, Russo I, Morellato A, Matteoli M, Menna E, **Boeri Erba E**, McLean C, Armstrong JD, Ala U, Buxbaum JD, Brusco A, Couté Y, De Rubeis S, Turco E, Defilippi P. Front Mol Neurosci. 2017;10:212. doi: 10.3389/fnmol.2017.00212. eCollection 2017

4. Sauer P, Timm J, Liu D, Sitbon D, **Boeri Erba E**, Velours C, Mücke N, Langowski J, Ochsenbein F, Almouzni G, Panne D. Insights into the molecular architecture and histone H3-H4 deposition mechanism of yeast chromatin assembly factor 1. eLife 2017; 6. pii: e23474. doi: 10.7554/eLife.23474.

5. Macek P, Kerfah R, **Boeri Erba E**, Crublet E, Moriscot C, Schoehn G, Amero C, Boisbouvier J. Unraveling Self-Assembly Pathways of the 468 kDa Proteolytic Machine TET2. Sci Adv. 2017; 3(4):e1601601. doi: 10.1126/sciadv.1601601

6. Lopez S, Rondot L, Cavazza C, Iannello M, **Boeri Erba E**, Burzlaff N, Strinitz F, Jorge-Robin A, Marchi-Delapierre C, Ménage S. Efficient conversion of alkenes to chlorohydrins by a Ru-based artificial enzyme. Chem. Commun., 2017; DOI: 10.1039/C6CC08873B

7. Baquero E, Albertini AA, Raux H, Abou-Hamdan A, **Boeri Erba E**, Ouldali M, Buonocore L, Rose JK, Lepault J, Bressanelli S, Gaudin Y. Structural intermediates in the fusion-associated transition of vesiculovirus glycoprotein. EMBO J. 2017; 36(5):679-692. doi: 10.15252/embj.201694565

8. Yee AW, Moulin M, Breteau N, Haertlein M, Mitchell EP, Cooper JB, **Boeri Erba E***, Forsyth VT*. Impact of Deuteration on the Assembly Kinetics of Transthyretin Monitored by Native Mass Spectrometry and Implications for Amyloidoses. *Angew Chem Int Ed Engl.* 2016; 55(32):9292-6. doi: 10.1002/anie.201602747

9. Rasmussen KK, Frandsen KE, **Boeri Erba E**, Pedersen M, Varming AK, Hammer K, Kilstrup M, Thulstrup PW, Blackledge M, Jensen MR, Lo Leggio L. Structural and dynamics studies of a truncated variant of CI repressor from bacteriophage TP901-1. *Sci Rep.* 2016; 6:29574. doi: 10.1038/srep29574

10. **Boeri Erba E***, Klein PA, Signor L. Combining a NHS ester and glutaraldehyde improves crosslinking prior to MALDI MS analysis of intact protein complexes. *J Mass Spectrom.* 2015; 50(10):1114-9

11. Caillat C, Macheboeuf P, Wu Y, McCarthy AA, **Boeri Erba E**, Effantin G, Göttlinger HG, Weissenhorn W, Renesto P. Asymmetric ring structure of Vps4 required for ESCRT-III disassembly. *Nat Commun.* 2015; 6:8781 10.1038/ncomms9781

12. **Boeri Erba E***, Petosa C. The emerging role of native mass spectrometry in characterising the structure and dynamics of macromolecular complexes. *Protein Sci*., 2015; 24(8):1176-92 10.1002/pro.2661

13. Lebrette H, Borezée-Durant E, Martin L, Richaud P, **Boeri Erba E**, Cavazza C.
Novel insights into nickel import in Staphylococcus aureus: the positive role of free histidine and structural characterization of a new thiazolidine-type nickel chelator. *Metallomics*, 2015; 7(4):613-21

14. **Boeri Erba E***. Investigating macromolecular complexes using top-down mass spectrometry. *Proteomics*, 2014; 14(10):1259-70 10.1002/pmic.201300333

15. Signor L, **Boeri Erba E***.
Matrix-assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) Mass Spectrometric Analysis of Intact Proteins Larger than 100 kDa. *J Vis Exp,* 2013; *(79), e50635 10.3791/50635*

16. Repetto D▲, Aramu S▲, **Boeri Erba E▲**, Sharma N, Jensen O, Cabodi S, Turco E, Di Stefano P, Defilippi P. Identification of phosphorylation sites on the adaptor protein p140Cap: p140Cap mediates binding to the C-terminal Src kinase through the EPLYA and EGLYA phosphorylated tyrosines. *Plos one, 2013; 8(1):e54931*
▲ authors equally contributed to this investigation

17. Pellegrino S, Radzimanowski J, de Sanctis D, **Boeri Erba E**, McSweeney S, Timmins J.
Structural and functional characterization of an SMC-like protein RecN: New insights into double-strand break repair. *Structure,* 2012; 20(12):2076-89

18. Brand GD, Salbo R, Jørgensen TJ, Bloch C Jr, **Boeri Erba E**, Robinson CV, Tanjoni I, Moura-da-Silva AM, Roepstorff P, Domont GB, Perales J, Valente RH, Neves-Ferreira AG. The interaction of the antitoxin DM43 with a snake venom metalloproteinase analyzed by mass spectrometry and surface plasmon resonance. *J Mass Spectrom*. 2012; 47(5):567-73.

19. Mädler S, **Boeri Erba E**, Zenobi R.
MALDI-ToF Mass spectrometry for studying noncovalent complexes of biomolecules. *Topics in Current Chemistry* 2012; 1-36.

20. Mädler S, Barylyuk K, **Boeri Erba E**, Nieckarz R J, Zenobi R.
Compelling advantages of negative ion mode detection in high-mass MALDI-MS for the analysis of homomeric protein complexes*. J Am Soc Mass Spectrom.* 2012; 23(2):213-24.

21. **Boeri Erba E**, Barylyuk K, Yang Y, Zenobi R.
Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal Chem.* 2011; 83(24):9251-9.

22. **Boeri Erba E**, Zenobi R.
Mass Spectrometric Studies of Dissociation Constants of Noncovalent Complexes. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2011; 107, 199-228.

23. **Boeri Erba E**, Ruotolo B, Barsky D, Robinson C V.
Ion Mobility-Mass Spectrometry Reveals the Influence of Subunit Packing and Charge on the Dissociation of Multi-protein Complexes. *Anal Chem*. 2010; 82(23):9702-10.

24. Sharon M, Mao H, **Boeri Erba E**, Stephens E, Zheng N, Robinson CV.

Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality. *Structure*. 2009; 17(1):31-40.

25. Levy ED, **Boeri Erba E**, Robinson CV, Teichmann SA.
Assembly reflects evolution of protein complexes. *Nature*. 2008; 453(7199):1262-5

26. **Boeri Erba E**, Matthiesen R, Bunkenborg J, Schulze WX, Di Stefano P, Cabodi S, Tarone G, Defilippi P, Jensen ON.
Quantitation of multisite EGF receptor phosphorylation using mass spectrometry and a novel normalization approach. *J Proteome Res*. 2007; 6(7):2768-85

27. Voena C, Conte C, Ambrogio C, **Boeri Erba E**, Boccalatte F, Mohammed S, Jensen ON, Palestro G, Inghirami G, Chiarle R.
The tyrosine phosphatase Shp2 interacts with NPM-ALK and regulates anaplastic lymphoma cell growth and migration. *Cancer Res*. 2007; 67(9):4278-86

28. Durán MC, **Boeri Erba E**, Mohammed S, Martín-Ventura JL, Egido J, Vivanco F, Jensen ON. Characterization of HSP27 phosphorylation sites in human atherosclerotic plaque secretome. *Methods Mol Biol*. 2007; 357:151-63

29. Ambrogio C, Voena C, Manazza AD, Riera L, Piva R, Barberis L, Costa C, Tarone G, Defilippi P, Hirsch E, **Boeri Erba E**, Mohammed S, Jensen ON, Palestro G, Inghirami G, Chiarle R.
p130cas mediates the transforming properties of the anaplastic lymphoma kinase. *Blood*. 2005; 106: 3907-3916

30. **Boeri Erba E**, Bergatto E, Cabodi S, Silengo L, Tarone G, Defilippi P, Jensen ON.
Systematic analysis of the EGF receptor by mass spectrometry reveals stimulation-dependent multisite phosphorylation. *Mol Cell Proteomics*. 2005; 4:1107-1121

31. Cabodi S, Moro L, Bergatto E, **Boeri Erba E**, Di Stefano P, Turco E, Tarone G, Defilippi P. Integrin regulation of epidermal growth factor (EGF) receptor and of EGF-dependent responses. *Biochem Soc Trans*. 2004; 32(Pt3):438-42

32. Di Stefano P, Cabodi S, **Boeri Erba E**, Margaria V, Bergatto E, Giuffrida MG, Silengo L, Tarone G, Turco E, Defilippi P.
p140Cap (p130Cas associated protein) as a new tyrosine phosphorylated protein involved in cell spreading. *Mol Biol Cell*. 2004; 15:787-800

33. Moro L, Dolce L, Cabodi S, Bergatto E, **Boeri Erba E**, Smeriglio M, Turco E, Retta F, Giuffrida G,. Venturino M, Godovac-Zimmermann J, Conti A, Schaefer E, Beguinot L, Tacchetti C, Gaggini P, Silengo L, Tarone G, Defilippi P.
Integrin-induced epidermal growth factor (EGF) receptor activation requires c-Src and p130Cas and leads to phosphorylation of specific EGF receptor tyrosines. *J Biol Chem*. 2002; 277:9405-9414

Part II: previous research (2001-2018)

Since 2001 I have been interested in applying mass spectrometry (MS) to investigate relevant biological topics in an innovative manner. Below I describe my main projects carried out until 2011 as a PhD student and a postdoctoral fellow, working in Italy, Denmark, UK and Switzerland. As a PhD student, I investigated the epidermal growth factor receptor and an oncosuppressor known as p140. As a postdoctoral fellow, I utilised native MS to study several protein complexes and answer different questions regarding their evolution, their two dimensional organisation and their behaviour in the gas phase.
In 2011 I moved to the IBS in Grenoble where I continue investigating macromolecular complexes using MS. Specifically, I studied the human Transthyretin and the yeast CAF1 complex. Moreover, I pursued MS methodological studies to enhance the ability of MS to analyse challenging samples, contributing to determine the structure and function of large assemblies. Over the years, my enthusiasm about the benefit of applying MS in many distinct projects has remained intact.

Brief introduction about biological MS: the different approaches I used during my career

Since the beginning of my PhD studies I used MS to provide novel insights into biological systems. Over time, I have analysed mainly proteins using different MS-based approaches. I briefly illustrate the four main approaches utilised in my work. Namely, the approaches are “bottom-up proteomics”, “native MS”, “top-down MS” and “NEMS” (i.e., nanomechanical resonators arrays for mass sensing).

During my PhD I used “**bottom-up proteomics**” [1]. This means that target proteins are proteolysed and then peptides are analysed by MS. This approach has been introduced at the beginning of the 90s [2] and it has enormously expanded the use of MS to investigate biological events. A turning point has been the development of two techniques: electrospray ionisation and matrix-assisted laser desorption/ionisation. They allow the ionisation of large, non-volatile and thermally-labile molecules such as proteins (Figure 1). Remarkably, these ionisation methods have enabled MS to investigate intact biomolecules.

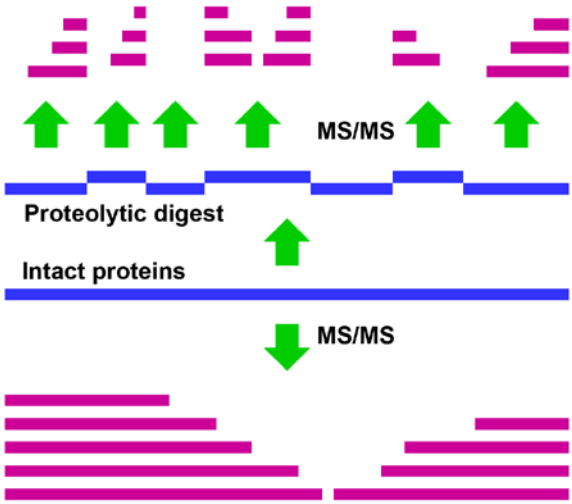


Figure 1. Analysis of proteins by MS. In bottom-up proteomics (top), proteins are proteolysed using an enzyme (such as trypsin), and the resulting peptides are analysed in two stages. In the first stage (known as “MS mode”), the masses of the intact tryptic peptides are determined. In the second stage (aka “MS/MS mode” or “tandem MS mode”), these peptide ions are further fragmented to obtain information about the sequence of the digested proteins and their post-translational modifications (PTMs). Using the top-down approach (bottom), intact protein ions are fragmented in the mass spectrometer. This provides the molecular mass of the proteins and of their fragment ions. Figure was adapted from [3] and Wikipedia.

The electrospray ionisation (ESI) allows the ionisation of analytes from a solution [4]. In very simple terms, a high voltage (1-5 kV) is applied to the liquid supplied through a metal needle. The applied voltage induces an electric field between the needle and the inlet of the mass spectrometer. This high field causes the exposed liquid surface to form an extended structure (known as the Taylor cone [5]) at the needle tip. When the Coulombic repulsion of the surface charges exceeds the surface tension of the solution, charged droplets are emitted from the Taylor cone. Droplets desolvate as they travel down pressure and potential difference toward the mass analyser.

Matrix-assisted laser desorption/ionisation (MALDI) presents a key ionisation approach to analyse biomolecules [6]. MALDI requires the use of a laser and a “matrix”, an organic acid with a strong absorption at the ultraviolet (or less commonly infrared) wavelength [7]. The analyte is mixed with a large molar excess of the matrix and then the analyte-matrix dried mix is irradiated by intense pulses of UV-laser for a short time. The analyte molecules are ionised, leading to the formation of the typical $[M+H]^+$ species. The discovery of these two

ionisation techniques granted to some of their developers the Nobel Prize in Chemistry in 2002.

"Bottom-up proteomics" has become a fundamental tool to identify many proteins with high accuracy and sensitivity. However, important information is missed when proteins are proteolysed. For example, the ability to assess the presence of different isoforms is lost. Moreover, "bottom-up proteomics" requires denaturing conditions and information regarding the stoichiometry of non-covalent assemblies and the direct protein-protein interactions are inaccessible.

Taking into account these limitations of "bottom-up proteomics", for my post-doc I decided to continue learning about biological MS, and to focus my attention on the investigation of macromolecular complexes. Therefore, I joined the laboratory of Prof. Dame Carol Robinson (Cambridge University, UK) where I learned about "**native MS**". This approach allows you to preserve non-covalent bonds and to assess the stoichiometry of macromolecular assemblies and their two-dimensional interactions.

The mass of intact macromolecular assemblies, their stoichiometry, od[8] (Figure 2). Unlike other approaches used to study inter-subunit interactions (e.g., pull-down assays, yeast two-hybrid), native MS directly monitors the interactions between subunits, without the need to introduce mutations or tags. Native MS allows one to gain structural information on macromolecular complexes because it is able to analyse a wide mass range (from 20 Da to 18 MDa [9]) with high sensitivity and accuracy. Native MS presents excellent selectivity. This means that several species with different masses can be simultaneously analysed. This represents a great advantage compared to many structural biology approaches, which require homogenous samples.

"Native MS" represents an exciting field where I continue working because this MS technique represents an emerging tool in "integrative structural biology" (see below). Indeed, native MS helps determining the architecture of large and heterogeneous macromolecular complexes when high resolution data are missing [8, 10].

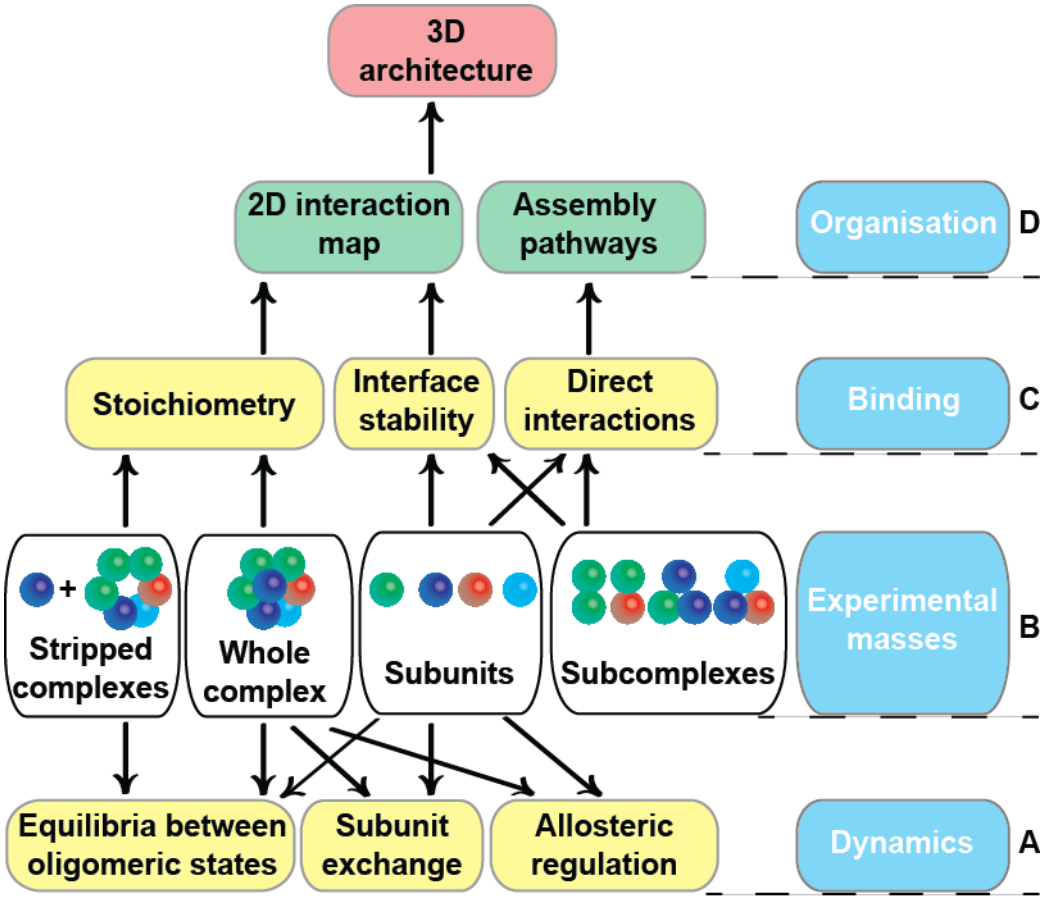


Figure 2. Native MS provides diverse information about macromolecular complexes. Primary data concerning the masses of the holo-complex, subcomplexes, and individual subunits (row B) provide access to a wealth of secondary information, such as binding interactions and dynamic properties (rows A, C and D). Rows B-C: The experimental masses of individual subunits of a complex are determined following chromatographic separation under denaturing conditions [10b]. The use of MS conditions optimised for preserving noncovalent interactions (i.e., native MS) allows the mass of the intact complex to be measured. The combined data about subunits and a holo-complex reveal the stoichiometry of subunits, which can be further confirmed by a series of tandem MS spectra (i.e., "gas-phase dissociation" to generate stripped complexes [11]). The use of perturbing conditions or destabilizing agents ("in-solution dissociation") allows one to generate overlapping subcomplexes (dimers, trimers, etc.), whose composition reveals the direct interactions between subunits [12] and the stability of interfaces [13]. Row D: The above data allow one to determine the two dimensional (2D) interaction network of subunits within a protein complex and, when combined with structural data (EM or SAXS envelopes), to deduce the 3D architecture of a macromolecular assembly (e.g., [14]). In addition, individual subunits (or subcomplexes) can be mixed in solution and a mass shift can be detected if a

(sub-)complex is generated, allowing the assembly pathway of a macromolecular complex to be identified. Row A: Native MS can provide information regarding the dynamic behavior of complexes. First, it can reveal the presence of different oligomeric states and monitor changes in equilibrium induced by different solution pH values and concentrations [15]. Second, the subunit composition of intact complexes can be varied and monitored as a function of time by incubating light and heavy isoforms of a protein (e.g., labeled with ^{13}C and ^{15}N). The kinetics of subunit exchange can reveal distinct pathways for wild-type and mutant proteins (e.g., involved in amyloidosis [16], or assessing the effect of substrates and products). Third, native MS can relatively quantify the populations of macromolecular complexes containing different numbers of bound molecules, providing information about allostery [17]. Figure from [8].

When you analyse a non-covalent complex, the accurate mass of the subunits forming the assembly should be assessed. Therefore, part of our work is dedicated to determine the mass of the proteins under denaturing conditions. When the biomolecules are soluble and smaller than 100 kDa, we use high performance liquid chromatography (HPLC) coupled with an ESI instrument [10b]. The use of HPLC-MS has the advantage of separating different species within a sample (Figure 3).

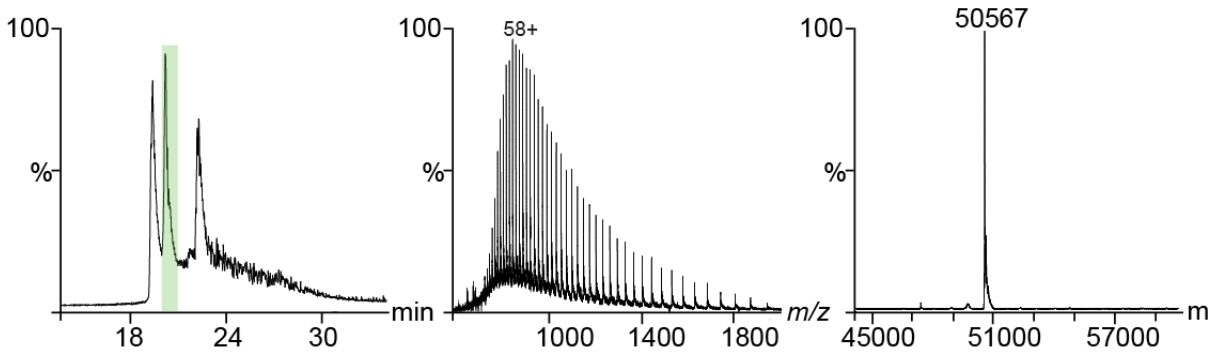


Figure 3. Example of a HPLC-ESI-MS experiment. A trimeric complex (known as the yCAF1 complex [18]) was analysed by MS under denaturing conditions. Left panel: the three subunits of the complex (i.e., Cac1, Cac2, and Cac3) were chromatographically separated, yielding three distinct retention times (RTs). Central panel: MS spectrum of Cac3 (RT = 20.1 min). Right panel: deconvoluted spectrum of Cac3. The average mass was calculated from the protein sequence (50525 Da). The detected mass (50567 Da) was higher than the expected one; this mass difference indicated the presence of an acetylation mark (42 Da) [10b].

When proteins are larger than 100 kDa or they are membrane proteins requiring detergents to be soluble, we use a MALDI instrument [19]. Recently, the IBS has acquired a new MALDI mass spectrometer that allows us to perform "**top-down MS**" experiments [1] (Figure 1). This type of MS is based on the fragmentation of intact proteins inside a mass spectrometer. Fragmentation experiments allow the sequencing of N- and C- termini of proteins [20]. A sequence can be determined because two following neighbour fragments differ by the mass of one amino acid. The fragmentation data provide information not only about the protein primary sequences, but also about their PTMs. In simple terms, the modifications are highlighted by a mass shift compared to the mass of a "naked" amino acid.

In addition to the MS-based methods I am using and I described above, I would like to briefly mention **NEMS** (nanomechanical resonators arrays for mass sensing of bioparticles) (Figure 4) [21].

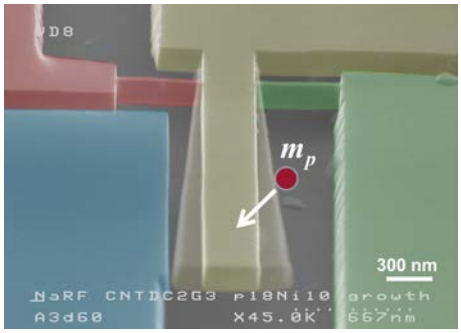


Figure 4. NEMS principle. Specifically designed nanostructures (e.g. 1.8 µm x 120 nm x 100 nm) present high mechanical vibrational frequencies. When a molecule is added onto this structure, the total mass of the resonator is affected and thus the resonance frequency of the oscillation changes. By detecting this frequency shift, the mass of added species on the device can be measured.

This innovative type of MS is developed by colleagues at BIG (Biosciences and Biotechnology Institute of Grenoble) and at LETI (Institute for electronics and information technologies, Grenoble). We are a group of scientists, who aim to use native MS and NEMS to provide novel insights on biomedically-relevant specimens such as bacteriophages (see part III, below).

The emerging role of MS in structural biology

Biological MS has a wide application from the identification of tumour biomarkers [22] to the characterisation of metabolites [23]. Since two decades MS has emerged as a tool in integrative structural biology [8, 10, 24] (Figure 5). MS represents a valuable approach in hybrid structural investigations thanks to its sensitivity, wide applicability, and speed of analysis. Moreover, MS selectivity allows scientists to simultaneously analyse and separate several species with different masses. Mass selectivity represents a key advantage compared to other structural biology approaches that require homogenous samples.

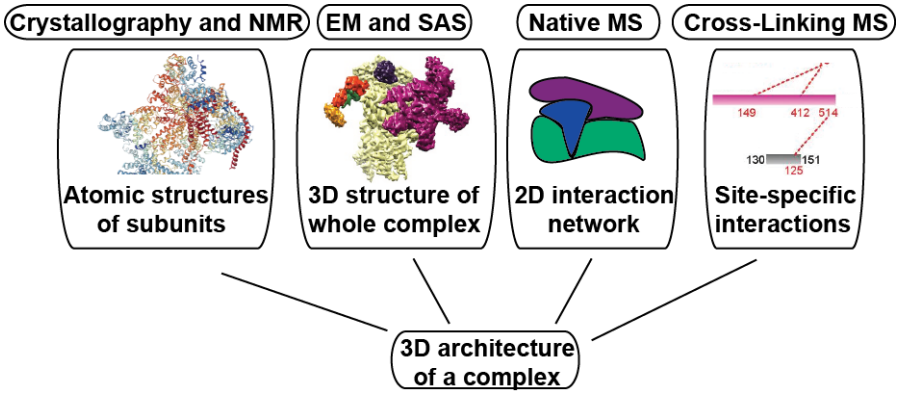


Figure 5. Integrative structural biology. When scientists investigate the structure of dynamic macromolecular assemblies, they require data obtained through a 'hybrid approach'. They can use distinct methods including high-resolution approaches (such as NMR) and lower-resolution methods such as small-angle X-ray scattering and MS (e.g., cross-linking MS [25] and native MS).

To date, only a few laboratories work on "structural MS" and there is a great need for improvement in terms of sample preparation and instrumentation. In the next years the MS laboratory at the IBS aims to further strengthen the role of MS in structural biology (see part III).

University of Turin (Italy) and University of Southern Denmark (2001-2006)

Study of the epidermal growth factor receptor by bottom-up MS-based proteomics

Introduction

The epidermal growth factor receptor (EGFR) plays a fundamental role in human development, proliferation, and differentiation. When activated, this 180 kDa protein tyrosine

kinase determines mitogenesis, apoptosis and migration. In response to ligand binding or integrin-mediated adhesion, it is phosphorylated and mediates cellular responses in normal biological processes and in pathological states. Mutations of the EGFR sites and its overexpression are implicated in a variety of cancers, including mammary carcinomas, squamous carcinomas, and glioblastomas [26]. Monoclonal antibodies inhibiting the EGFR are used to treat metastatic colorectal cancer [26].

My research

- **Boeri Erba E**, *et al.* Quantitation of multisite EGF receptor phosphorylation using mass spectrometry and a novel normalization approach. J Proteome Res. 2007; 6(7):2768-85
- **Boeri Erba E**, *et al.* Systematic analysis of the EGF receptor by mass spectrometry reveals stimulation-dependent multisite phosphorylation. Mol Cell Proteomics. 2005; 4:1107-1121
- Moro L, Dolce L, Cabodi S, Bergatto E, **Boeri Erba E**, *et al.* Integrin-induced epidermal growth factor (EGF) receptor activation requires c-Src and p130Cas and leads to phosphorylation of specific EGF receptor tyrosines. J Biol Chem. 2002; 277:9405-9414.

Using bottom up MS-based proteomics, I performed a sensitive, qualitative and quantitative study of multiple phosphorylation sites of the human EGFR in normal and cancer cells. I investigated the EGFR phosphorylation when the receptor was activated by several types of stimulation (e.g. by EGF, integrin-dependent adhesion and suspension). For example, we assessed the phosphorylation profile of EGFR using MS combined with SILAC (i.e., stable isotope labeling by amino acids in cell culture) [27]. We investigated the influence of cell adhesion on the EGF-dependent phosphorylation. In particular, we detected and relative quantified the phosphorylation of eleven EGFR residues. The levels of phosphorylation obtained in adherent and suspended cells treated with EGF were compared to those obtained in sodium pervanadate- (Na-perv) treated cells (Figure 6). The extent of phosphorylation induced by Na-perv was defined as a reference for 100% phosphorylation.
My study revealed that the EGF stimulation of **adherent cells** induced higher levels of tyrosine phosphorylation and lower levels of serine and threonine phosphorylation than the EGF stimulation of **suspended cells**.
To conclude, our findings support the hypothesis that integrin-mediated adhesion modulates the phosphorylation of plasma membrane receptor tyrosine kinases, which controls EGF-induced signal transduction. Since integrins are involved in tumour initiation and progression

[28], it is necessary to better characterise integrins, receptor tyrosine kinases and their crosstalk, which controls cell migration and invasion.

		differential phosphorylation		fold change	experimental	control
P sites	mw	Na-perv/Adh+ EGF	Na-perv/Susp+ EGF	Adh+EGF/Susp+ EGF	Adh+EGF/Susp+ EGF	Na-perv/Na-perv
pT669	2113.05	3.8 σ 0.2	1.5 σ 0.1	0.4	ND	ND
pY845	1709.77	16.5 σ 0.1	>30	>1.8	1.5 σ 0.1	1.0 σ 0.1
	2147.07	17.7 σ 0.6	>30	>1.7	ND	1.1 σ 0.1
	1406.66	16.5 σ 0.1	>30	>1.8	ND	1.0 σ 0.1
pS967	1659.68	12.1 σ 0.3	4.4 σ 0.1	0.4	ND	1.0 σ 0.1
	1643.69	6.0 σ 1.2	4.1 σ 0.2	0.7	0.3 σ 0.1	1.0 σ 0.1
pY974	1659.68	5.3 σ 0.2	10.8 σ 0.3	2	1.2 σ 0.2	0.9 σ 0.1
	1643.69	5.0 σ 0.1	10.8 σ 0.3	2.2	ND	1.0 σ 0.1
pY1086	2478.18	1.3 σ 0.1	2.5 σ 0.1	1.9	1.7 σ 0.1	1.2 σ 0.1
pY1148	2314.99	1.6 σ 0.1	2.5 σ 0.1	1.6	1.8 σ 0.2	0.9 σ 0.1
pY1173	1289.53	1.4 σ 0.1	2.5 σ 0.1	1.8	1.9 σ 0.1	0.9 σ 0.1

Figure 6. Quantitation of EGFR phosphopeptides. Eleven EGFR phosphopeptides were analysed by HPLC-MS. In a first experiment, the phosphorylation of EGFR induced by EGF in adherent cells (Adh+EGF) was compared to the maximal level of phosphorylation obtained by sodium pervanadate (Na-perv) (see the "Na-perv/Adh+EGF" column). The phosphorylation level induced by Na-perv was considered as a batch mark of 100% phosphorylation. In a second experiment, the level of phosphorylation obtained in suspended cells treated with EGF was compared to those obtained in Na-perv-treated cells (see the "Susp+EGF/Na-perv" column). To determine the relative fold-change of the phosphopeptide abundance, the experimental ratios Na-perv/suspension+EGF were divided by the corresponding ratios Na-perv/adhesion+EGF (see "fold-change, Adh+EGF/Susp+EGF" column). In a third experiment, adherent and suspended cells both stimulated with EGF were directly compared ("experimental, Adh+EGF/Susp+EGF" column). The use of "fold-change" of Adh+EGF/Susp+EGF allowed me to investigate eleven EGFR phosphopeptides, whereas the "experimental Adh+EGF/Susp+EGF" data only provided information about six phosphopeptides. The standard deviation of three measurements (indicated as δ) was reported. Figure from [27].

Identification and characterisation of the oncosuppressor p140Cap by proteomics

Introduction

In cancer cells, cellular transformation, migration and invasion are controlled by integrins, growth factor receptors and special proteins named "integrin signalling adaptors" [28]. Among these adaptors, p130Cas coordinates cell adhesion signalling based on receptor tyrosine kinases. Its overexpression determines anti-estrogen resistance in breast cancer cells.

My research

- Alfieri A, Sorokina O, Adrait A, Angelini C, Russo I, Morellato A, Matteoli M, Menna E, **Boeri Erba E**, et al. Synaptic interactome mining reveals p140Cap as a new hub for PSD proteins involved in psychiatric and neurological disorders. *Front Mol Neurosci.* 2017 Jun 30;10:212. doi: 10.3389/fnmol.2017.00212.
- Repetto D*, Aramu S*, **Boeri Erba E***, *et al.* Identification of phosphorylation sites on the adaptor protein p140Cap: p140Cap mediates binding to the C-terminal Src kinase through the EPLYA and EGLYA phosphorylated tyrosines. *Plos one,* 2013;8:e54931;* authors equally contributed to this investigation
- Ambrogio C, Voena C, Manazza AD, Riera L, Piva R, Barberis L, Costa C, Tarone G, Defilippi P, Hirsch E, **Boeri Erba E**, *et al.* p130cas mediates the transforming properties of the anaplastic lymphoma kinase. *Blood*. 2005; 106: 3907-3916
- Di Stefano P, Cabodi S, **Boeri Erba E**, *et al.* p140Cap (p130Cas associated protein) as a new tyrosine phosphorylated protein involved in cell spreading. *Mol Biol Cell*. 2004; 15:787-800.

Taking into account the importance of p130 for cell migration and invasion in transformed cells, we identified a novel p130Cas-associated protein and we named it as p140Cap, (i.e., p140 Cas-associated protein) known also as “SRC kinase signaling inhibitor 1” (SRCIN1). P140Cap is an adaptor protein involved in integrin- and EGF-dependent signalling [29]. It is expressed in some primary breast carcinomas and its presence significantly correlates with unfavourable overall survival [30].

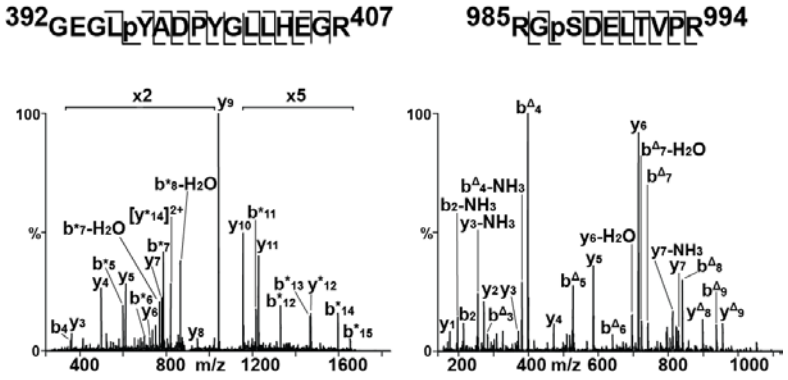


Figure 7. Spectra of *in vivo* phosphorylated p140 peptides. The protein p140 was immunoprecipitated from human cells, proteolysed and then its peptides were analysed by LC-MS. Each peptide was fragmented inside the mass spectrometer in order to allocate the position of the phosphorylated sites [31].

I studied the p140 *in vivo* phosphorylation identifying its modification sites (see Figure 7 and [32]). Further investigation carried out by the Defilippi *et al.* showed that p140Cap is an important negative regulator of the proto-oncogene tyrosine-protein kinase Src [33].

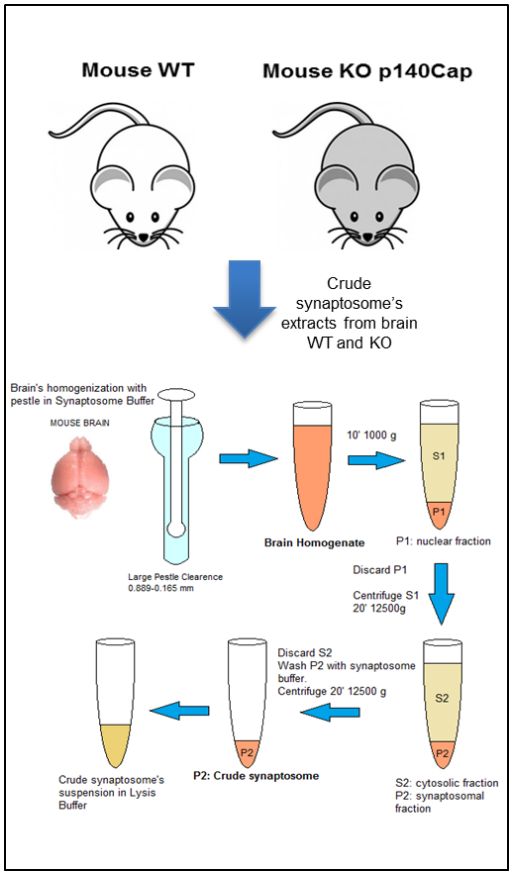


Figure 8. Workflow to characterise the p140Cap synaptic interactome. We analysed the p140Cap synaptic interactome using co-immunoprecipitation from crude mouse synaptosomes and MS-based proteomics. We identified 351 p140Cap interactors and found that they cluster to subcomplexes mostly located in the postsynaptic density (PSD). Figure from [34].

Src kinase activity increases in several tumour tissues and tumour cell lines such as colon carcinoma cells. Overall, our work shed new light on the function of p140Cap as a player in suppressing tumour cell properties, regulating Src kinase.

P140 acts not only as a tumour suppressor, but also plays a role in synaptic plasticity. It localises to neuron dendritic spines, where it contributes to microtubule and actin dynamics [32]. p140Cap knock-out mice show defects in short- and long-term memory, as well as in long-term potentiation and depression [32]. In 2016 we performed a MS-based proteomic screening to determine the interaction partners of p140 in mouse synapses (Figure 8).

Synaptic interactome reveals that p140Cap is a key hub for proteins involved in psychiatric and neurological disorders [34].

University of Cambridge (UK) and Swiss Federal Institute of Technology (ETH), Zurich (CH) (2006-2011)

Investigation of intact protein complexes by native MS

Introduction

In the last 20 years native MS has emerged as a remarkable tool in structural biology for investigating the organization of non-covalent complexes, when high-resolution structural data are unavailable [8, 10b, 35] (Figure 2). Compared to other approaches used to investigate protein-protein interactions (e.g., yeast two-hybrid), native MS monitors the interactions between subunits, without the need of tags or crosslinking reactions. It is sensitive, accurate and selective.

Below I describe my work performed using native MS. By investigating several protein complexes, I studied different aspects, ranging from the evolution of macromolecular assemblies to the gas-phase behaviour of complexes.

My research

- Levy ED, **Boeri Erba E**, et. Assembly reflects evolution of protein complexes. *Nature*. 2008; 453(7199):1262-5.

When I was at the University of Cambridge, I became interested in studying the evolution of homo-oligomeric protein complexes. Using native, MS I examined the structure, assembly and disassembly of homo-oligomeric complexes (Figure 9). In particular, I experimentally generated building blocks of homo-oligomeric complexes using appropriate destabilising conditions. In collaboration with Teichmann's group (Medical Research Council, Cambridge), we demonstrated that the assembly of macromolecular assemblies reflects their evolution *in vivo*. In other words, an intermediate in the assembly of a protein complex *in vitro* (i.e., an experimentally determined building block) is an intermediate in its evolution.

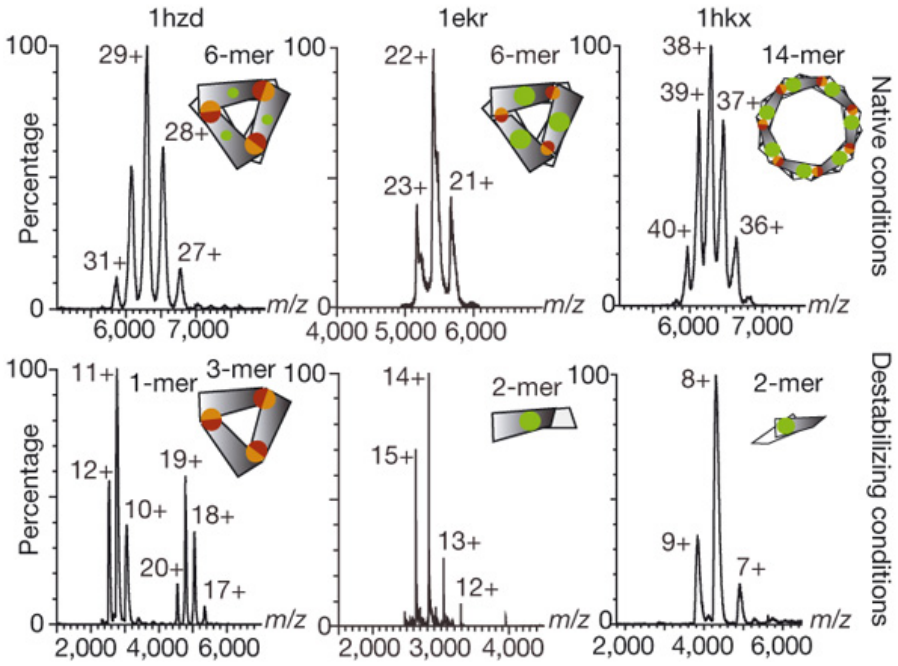


Figure 9. Native mass spectra of intact complexes (top panel) and of subcomplexes obtained after destabilisation in solution (bottom panel). These complexes were previously crystallised. Using destabilisation conditions I was able to selectively break the weakest interactions and preserve the strongest interactions. This means that the building blocks were isolated by native MS. My results were consistent with crystallographic data present in the PDB database. 1hzd: AUH, an RNA binding protein; 1ekr: molybdenum cofactor biosynthesis protein; 1hkx: Ca2+/calmodulin-dependent kinase II [13].

- **Boeri Erba E**, et al. Ion Mobility-Mass Spectrometry Reveals the Influence of Subunit Packing and Charge on the Dissociation of Multi-protein Complexes. *Anal Chem*. 2010; 82(23):9702-10.

As postdoctoral fellow at Cambridge University, I was interested in the behaviour of protein complexes in the gas phase (i.e., inside a mass spectrometer). Using native MS and ion mobility-MS I investigated the factors influencing the dissociation of protein complexes in the gas phase.

I focused my attention on the dissociation pathways of two dodecameric (12-meric) complexes, HSP16.9 and SP-1, and subjected both protein complexes to tandem MS using collision induced dissociation (CID) (Figure 10). I expected that they behaved similarly given their ring-shape topology (Figure 10). Like many other complexes, HSP16.9 dissociated generating a sequential loss of monomeric subunits. On the contrary, the behaviour of SP-1 was unexpected because there was the ejection of monomers (1-mers), dimers (2-mers), and tetramers (4-mers).

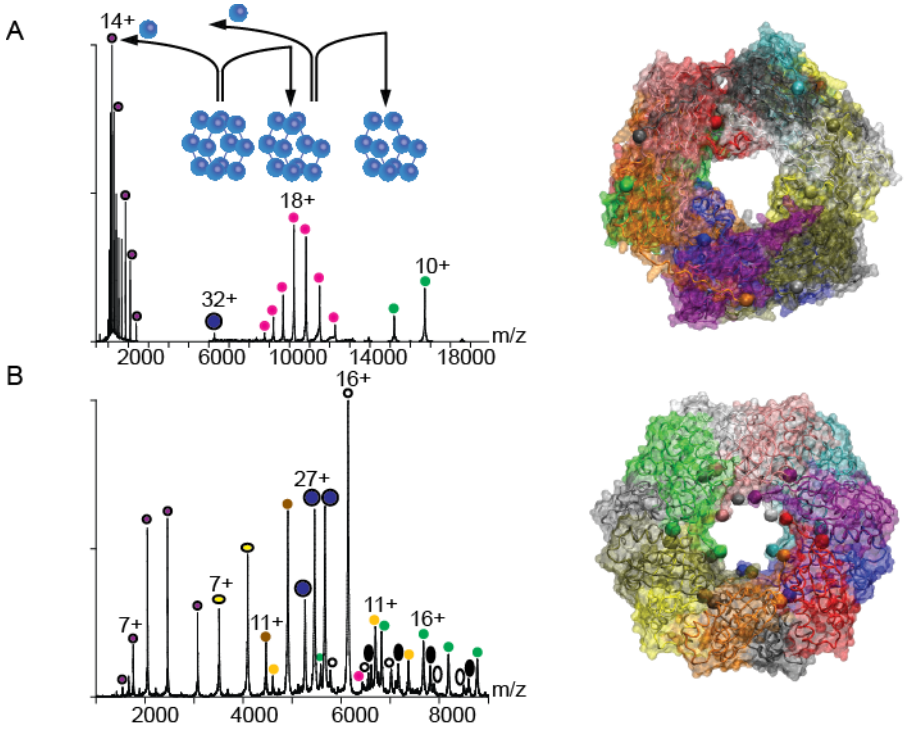


Figure 10. Features of the HSP16.9 and SP-1 macromolecular assemblies. (A) The HSP16.9 protein complex. Left panel: MS/MS spectrum of intact dodecameric (12-meric) HSP16.9 32+ ions. A schematic representation illustrates how this protein complex dissociates by forming lowly charged decamers (10-mers) and undecamers (11-mers) at high *m/z* and highly charged monomers at low *m/z.* This type of dissociation is defined as "typical" because there is sequential loss of monomeric (1-meric) subunits. 12-mer: blue; 11-mer: pink; 10-mer: green; and 1-mer: purple. Right panel: Crystal structure of the HSP16.9 protein complex (PDB code: 1gme). The N- and C-terminal amino acids of each protein within the protein complex are represented as spheres in order to clearly illustrate their positions. The N-termini of chains B, D, F, H, J, and L are disordered and not present in the crystal structure. (B) The SP-1 protein complex. Left panel: MS/MS spectrum of intact 12-meric SP-1 27+ ions. 1-mers, 2-mers, 3-mers, 4-mers, 5-mers and 11-mers, 10-mers, 9-mers, 8-mers, 7-mers, 6-mers are present at low and high *m/z* in the spectrum, respectively. The dissociation of SP-1 is "atypical" due the ejection of dimers (2-mers) and tetramers (4-mers). 12-mer, blue circles; 11-mer, pink circles; 10-mer, green circles; 9-mer, white ellipses; 8-mer, white circles; 7-mer, black ellipses; 6-mer, orange circles; 4-mer, brown circles; 2-mer, yellow ellipses; 1-mer, purple circles. Right panel: The crystal structure of the SP-1 protein complex (PDB code: 1tr0). Figure from [36].

To characterise the behaviour of SP-1, I investigated the correlation between the charge state of the parent ion and the dissociation behaviour of the protein complex. I performed MS/MS experiments and selected the SP-1 ions carrying 35+, 26+ and 23+ charges (Figure 11).

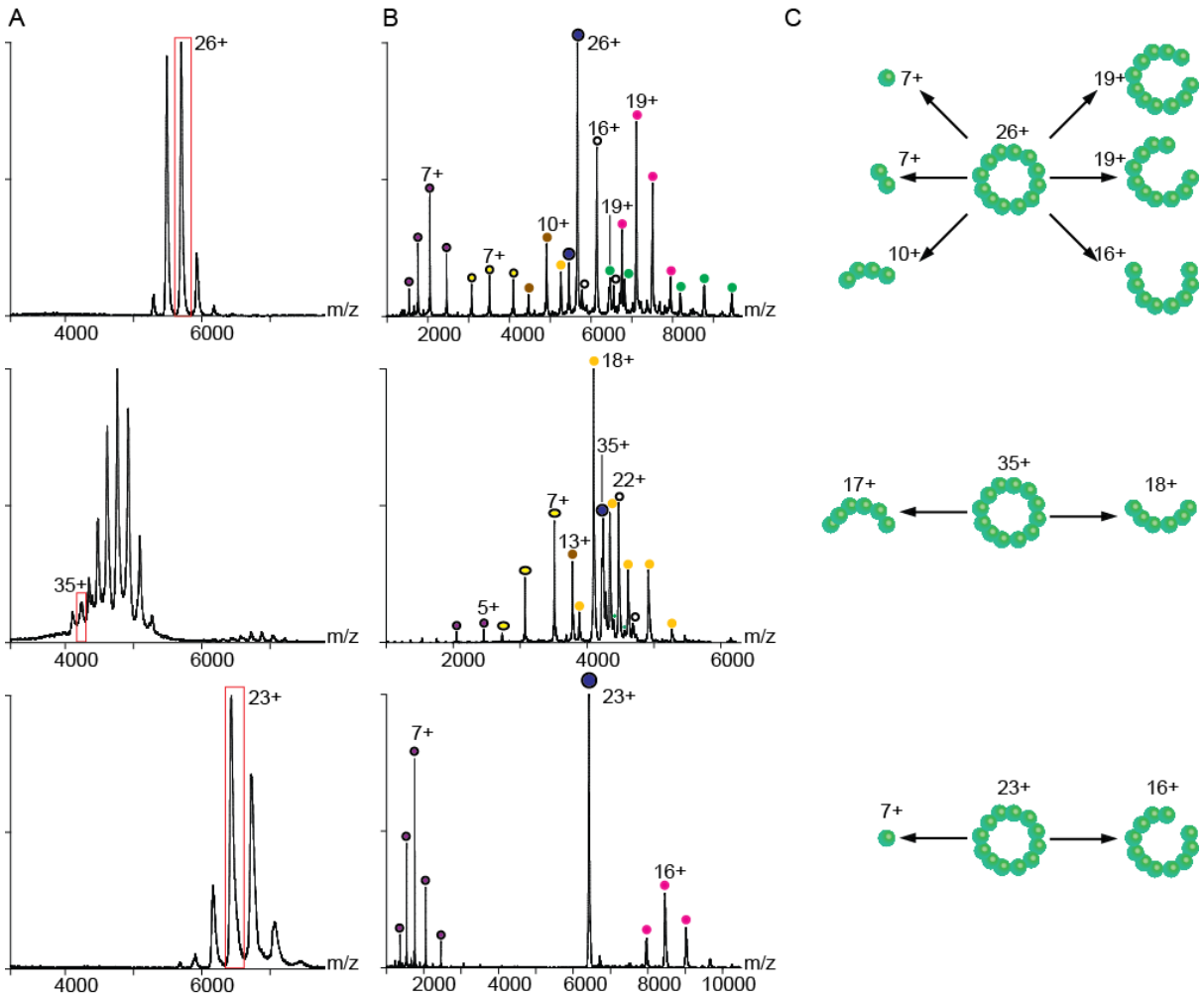


Figure 11. MS and MS/MS spectra of the SP-1 protein complex from different solution conditions. (A) MS spectra of the SP-1 protein complex in three different solution conditions: 100 mM ammonium acetate (top panel), 6.25% sulfolane (middle panel), 2.5% DMSO (lower panel). B) MS/MS spectra of three charge states of the SP-1 protein complex: 26+ (top panel), 35+ (middle panel), 23+ (lower panel). 12-mer, blue circles; 11-mer, pink circles; 10-mer, green circles; 8-mer, white circles; 6-mer, orange circles; 4-mer, brown circles; 2-mer, yellow ellipses; 1-mer, purple circles. (C) Schematic representations illustrate the dominant dissociation pathways when the SP-1 protein complex is in the three different charge states. The SP-1 26+ ions eject 1-mers, 2-mers, and 4-mers, forming 8-mers, 10-mers, and 11-mers. The SP-1 35+ ion dissociates predominantly via the formation of 6-mers. The SP-1 23+ ion ejects 1-mers, generating 11-mers. Figure from [36].

The SP-1 26+ ion dissociated via the ejection of 1-mers, 2-mers, and 4-mers. The dissociation of the 35+ ion generated 6-mers, which carried 17+ and 18+ charges. These 6-mers are half of the mass of the complex and hold half of the charge initially present on the intact protein complex. The SP-1 23+ ion expelled 1-mers, generating 11-mers, showing "typical" behaviour for CID experiments. Even though the SP-1 protein complex was able to dissociate atypically, the reduction of charges in this protein assembly induced it to behave typically. Therefore, the number of charges of the SP-1 protein complex effectively governed its dissociation.

I did further experiments using ion mobility-MS to measure the collisional cross section for several charge states prior to activation and unfolding. I also monitored the behaviour of the different charge states during dissociation into product ions, monitoring their unfolding as well as the energy necessary to convert the different charge state of SP-1 to dissociated and unfolded states. These experiments indicate that charge state of the parent ion highly influences the dissociation and unfolding of the SP-1 complex.

Finally, I examined the crystal structures of the SP-1 and HSP16.9 protein complexes to correlate their dissociation behaviour with their quaternary structure. I considered the location of the N- and C-termini as key factors that initiate unfolding in the gas phase. Indeed, the HSP16.9 protein complex could unfold by charge migration relatively easily since all the C-termini lie on the surface of the complex. On the contrary, the structure of the SP-1 protein complex has structural features that could prevent the ready unfolding of subunits. For example, all the N-termini of the SP-1 protein complex are toward the central cavity formed by the 12-mer, and all the C-termini are buried inside the structure. The N- and C-termini are woven into the protein in such a way (e.g., forming a middle strand in a β sheet) that they cannot be exposed to the surface without the near complete unfolding of the protein.

Overall, my work indicates that quaternary structure and charge density are two key factors governing the dissociation of protein complexes. This moves the scientific community closer to the goal of obtaining structural information on subunit interactions and packing from gas-phase experiments.

- **Boeri Erba E**, et al. Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal Chem.* 2011; 83(24):9251-9.

During my research period at the ETH (Zurich, 2010-2011) I exploited MS **for a sensitive measurement of equilibrium association constant (K_a) of protein-protein interactions**.

This approach aimed to overcome the limitations of traditional approaches such as isothermal titration calorimetry. Using native MS I investigated the association of a noncovalent homo-oligomeric complex (i.e. concanavalin A) (Figure 12). This allowed me not only to obtain K_a on subunit interactions in a very sensitive manner using a small amount of unlabeled sample, but also to gain information on composition, stoichiometry and subunit interactions of a protein complex. Moreover, my investigation supports the hypothesis that native MS measurements of the protein complexes (in the gas-phase) depict the solution-phase composition of macromolecular assemblies. This is shown by the fact that the signal intensity of the different oligomeric states changes, when the solution concentration of concanavalin and/or the buffer pH are modified.

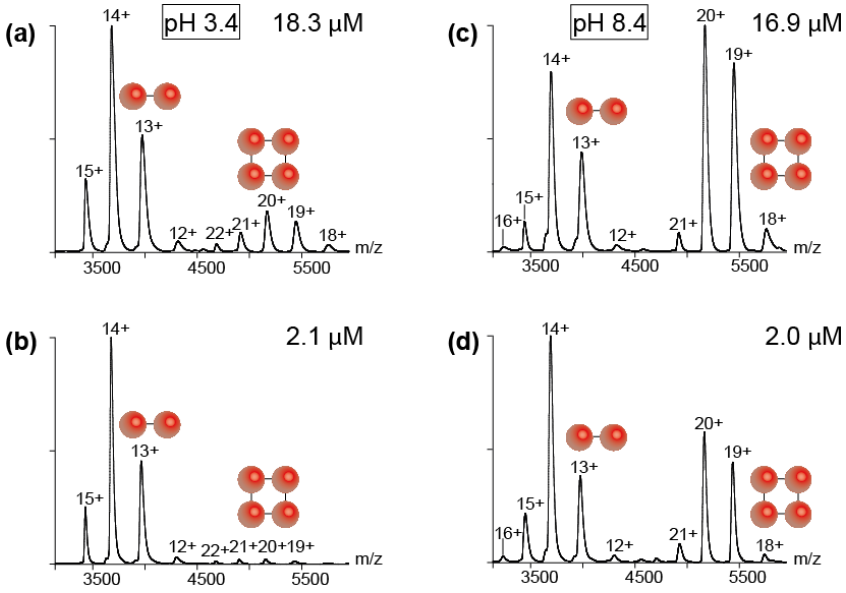


Figure 12. Relative abundance of dimers and tetramers versus solution concentration of concanavalin A. The solution pH highly influences the oligomeric state of concanavalin A. (a) Spectrum of intact concanavalin A at D_0= 18.3 µM and pH 3.4; (b) spectrum at D_0 = 2.1 µM and pH 3.4; (c) spectrum of concanavalin A at D_0 = 16.9 µM and pH 8.4; (d) spectrum of concanavalin A at D_0 = 2.0 µM and pH 8.4. Two spheres indicate the dimeric ions and four spheres indicate tetrameric ions. Figure from [15].

- Sharon M, Mao H, **Boeri Erba E**, et al. Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality. *Structure*. 2009; 17(1):31-40.

The ubiquitin-proteasome proteolysis pathway is essential for many crucial processes such as regulation of transcription and DNA replication [37]. The proteasome is an important cellular assembly that degrades damaged or unnecessary proteins, which are poly-

ubiquitinilated. The COP9 signalosome complex (CSN) controls the ligases that bind the ubiquitin to target proteins. It is composed of eight subunits forming a 450 kDa complex. For a long time, the CSN architecture has been studied to better understand the CSN function.

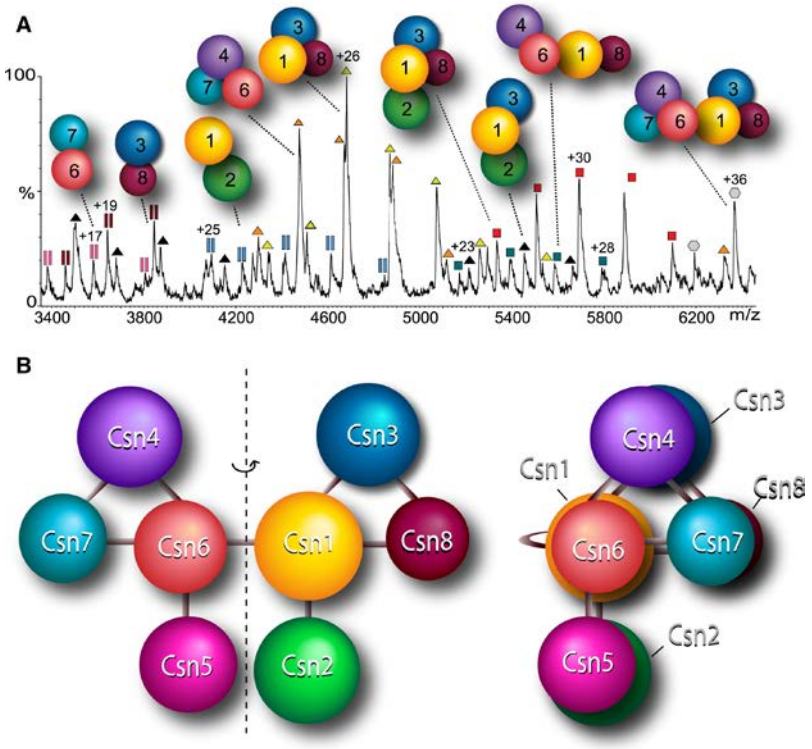


Figure 13. Subunit map of the human CSN Complex. (A) An 3400-6200 *m/z* range of the mass spectrum recorded when the CSN complex was in absence of the Csn5 subunit. Experimental conditions were optimised to induce the formation of subcomplexes (labeled with the topological arrangement of subunits). In particular, 5%–12.5% methanol (v/v) was added to generate subcomplexes. (B) The protein interaction map of the eight-subunit assembly was determined through the composition of the different subcomplexes. Two distinct modules containing Csn1/2/3/8 and Csn4/5/6/7 were identified. A single interaction between Csn1 and Csn6 could link the two modules. In each module (Csn1/3/8 or Csn4/6/7) each subunit interacted with the other two. On the contrary, the most peripheral subunits, Csn2 and Csn5, bound only a subunit. Figure was from [38].

Our native MS-based investigation allowed us to propose a model of the CSN structural organisation even though high-resolution structural data were not available. In particular, we built a comprehensive map of the interactions between CSN subunits (Figure 13). Our model also provided a convincing explanation to many *in vivo* observations of CSN subcomplexes. These small complexes may be biologically significant because CSN consists of two distinct submodules (Csn1/2/3/8 and Csn4/5/6/7) with possible independent roles. Overall, our

modular model represented a step forward in the scientific deciphering of the CSN structure and function.

Institute of Structural Biology (IBS), Grenoble (2011-2018)

Structural studies of biomedically-relevant protein complexes

Transthyretin amyloidogenesis

- Yee AW, Moulin M, Breteau N, Haertlein M, Mitchell EP, Cooper JB, **Boeri Erba E***, Forsyth VT*. Impact of Deuteration on the Assembly Kinetics of Transthyretin Monitored by Native Mass Spectrometry and Implications for Amyloidoses. Angew Chem Int Ed Engl. 2016; 55(32):9292-6. doi: 10.1002/anie.201602747, *: corresponding author

Transthyretin (TTR) is a homotetrameric protein responsible for transporting thyroxine and retinol and is implicated in amyloidosis. Aggregation of the wild-type (WT) TTR causes senile systemic amyloidosis, predominantly affecting the heart. Mutated TTR causes familial amyloid polyneuropathy and/or cardiomyopathy. Over the years, numerous techniques have been applied to study TTR and its misfolding that ultimately leads to the deposition of insoluble amyloid fibrils. Currently there are over 200 structures of TTR. Unfortunately none of these data fully explains the mechanism of TTR amyloid fibril formation.
We used MS, thioflavin T fluorescence, and crystallography to study the effect of deuteration on the WT TTR (Figure 14). We showed that while the X-ray structures of unlabeled (H-TTR) and deuterium-labeled TTR (D-TTR) are essentially identical, subunit exchange kinetics and amyloid formation are accelerated for the deuterated protein. However, a slower subunit exchange is noted in deuterated solvent, reflecting the poorer solubility of non-polar protein side chains in such an environment. These observations are important not only for the interpretation of kinetic studies involving deuteration, but also for uncovering the mechanism of amyloidosis. Undoubtedly, details of the solvent structure and amino acid protonation states of WT and mutated TTR are crucial for better understanding amyloidosis.

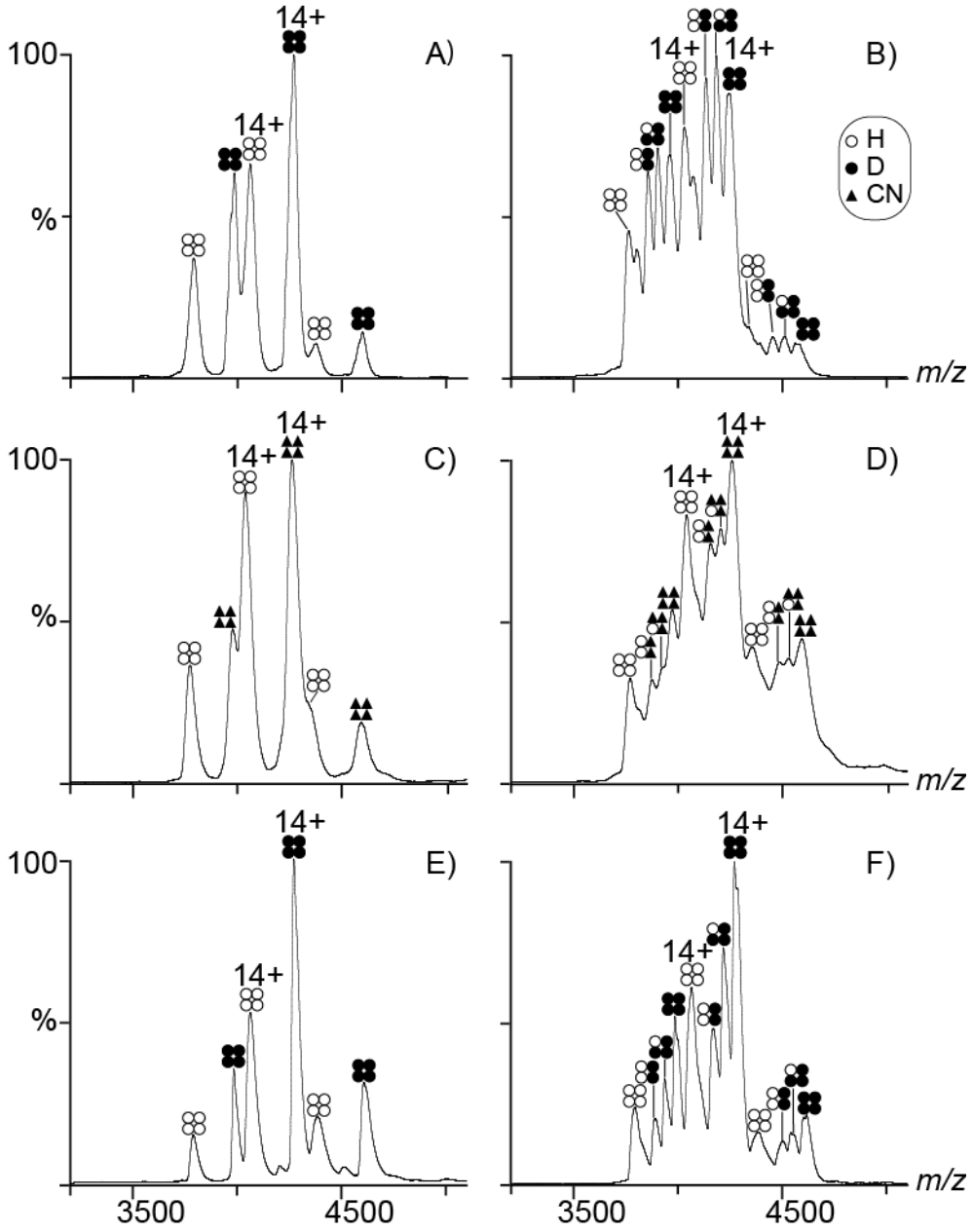


Figure 14. Native MS spectra of mixed unlabeled and isotope-labeled WT-TTR. A) 38 minutes after mixing H-TTR (unlabelled protein) and D-TTR (labelled). In the spectrum six peaks are present matching the masses of the 4H and 4D TTR homo-tetramers. B) TTR spectrum after 8 days of the H-TTR and D-TTR mixing. C) TTR spectrum 7 minutes after mixing 4H and 4CN (TTR labelled with ^{13}C ^{15}N isotopes). D) The same reaction was monitored after 10 days. H-ammonium acetate was used for the samples shown in A–D. E) 7 minutes after mixing HTTR and DTTR. F) TTR spectrum after 8.8 days of mixing HTTR and DTTR. D-ammonium acetate dissolved in D_2O was used for the samples in (E) and (F).

- Yee AW, Aldeghi M, Blakeley MP, Ostermann A, Mas P, Moulin M, De Sanctis D, , Bowler M, Mueller-Diekman C, Mitchell EP, Haertlein M, de Groot BL, **Boeri Erba E**, Forsyth VT*. A structural and dynamical analysis of transthyretin: a molecular mechanism for amyloidogenesis.

Many TTR mutations are implicated in fatal forms of amyloidosis. However, some mutations offer *in vivo* protective effects. To understand the molecular basis underlying the different behaviours of TTR mutants, we utilised native MS, neutron crystallography, and modelling studies. In particular, MS was used to follow the subunit exchange dynamics of TTR over several days so that the effect of two mutations (S52P and T119M), and that of the drug tafamidis, on the rate of dissociation of the tetramer species could be compared (Figure 15). These experiments relied on the use of hydrogenated (H) and deuterated (D) TTR to track otherwise identical protein chains. H- and D-TTR were mixed in an equimolar ratio and monitored over time to observe the rate of dissociation of the homo-tetramers (4H, 4D) and the rate of formation of the hetero-tetramers (2H2D, 1H3D, 3H1D) (Figure 15). Since we have shown that the kinetics of TTR subunit exchange is susceptible to isotope effects, a deuterated TTR species [i.e., D-S52P] was used as a reference to allow a meaningful comparison of the TTR variants. Overall, the MS experiments revealed that the S52P mutation increases the rate of dissociation of TTR tetramers by about a factor of three as compared with the WT. In strong contrast, both the T119M mutation and tafamidis effectively abolished tetramer dissociation.

By combining MS, neutron crystallography, and modelling studies, we proposed a mechanism, whereby TTR amyloid fibrils are generated. Fibrils may be formed via an equilibrium of partially unfolded monomers that proceeds toward amyloidogenic forms. Unfolding events may originate at a specific loop of the TTR monomers, and destabilising mutations in this region increases the speed of TTR fibrillation. Moreover, the TTR protective mutants may stabilise the protein in way similar to the binding of Tafamidis [39] used to treat adult patients affected by TTR amyloidosis to delay the peripheral neurologic impairment.

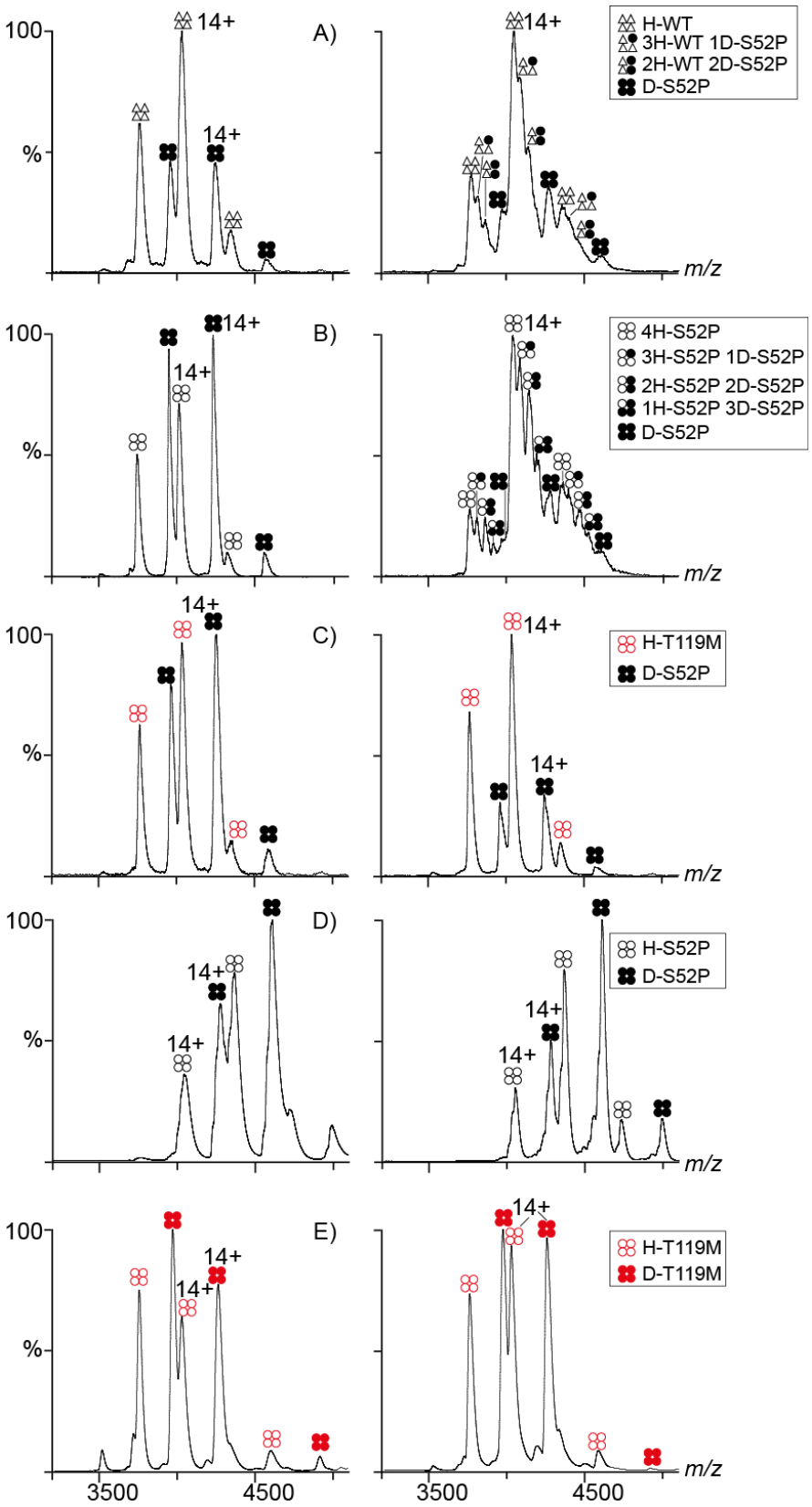


Figure 15. MS spectra of the subunit exchange of TTR. Reactions between deuterated S52P and hydrogenated (A) WT, (B) S52P, (C) T119M, and (D) S52P bound to tafamidis, were monitored. The shown spectra were recorded at the beginning of the reaction (left

panel) and after 11 days (right panel). It should be noted that the DMSO, in which the tafamidis was dissolved, reduced the charge states of TTR. (E) MS spectra recorded during the exchange between hydrogenated T119M and deuterated T119M. This was a control to assess that T119M tetramers were stable and not subjected to any subunit exchange. Unpublished data.

The assembly of eukaryotic nucleosome

- Sauer P, Timm J, Liu D, Sitbon D, **Boeri Erba E**, Velours C, Mücke N, Langowski J, Ochsenbein F, Almouzni G, Panne D. Insights into the molecular architecture and histone H3-H4 deposition mechanism of yeast chromatin assembly factor 1. eLife 2017; 6. pii: e23474. doi: 10.7554/eLife.23474

In eukaryotic cells, very long DNA molecules are tightly packaged and wrapped around histone proteins to form structures known as nucleosomes. While this is a useful way to store DNA, it also makes it inaccessible to many molecules that activate genes, copy DNA or perform other important cell processes. To make the DNA accessible, cells selectively disassemble particular nucleosomes and remove the histones. After DNA replication, the nucleosomes must reassemble to repackage the DNA.

A single nucleosome contains four pairs of histones (i.e., H2A, H2B, H3 and H4). Histone chaperones assemble nucleosomes in a two-step process. First, two of H3-H4 pairs (i.e., a tetramer) interact with DNA and form a macromolecular complex. Then, two pairs of H2A, H2B bind to complete the nucleosome. An enzyme called CAF1 is known to load the H3-H4 tetramers on DNA as the DNA is being copied. In this way, nucleosomes pack the newly made DNA. How CAF1 deposits H3-H4 tetramers onto the DNA is not known.

The EMBL laboratory lead by Daniel Panne, the MS laboratory of the IBS and scientists at Institut Curie and at Université Paris-Sud investigated the yeast CAF1 function using native MS (Figure 16), Size Exclusion Chromatography-Multiple Angle Laser Light Scattering (SEC-MALLS), Analytical UltraCentrifugation (AUC) and Small-Angle X-ray Scattering (SAXS). We showed that each CAF1 enzyme is able to bind a single H3-H4 pair. Then, two CAF1 enzymes bind to DNA and attach a H3-H4 tetramer onto it. The tetramer is formed in this way to ensure that the histones are correctly delivered to DNA after the DNA has been copied.

Overall, these findings explain the first steps of the nucleosome formation. Our data indicate the sequence of key events that take place when CAF1 attaches H3-H4 tetramers onto DNA (Figure 17).

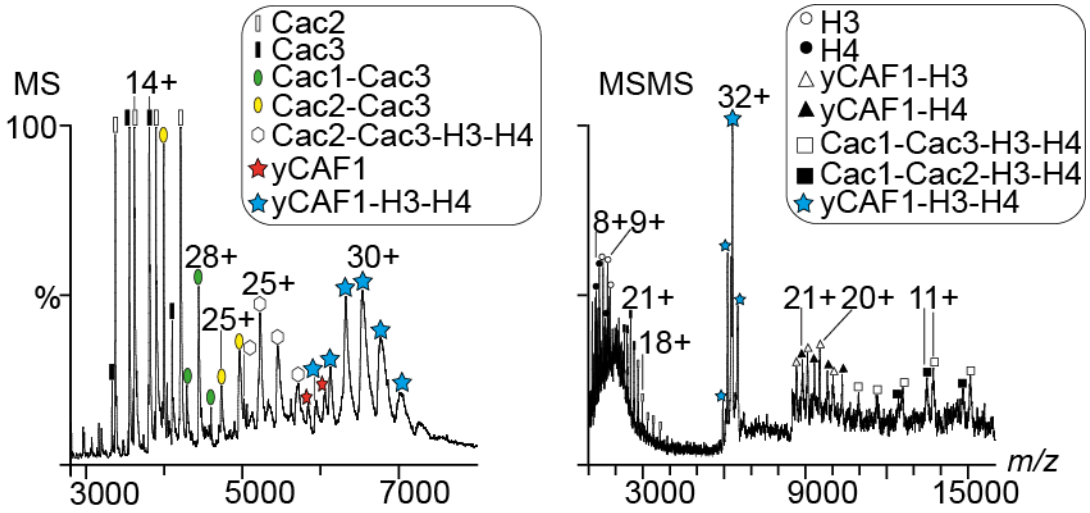


Figure 16. MS and MS/MS spectra of the yCAF1-H3-H4. Left panel: MS spectrum of the heteroassembly. In addition to the intact CAF1, trimer in the unbound state (174 kDa) and bound to H3-H4 (201 kDa) (labeled as red and cyan stars, respectively), dimers such as Cac1-Cac3 (green ellipses) and Cac2-Cac3 (yellow ellipses), trimers such as Cac2-Cac3-H3-H4 (white hexagon), and single monomeric proteins (blank and white rectangles) were also detected. Most likely, these subcomplexes were present in solution and not generated in the gas phase. Right panel: MS/MS spectrum of the 32+ ions confirmed a 1:1:1:1:1 stoichiometry of the yCAF1-H3-H4 complex. Figure from [18].

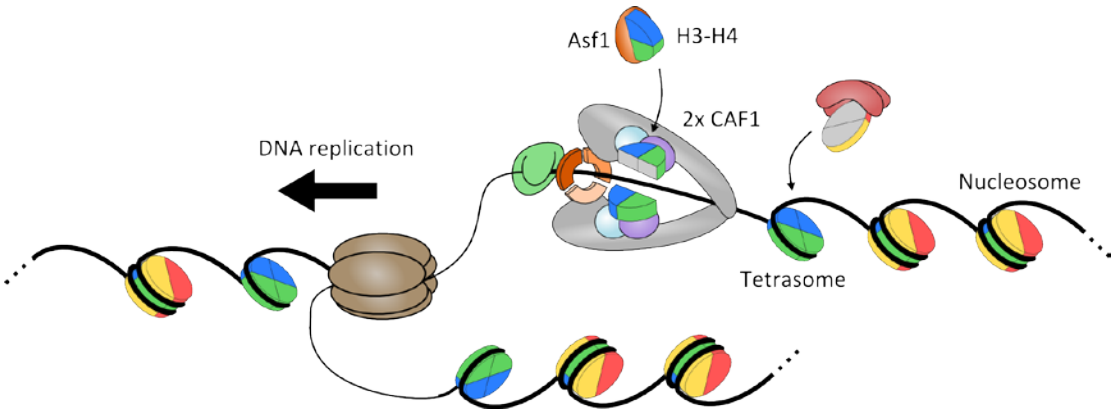


Figure 17. Model for yCAF1 recruitment and H3-H4 deposition. A yCAF1 complex (labelled in grey, light blue and violet) is loaded with dimeric H3-H4 through association of yAsf1. Two yCAF1-H3-H4 complexes bind cooperatively to an extended DNA sequence and deposit two H3-H4 dimers and form tetrasomes. Additional proteins (such as the one labelled in orange) are required for deposition of H3-H4 tetramers. Then, H2A-H2B chaperones recognise the H3-H4 tetrasome and deposit two copies of H2A-H2B to form a complete nucleosome (i.e., 2 H2A, 2 H2B, 2 H3 and 2 H4).

Methodological development for studying intact proteins and macromolecular complexes by MS

Biological MS helps to answer many different scientific questions. To do that, it is necessary to continue improving MS approaches. A key aspect is the ”sample preparation”. This indicates the steps necessary to prepare the analyte prior to the MS analyses. Below I describe our work done to improve the investigation of intact proteins and macromolecular assemblies.

My research

- Signor L, **Boeri Erba E***. Matrix-assisted Laser Desorption/Ionization Time of Flight Mass Spectrometric Analysis of Intact Proteins Larger than 100 kDa. *Journal of visualized experiments*, 2013; 79: e50635.

We developed a novel sample preparation for the MALDI MS analysis of large proteins and protein complexes. MALDI-MS represents an alternative to ESI-MS because its performance is less affected by buffer components, detergents and contaminants. Two key factors influence the quality of the MALDI spectra: the matrix and the technique used for the matrix deposition. As matrix, we utilised a mixture of two acids: 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (α-CHCA). As deposition technique, we used the thin layer method where the first layer was α-CHCA dissolved in acetone. Compared to the other approaches traditionally used (e.g., sinapinic acid as matrix, SA), we observed better resolution and higher sensitivity (Figure 18).

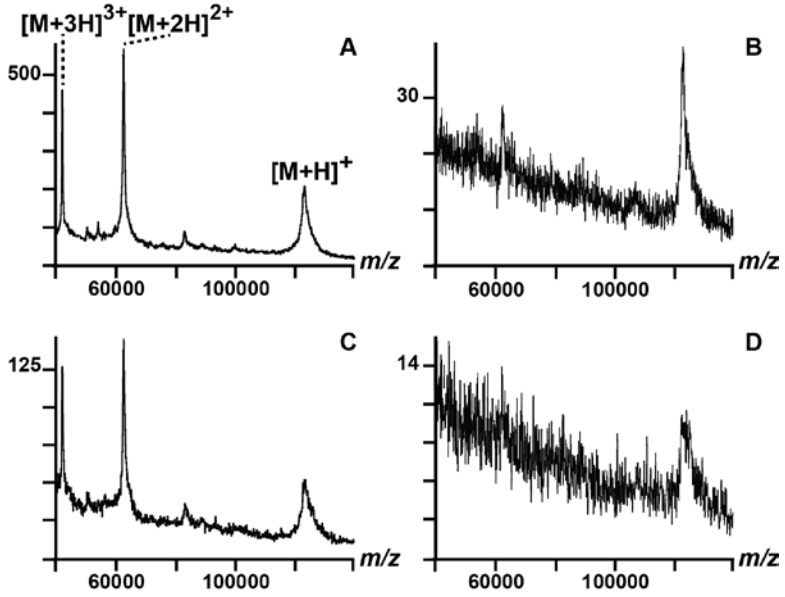


Figure 18. MALDI-TOF analysis of the intact Chromosome region maintenance 1 protein (Crm 1). The molecular weight of this protein is 123386 Da. A) 1 pmole of Crm1 was analysed using the DHB_CHCA mix as matrix. B) amount: 1 pmole; matrix: SA. C) amount:

0.5 pmole; matrix: DHB_CHCA. D) amount: 0.5 pmole; matrix: SA. The signal of the peaks, expressed in an arbitrary intensity unit scale, is normalized to the maximum value present in each spectrum.

- **Boeri Erba E***, Klein PA, Signor L. Combining a NHS ester and glutaraldehyde improves crosslinking prior to MALDI MS analysis of intact protein complexes. *J Mass Spectrom. 2015 Oct;50(10):1114-9. doi: 10.1002/jms.3626*

We established a novel crosslinking (XLing) approach [25b] to study macromolecular assemblies using MALDI-MS. We tested two different crosslinkers (XLers) [i.e. bis(sulfosuccinimidyl)suberate known as BS^3 and glutaraldehyde], used separately or in combination, utilising gentle agitation and ultracentrifugation (Figure 19).

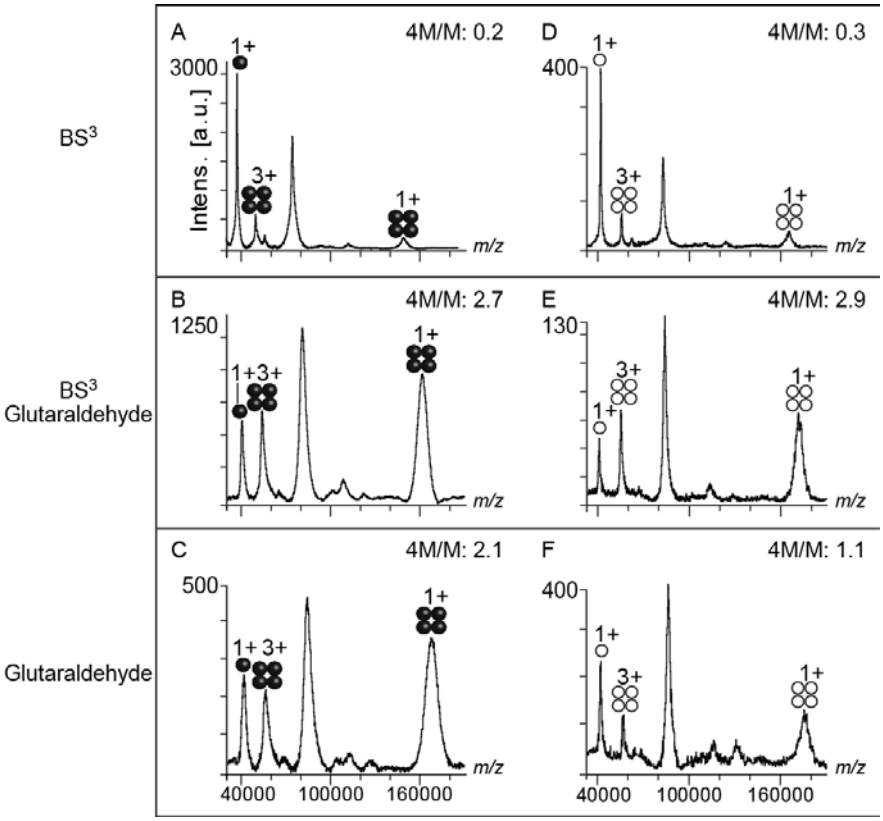


Figure 19. Combining two XLers using gentle agitation in a thermomixer at pH 7.2. We XLed the ADH using A) BS^3 alone, C) glutaraldehyde alone or B) a combination of the two reagents. The aldolase was XLed utilizing D) BS^3, F) glutaraldehyde or E) both. The ADH and aldolase subunits are labeled as black and white spheres, respectively. The ratio 4M/M is a semi-quantitative indicator of XLing efficiency. Four spheres show the tetrameric peaks. The ratio 4M/M is the sum of the intensities of the tetrameric ions (i.e. $[4M+ H]^{1+}$ and [4M+

3H]$^{3+}$) divided by M, the intensity of the peak at the lowest *m/z* (i.e. less XLed, [M+ H]$^{1+}$). The intensity of each spectrum is expressed as an arbitrary unit (a.u.). Intens = intensity. Figure from [40].

We XLed two tetrameric protein complexes [i.e. alcohol dehydrogenase (ADH) and aldolase] larger than 140 kDa at two pH values (i.e., 7.2 and 8.0). Our data shows that the pH influenced the XLing when using a single XLer. The use of two combined XLers was more efficient than using a single reagent. In particular, two combined XLers determined a higher degree of XLing and lower mass shift. This observation suggested a ranking in the availability of the target amino acids. First, BS^3 binds residues at specific distances, then glutaraldehyde links still available amino acids. Ultracentrifugation and gentle agitation both provided similar degrees of XLing, but the former determined a lower mass increment resulting from redundant XLing. Overall, we developed a dual XLing approach useful for determining mass and stoichiometry of protein assemblies by MALDI-MS.

- Macek P, Kerfah R, **Boeri Erba E**, Crublet E Moriscot C, Schoehn G, Amero C, Boisbouvier J. Unraveling self-assembly pathways of the 468-kDa proteolytic machine TET2. *Sci Adv. 2017 Apr 7;3(4):e1601601. doi: 10.1126/sciadv.1601601.*

Using native MS, nuclear magnetic resonance spectroscopy (NMR), and electron microscopy (EM), we investigated the assembly pathway of Tetrahedral aminopeptidase 2 (TET2), a 468 kDa homo-complex. During an assembly, subunits undergo appropriate conformational rearrangements to form a supramolecular structure. Tracking such rearrangements represents a challenge because of the low-abundance of intermediate states changing over time. Although the self-assembly is a time-dependent process occurring at molecular level, its current understanding originates from static structures, low resolution techniques and modelling. NMR spectroscopy is able to monitor structural changes at the atomic level in real-time, however its size and time resolution constraints represent a practical challenge. We utilised a methyl specific labelling in an otherwise deuterated protein combined with relaxation optimized, fast acquisition real-time NMR to overcome both size and time scale limits [41]. Time-resolved EM was used to identify the shape of large oligomeric intermediates appearing during the self-assembly pathway. However, the initial intermediate in the TET2 self-assembly was of small size, undetectable by EM.

To better monitor the first events during the self-assembly, we used “isotopic hybridization” and native MS (Figure 20). By doing that, we determined the mass and stoichiometry of the hybrid complex and identified the intermediate present at the beginning of the self-assembly

reaction. Our results indicated that the intermediate was a monomer. Overall, MS well complemented NMR and EM used to probe structural changes of subunits during the assembly of a large macromolecular complex.

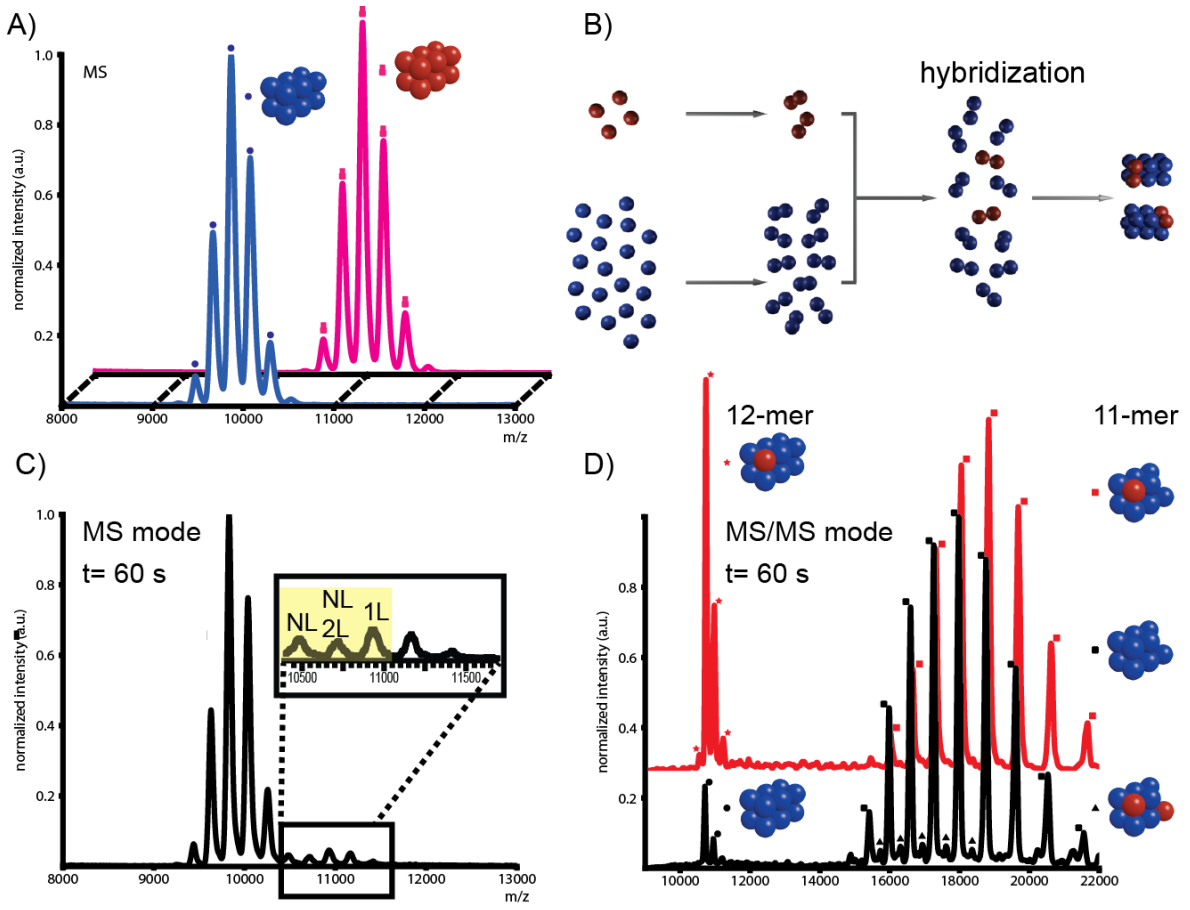


Figure 20. Self-assembly mechanism of TET2. A) MS spectra of the non-labelled (NL) TET2 12-mer (blue) and the [^{15}N, ^{13}C, ^{2}H] fully labelled (L) (red) TET2. The signals of the NL and L dodecamers (12-mers) do not overlap due to a significant mass difference. B) To do "isotopic hybridization" experiments we utilised NL and L 1-mers (i.e., disassembled complexes) as starting material. 60 seconds after the beginning of the self-assembly reaction, the NL and L were mixed. In this way, L intermediates were incorporated into final "isotopically hybrid" complexes which were further analysed by tandem MS. C) A MS spectrum of an isotopically hybrid 12-mer. D) Tandem MS spectra of two isotopically hybrid TET2 12-mers whose *m/z* were 10700 and 10900, respectively.

Part III: current and future research at the IBS

Since two decades MS has emerged as a tool in integrative structural biology thanks to its sensitivity, wide applicability, selectivity and speed of analysis [10a, 24]. In the next years the MS laboratory at the IBS aims to develop the analysis of antibodies because they are becoming pivotal drugs and MS can play an important role in their characterisation. We will continue working on the native MS, focusing our research on a large RNA polymerase complex involved in the chloroplast biogenesis. We also intend to investigate bacteriophages using a combination of native MS and NEMS.

Characterisation of therapeutic antibodies using MS

Introduction

Antibodies: their importance in the immunoresponse and as drugs

Biopharmaceuticals are medicinal products containing active biological agents utilised in *vivo* diagnostics, prophylaxis and therapy [42]. Among them, antibodies (Abs) are emerging as key drugs in the case of inflammatory diseases (such as rheumatoid arthritis) and cancer [43]. The 2018 Nobel Prize in Physiology or Medicine was awarded to two immunologists, James Allison and Tasuku Honjo. They pioneered immunotherapy, which harnesses the body's immune system to fight cancer.

Monoclonal Abs are the largest class of recombinant proteins utilised in human therapy and *in vivo* imaging of many diseases [44]. Antibody-drug conjugate (ADC) composed of an Ab bound to cytotoxic drug become increasingly important as oncology therapeutics [43b, 45].

A key advantage of Abs is their great accuracy to reach their target because they specifically bind diseased cells, neglecting healthy ones. Abs cause the diseased cells to self-destruct or activate the immune system against them [43a]. Naturally Abs are part of the adaptive immune response. They are synthesised in response to the presence of molecules that are extraneous to the animal host (i.e., antigens, Ags). Specifically, Abs are secreted by B lymphocytes cells to protects animals from infectious organisms and their toxic products. After secretion, Abs circulate throughout the blood and lymph where they specifically bind Ags. Then Ab-Ag complexes are removed from circulation by macrophage phagocytosis or elicit other responses from the immune system.

Abs are part of a large family of glycoproteins that present common features. Abs are also known as immunoglobulins (Igs). Igs are group of animal proteins that are structurally similar and may or may not provide immunity. Therefore, all the Abs are Igs, but not all the Igs are Abs. However, the two terms are often used interchangeably. Structurally, Abs are characterised by specific features (see below). Functionally, Abs are able to bind not only

Ags, but also specialised proteins or cells of the immune system. Monoclonal Abs (mAbs) present so called “monovalent affinity”. This means that they bind an identical epitope. MAbs are generated by identical immune cells that are all clones of a unique parent cell. On the contrary, polyclonal antibodies (pAbs) are secreted by different B cell lineages. They are a pool of Igs, reacting against a specific Ag, but identifying different epitopes of the Ag.

The structure of Abs

Abs are formed of one or more copies of a basic unit that resembles a “Y”. Y-like units are composed of two ≈25 kDa polypeptide chains (light chains) and of two ≈50 kDa heavy chains [46]. These four chains are linked by covalent and non-covalent interactions to form a Y-like unit.

The Five Immunoglobulin (Ig) Classes

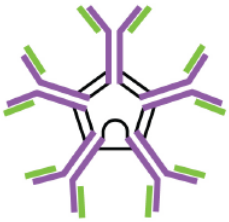
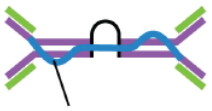
	IgM pentamer	IgG monomer	Secretory IgA dimer	IgE monomer	IgD monomer
					
Heavy chains	μ	γ	α	ε	δ
Number of antigen binding sites	10	2	4	2	2
Molecular weight (Daltons)	900,000	150,000	385,000	200,000	180,000
Percentage of total antibody in serum	6%	80%	13%	0.002%	1%

Figure 21. The different classes of Igs and their features. Figure adapted from Lakna Panawala’s work.

Abs can be classified into five families (IgG, IgM, Ig A, IgE, IgD) according to the number of Y-like units and to the type of heavy chain they contain. Regarding the number of Y-like units that join to form a complete Ig, it varies according to the type of Ab. For instance, IgGs are formed by a single Y-like unit with two Ag-binding sites (paratopes) (Figures 21 and 22). Instead, IgMs are composed of five Y units carrying ten Ag sites. Both light and heavy chains usually cooperate to form the Ag-binding surface.

Each type of Ig has a specific heavy chain known as γ chain (for IgG), μ chain (for IgM), α chain (for IgA), ε chain (for IgE) and δ chain (for IgD). The different heavy chains give distinctive conformations to the Abs, influencing their reactivity during the immunoresponse. Regarding the light chain in mammals, there are only two types, named as lambda (λ) and kappa (κ).

Among the five different types of Igs, the most used ones in therapy are the IgG (≈150 kDa) [46-47]. The light and heavy chains of IgGs have variable (V) sequences at their N-termini and constant (C) sequences at their C-termini (Figure 22). Specifically, each heavy chain presents a variable region (VH), three constant domains (CH1, CH2, CH3), and a "hinge region" localised between CH1 and CH2 (Figure 22) [46, 48]. Heavy chains are heavily glycosylated and this glycosylation affects Ab function and pharmacokinetics [49]. Each light chain consists of a constant domain (CL) and a variable domain (VL). Two heavy chain–light chain heterodimers (HL) are bound via disulfide bonds in the hinge region and via non-covalent interactions between the CH3 domains into a single antibody molecule (H2L2).

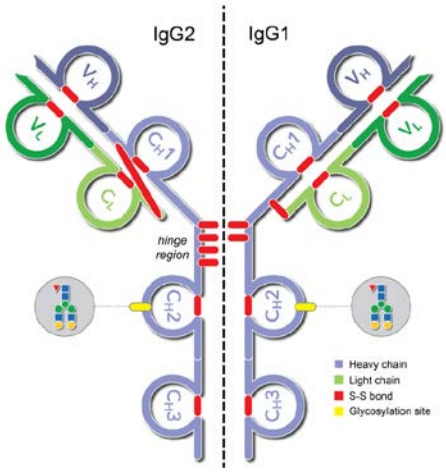


Figure 22. IgG structure. Adapted from [50]

The different domains of IgGs can be classified also according to the fragments that are generated by enzymatic proteolysis using papain and pepsin [47]. Treatment with the first enzyme yields two kinds of products, Fab and Fc fragments (Figure 23). Fab (~50 kDa) is the region of the Abs that binds Ag (Fab = fragment ag-binding). It is formed by the light chain associated with the VH and CH1 domains [46]. Fc (~50 kDa) is composed the lower hinge region, the CH2 and the CH3 [46]. It cannot bind Ag and is crystallisable [46]. Treatment with pepsin yields a single fragment, the F(ab')2. Using LysC-based limited proteolysis and reduction, the fragments Fc, Fab, 1/2Fc, Fd and light chain are generated [51]. Fd is a portion of the heavy chain, which is included in the Fab fragment.

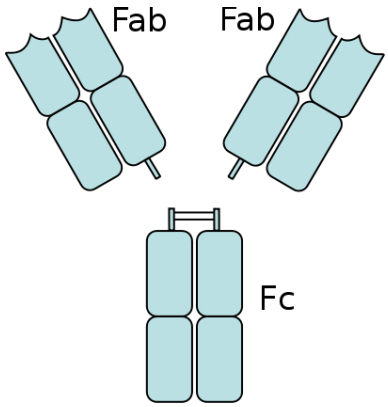


Figure 23. An antibody proteolysed by papain. It yields three fragments: two Fab fragments and one Fc fragment. Adapted from Wikipedia.

Manufacturing recombinant mAbs

Over the years, several distinct approaches have been developed to produce mAb [44, 52]. Three ways of mAb manufacturing are briefly described. (i) Traditionally, a mouse is injected with an Ag to elicit the animal immune response. Ab-producing B cells are extracted from the mouse spleen, then fused with fast-growing cancer cells (i.e., a myeloma), to generate a hybrid cell line called hybridoma, which can be grown in culture [53]. Finally, the Abs are isolated from hybridomas. (ii) Alternatively, phage display can be used [52a]. In this case, Ab sequences are inserted in bacteriophage DNA. Phages are selected from a population by exploiting their binding to the Ag. Then, the DNA from selected phages is transfected into *E. coli* and Abs are produced [52a]. (iii) Mammalian expression vectors and Chinese hamster ovary (CHO) cells represent a popular choice for the expression and production of recombinant Abs [52a, 54].

When mAbs are manufactured, the codons of the mAb transgene can be expressed in different cell systems and determine the mAb primary sequence [44]. The primary sequence of a mAb is critical for its function, but also PTMs are of key importance because they determine mAb heterogeneity, affecting the mAb stability, solubility, pharmacokinetics (PK), pharmacodynamics (PD), potency, and biological activities [47]. Over the years, the structure−function relationships of Abs have been studied to identify Ab microvariants and to investigate their influence on Ab features such as their ability to bind Ags. Structure−function relationships have been helpful to increase Ab homogeneity and to optimise the “chemistry, manufacture, and control” (CMC) development of preclinical Abs using genetic engineering. For instance, mutations affecting instability or aggregation have been removed. Moreover,

hinge-stabilized or aglycosylated IgG4 have been designed to improve Ab pharmaceutical properties [47].

Regulatory agencies such as European Medicines Agency (EMA) and Food and Drug Administration (FDA) require comprehensive characterisation and validation of primary sequence and PTMs of mAbs because these features influence Ab function and efficacy (for example see the following EMA guideline, EMA/CHMP/BWP/532517/2008). MS represents an emerging tool to study Abs [47].

The role of MS to analyse Abs

At the IBS we aim to develop sensitive, robust and rapid methods to characterise Abs using MS [47]. Our work focuses on major aspects, including (i) Ab mass heterogeneity and identification of primary amino acid sequence variants, (ii) localisation of disulfide bridges and PTMs, (iii) characterisation of antibody-drug conjugates (ADCs), (iv) investigation of epitope-paratope interactions.

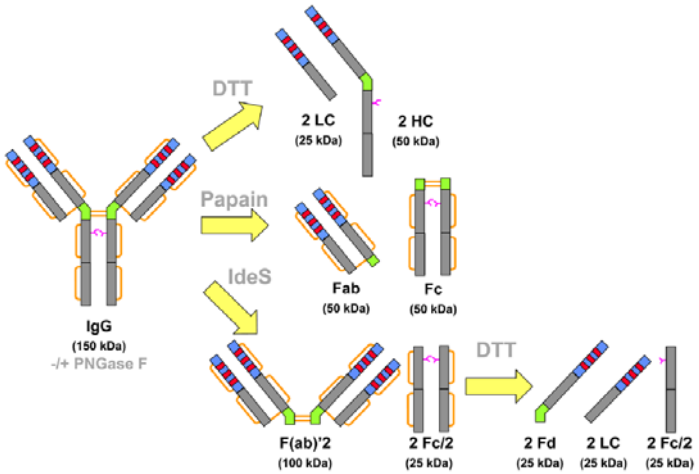


Figure 24. Abs can be characterised using MS according to a "middle-down" approach [55]. Abs are analysed intact before and after deglycosylation treatment using the enzyme PNGase. Abs are also investigated as fragments generated after enzymatic digestion and/or after disulfide bridge reduction. Adapted from [47].

(i) Ab mass heterogeneity and primary amino acid sequence variants.

These Ab features are of key importance because they influence the function of Abs such as their PK and reproducibility of the treatment. Utilising ESI- and MALDI-mass spectrometers, we intend to analyse Abs a) intact, b) after reduction of disulfide bonds, (also called an S-S bond, or disulfide bridge), after proteolysis with c) papain and d) immunoglobulin degrading

enzyme from Streptococcus pyogenes (IdeS [56], trade name FabRICATOR) (Figures 24 and 25).

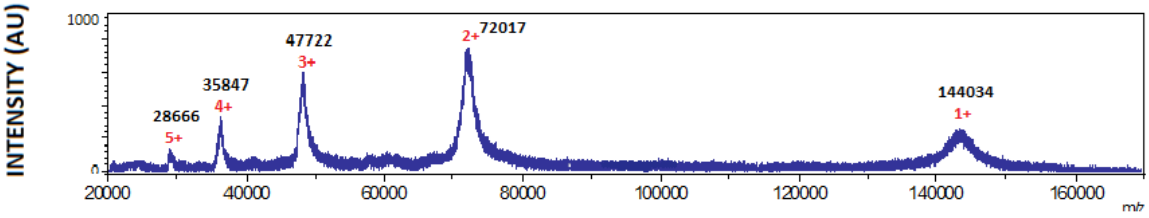


Figure 25. An intact Ab was analysed by MALDI-TOF. This analysis generated Ab peaks with five distinct charge states ranging from 1+ to 5+. The MALDI-TOF approach is sensitive because a picomole of sample is sufficient for a successful analysis. However, the resolution is suboptimal. Moreover, it is difficult to assess the charge of the species and not all the peaks could be assigned with high confidence. Therefore, ESI-MS complements the MALDI-based approach.

Regarding the characterisation of intact Abs, the sample will be analysed glycosylated and after the treatment with Peptide:N-glycosidase F (known as PNGaseF) [57] (Figure 24). This enzyme cleaves the bond between an N-linked glycan and an asparagine residue. It deminates the amide nitrogen of the asparagine, generating an aspartic acid. Enzymatic deglycosylation of Abs provides information about the contribution of the glycosylation to the heterogeneity of Abs. Moreover, deglycosyled sample provide information about primary amino acid sequence variants (Figure 26). This information is important to characterize the batch-to-batch variability and to assess the effect of suboptimal storage conditions.

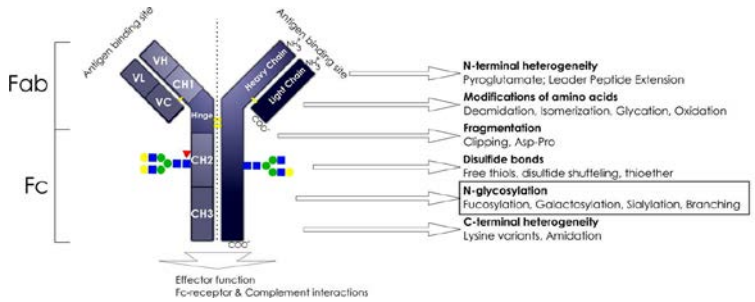


Figure 26. Possible modifications and alterations of an IgG. Figure adapted from [49a]

(ii) Assignment of disulfide bridges and PTMs.

The comparison of masses of Abs before and after reduction of disulfide bonds provides the number of the disulfide bridges. Further MS analyses indicate the position of the bonds. This

important information is used to distinguish the different subclasses of Igs. For example, IgG1, IgG2, IgG3, IgG4 differ in number of inter-heavy chain disulfide bonds [46]. There are disulfide bridges between heavy and light chains, in the hinge region and also there are intramolecular bonds in heavy and light chains.

The MS analyses of Abs will be performed before and after digestion with papain and Ides. After treatment with papain, the Fab and the Fc fragments can be obtained. These fragments can be analysed intact and after reduction with DDT, and then fragmented using MALDI-TOF/TOF. This provides information about sequence micro variants and the position of the disulfide bridges and PTMs. Regarding the analysis of Abs after treatment with Ides, the enzyme cleaves IgG in a sequence specific manner at neutral pH and the F(ab)'2 and the Fc fragments can be obtained. This second type of proteolysis complements data obtained with papain.

(iii) Analysis of antibody-drug conjugate (ADC) utilised in cancer therapy.

We intend to characterise the ADCs using MS. Namely, the purity of ADCs, their homogeneity/heterogeneity, their oligomeric state/aggregates, the Drug-to-antibody ratio (DAR) the relative quantification of the different species and the position of the bound drug can be assessed.

Firstly, the masses of the intact Ab (with no bound drug) and its related ADC species are measured by MS. The comparison of the masses of the "naked" Abs and ADCs allows the determination of the average DAR of ADC and of the drug load distribution (i.e., number of bound drugs). Secondly, we perform the "middle-down" analysis of the ADCs using mild enzymatic proteolysis and/or disulfide bond reduction (Figure 24). The generated fragments are further analysed by MS (e.g., by MALDI-TOF/TOF) to obtain structural information such as the position of the drugs.

(iv) Characterisation of the epitope-paratope interactions

The investigation of epitope-paratope bonds can be performed using MS in native and denaturing conditions. By native MS we can verify the existence of interactions between epitope-paratope and assess their stoichiometry and affinity. Using a mild proteolysis and MALDI-TOF-TOF analysis, we can sequence the fragments of the Ag to identify the amino acids of the epitope [58] (Figure 27).

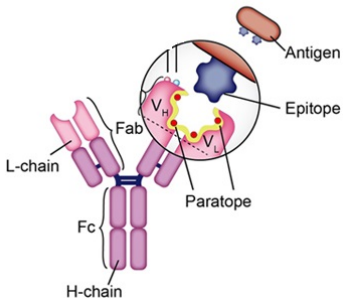


Figure 27. Schematic representation of the interactions between epitope and paratope. Figure adapted from Creative Diagnostics.

Native MS to study biomedically- and biotechnologically-relevant complexes

Studying a large RNA polymerase involved in the chloroplast biogenesis

Introduction

Plants convert sunlight into chemical energy through a chloroplast-dependent mechanism known as photosynthesis. This represents not only a key source of renewable bioenergy, but also a process that improves the Earth's atmosphere because it utilises carbon dioxide and releases dioxygen. Plants are also a source of food, of biofuels and of fibres used for clothing and construction. In the near future a deep knowledge of photosynthesis will greatly contribute to solve challenges as severe famine, environmental pollution and climate change. This requires a clear understanding of the chloroplast biology and its biogenesis.

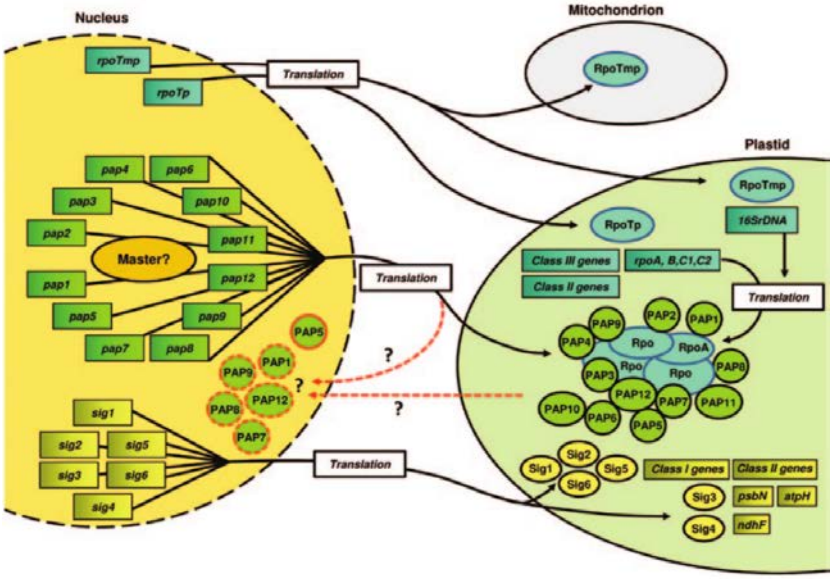


Figure 28. Transport and localization of the PAP proteins and of the composition of the PEP complex.

A chloroplast is an organelle characterised by two membranes and a high concentration of a photosynthetic pigment, the chlorophyll. This organelle contains its own DNA (aka, plastome) that encodes for 75-80 proteins. Other 3500-4000 chloroplastic proteins are synthesized from nuclear genes. Therefore, the chloroplast biogenesis requires on a coordinated expression of the nuclear genome and plastome, relying upon a bidirectional communication between the nucleus and the chloroplast.

Chloroplast biogenesis depends on the assembly of a RNA polymerase complex (i.e., plastid-encoded polymerase, PEP, 1.1 MDa) [59]. PEP is formed by 16 different proteins. PEP core components are 4 rpo subunits (rpo, RNA Polymerase Holoenzyme) and they are encoded in the chloroplasts. Moreover, 12 major PEP associated proteins (PAPs) are encoded in the nucleus (Figure 28). Mutations in pap genes yield albino plants incapable of photosynthesis, indicating that PEP is a key player in chloroplast biogenesis and function [60].

Since 2017, the MS laboratory is part of an ambitious project that aims to study the formation, regulation, and dynamics of the PEP and the PAPs. Scientists from the IBS, BIG (Institut de Biosciences et Biotechnologies de Grenoble) and IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire, Strasbourg) use complementary approaches to investigate the PEP and PAPs.

Using biochemistry and molecular biology we aim to recombinantly produce PEP and PAPs and also purify them from plants. By combining MS, EM, X-ray crystallography and NMR, we intend to solve the 3D structures of PAPs and PEP. Using plant genetics, colleagues study PEP and PAP mutants, their respective function in different cell types and organelles, and the signalling between the nucleus and the chloroplast. This project will bring basic knowledge that can potentially be applied in biotechnology and agriculture to better face environmental, social and economic challenges of the 21st Century.

The MS laboratory is in charge of two specific tasks of the project. (i) We utilise native MS to investigate of the PAP subcomplexes and the intact PEP. Specifically, we determine the mass of intact assemblies, their stoichiometry and the direct interactions between subunits. By mixing subunits in a stepwise manner, a hierarchy in the assembly pathway can be determined. Native MS is typically carried out using volatile buffers such as ammonium acetate and usually necessitates exchanging the sample buffer prior to analysis. In the case of PEP, noncovalent interactions may be weak and dissociated during purification or buffer exchange. To make particles more stable, the sample can be crosslinked using homobifunctional, water soluble crosslinkers [e.g., Bis(sulfosuccinimidyl)suberate] [40].

(ii) We aim to characterise the PTMs and their influence on the assembly of PEP. To date, the PTMs of PAPs have been marginally characterised because data have been mainly obtained during large scale proteomics studies [61]. We will focus our attention on PTMs of

PAPs, which are transiently present because they are temporarily expressed in young plants.

Using MS in denaturing conditions we combine the use of HPLC-ESI-mass spectrometer with the new instrument MALDI-TOF/TOF. Using these instruments we aim to assess: (i) the purity of the proteins, (ii) their homogeneity/heterogeneity, (iii) their PTMs and (iv) mutations. Protein fragments generated by limited proteolysis can be analysed to identify the flexible regions of biomolecules before crystallization studies. Since the HPLC apparatus is coupled with the mass spectrometer, complex mixtures and protein assemblies can be separated before being analysed. The HPLC separation and ESITOF analyses of proteins larger than 80 kDa or those which are post-translationally modified (e.g. glycosylated) may be difficult. In these cases, we will use the MALDI mass spectrometer, whereby the ionisation of samples is more efficient and no HPLC separation is required.

Overall, the PEP project will take great advantage from the MS data, which will complement the knowledge, obtained using structural biology and biochemistry.

Improving MS to investigate large macromolecular complexes

Investigating bacteriophages using native MS and NEMS

Introduction

The viruses infecting bacteria known as bacteriophages are the most abundant biological entity of Earth. Since they are natural predator of bacteria, they have an important role in the regulation, diversity and evolution of all prokaryote populations. Remarkably, phage therapy is an attractive alternative to treat antibiotic-resistant bacterial infections in humans, animals and plants [62]. Phages contain extremely efficient nanomachines used to infect their host very specifically, and to highjack the bacterial machinery for the production of new virions. Their study has led to the development of modern molecular biology, and is a source of inspiration and tools for biotechnologies. As a consequence, improving our understanding of phage structure and assembly will undoubtedly lead to key discoveries. 96% of all phages share a common structure composed of a **capsid** enclosing their genome, and a **tail** that recognise the host and delivers DNA into the host cytoplasm.

Bacteriophage T5 is a representative member of a very large family of phages infecting Gram-negative bacteria (*e.g. E. coli, Salmonella*). A 90 nm T5 capsid encloses a 121 kb double stranded DNA genome, and a 250 nm-long tail safely channels the viral DNA from the T5 capsid to the bacterial cytoplasm (Figure 29 A). Assembly pathways of the capsid and the tail are independent, and mutants allow the purification of functional tails or capsids at different stages of maturation.

The phage T5 capsid is an icosahedral protein shell that protects the viral genome. Capsid assembly follows an ordered sequence of tightly regulated events at the biochemical and genetic level (Figure 29), which have been established by combining genetic, biochemical and structural approaches [63]. However, the detailed knowledge of the exact stoichiometry and protein-protein interactions that control this assembly process is still lacking.

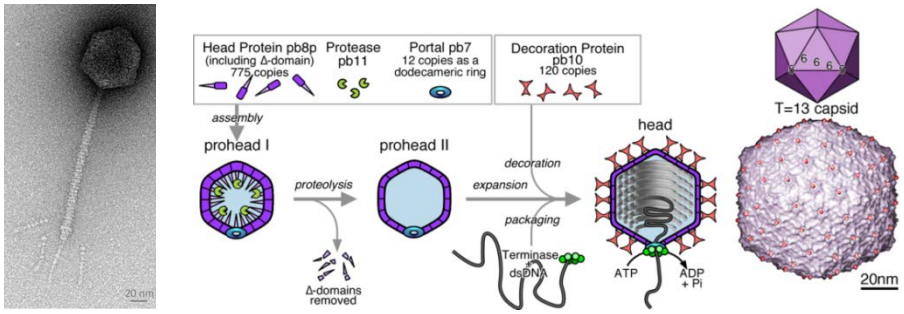


Figure 29. Information about the bacteriophage T5. A) Negative-stain electron microscopy image of phage T5 (diameter of the capsid: 90 nm). B) Scheme of phage T5 head assembly pathway. First, 775 copies of the head protein (i.e., pb8), 12 copies of the portal protein pb7 and unknown number of the head maturation protease pb11 assemble into a pro-capsid (**prohead I**). Then pb11 is activated, it cleaves itself and the N-termini of pb8 and of pb7, yielding an empty **prohead II**. Finally, the packaging of the 121 kb double-stranded (ds) DNA into prohead II is fuelled by an ATP-dependent molecular motor. The dsDNA packing triggers a two-fold increase of the capsid volume. The expanded mature capsid is then sealed by the head completion protein p144, and decorated with 120 copies of pb10 at the capsid surface.

The T5 tail is composed of eleven proteins (Figure 30 A, B). Among these subunits, the receptor-binding protein, pb5, is located at the tip of the distal end of the tail.

Interaction of pb5 with its receptor FhuA, an outer membrane protein of the *E. coli* host, may affect a protein, pb2 (121 kDa) [63c, 64] (Figure 30 C, D). This multifunction protein defines the length of the tail and is thought to perforate the cell wall [65]. The pb2 stoichiometry within the tail remains elusive, although some biochemical studies suggest 5 or 6 copies per phage particle. After its interaction with the receptor, pb2 might be subjected to partial proteolysis.

To date, phages have been studied using various biochemical and structural biology approaches, but few MS-based studies have been reported. Open questions still remain regarding the composition of the intact T5 virion, its capsid and tail, and novel insights could be provided by highly innovative MS technologies.

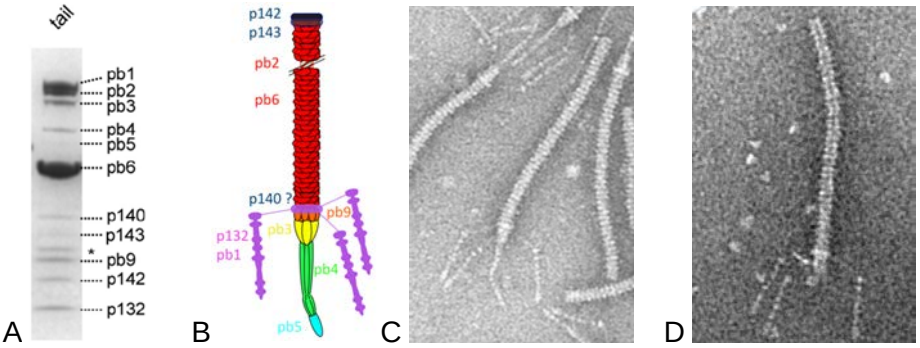


Figure 30. Tail of phage T5. A) Subunits forming the tail were analysed by SDS-gel electrophoresis. The eleven proteins have the following expected masses. Pb1, 148 kDa; pb2, 121.9 kDa; pb3, 107.3 kDa; pb4, 74.9 kDa; pb5, 67.8 kDa; pb6, 50.3 kDa; p140, 34.5 kDa; p143, 27.8 kDa; p9, 22.7 kDa; p142, 18.4 kDa; p132, 15 kDa [63c]. B) Schematic representation of T5 tail [63c]. C) and D) negative stain EM images of T5 tail before (C) and after (D) its interaction with FhuA. The tip of the tail seems to disappear, and it is unknown whether some of the T5 proteins may dissociate from the tail. Diameter of the tail tube: 10 nm.

Since 2018 I am part of a group of scientists who aim to improve our fundamental understanding of bacteriophage T5. We are supported by a CEA grant (i.e., DRF impulsion). We aim to study T5 because it is an ideal virus for investigating the structure and assembly process of phages. Indeed, *in vitro* intermediate particles of the assembly pathway (such as tails, empty or DNA-filled capsids) can be isolated and successfully purified.

Among the scientists supported by DRF impulsion, there are two experts of T5 and two mass spectrometrists. Over more than a decade, Pascal Boulanger and her team at I2BC (Institute of Integrative Cell Biology) has acquired unique expertise on the assembly of phage T5 capsid. At the IBS, Cécile Breyton and her team focus their studies on the T5 tail. Christophe Masselon and myself are both mass spectrometrists with complementary expertise. At BIG (Biosciences and Biotechnology Institute of Grenoble), the Masselon team has recently developed a new nano-electromechanical systems (NEMS)-based mass spectrometer in collaboration with LETI (Institute for electronics and information technologies, Grenoble) [21a, 21b] (Figure 31). Using this system, they demonstrated the detection of heterogeneous nanoparticles larger than 30 MDa. To complement intact mass determination, MS-based proteomics is performed at BIG using more conventional high-

performance mass spectrometers. At the IBS, our work on native MS enables the investigation of intact macromolecular assemblies up to 10 MDa.

Using a range of MS techniques, we intend to investigate the intact T5 virion (≈119 MDa), the DNA-full capsid (≈105-130 MDa) and the tail (≈9 MDa). Since these macromolecular complexes cover a large range of masses, we plan to utilise three different types of mass spectrometers that can handle biomolecules of different size:

1) NEMS-MS to investigate T5 proheads and mature capsids with masses up to 130 MDa
2) Native MS to analyse intact tails and its sub-complexes (up to 10 MDa)
3) MS-based proteomics to study T5 and host proteins (up to 150 kDa).

Using NEMS-MS and proteomics we aim to explore the mechanism of capsid assembly. Specifically, we will investigate the protease pb11 and its role in the maturation of the T5 prohead I (Figure 31). We will assess the stoichiometry of pb11 in prohead I. Since wild type proheads I are metastable intermediates of short lifetime [63d], we will mutate different residues of the pb11 active site or processing sites. Thus, we will produce variant proheads I unable to undergo proteolytic maturation.

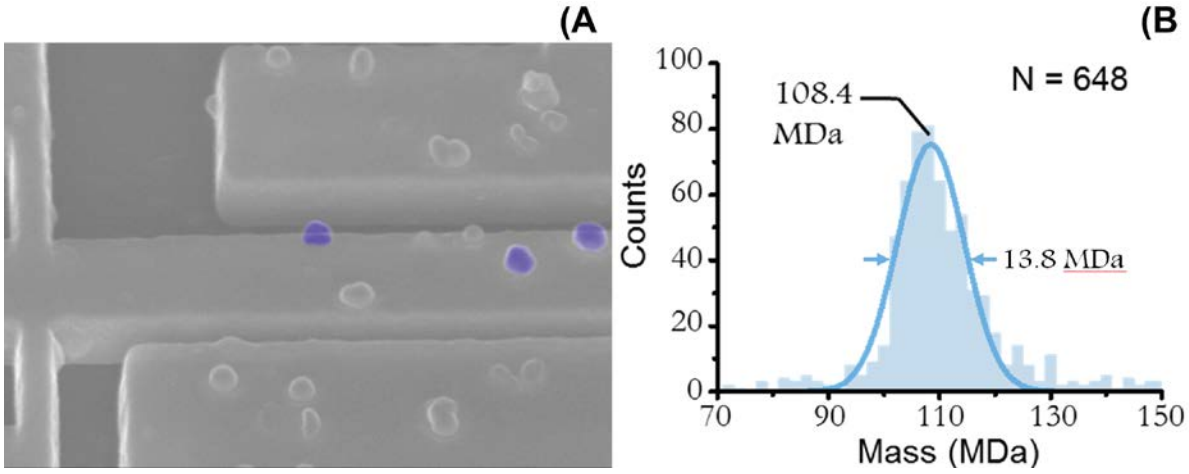


Figure 31. **Preliminary NEMS results.** (A) Scanning electron microscope image of bacteriophage T5 capsids (blue, false colour) observed on NEMS beam after particle landing and mass sensing. Polymer particles are grey. (B) NEMS-MS spectrum of intact DNA filled T5 capsids.

The IBS main activity will focus on native MS, but we will have the opportunity to learn more about NEMS and to discuss about the results of the capsid assembly obtained by proteomics. Using native MS, we will study the tail of T5 and its interaction with its *E. coli* receptor FhuA to determine the stoichiometry of each protein and to understand the re-

organization of the tail upon its interaction with the receptor. This represents a great challenge for native MS because the mass change might represent only ~10 % of the total mass of the tail. These data will be confirmed by MS-based proteomics, whereby we can assess the composition of the tail before and after receptor interaction.

We will face the challenge of characterizing MDa viral assemblies and their components with high accuracy and resolution. This will provide new stimulus for methodological development. Moreover, the MS will provide novel insights on T5, contributing to the exploitation of this phage to fight antibiotic-resistant infections.

References

[1] R. Aebersold, M. Mann, *Nature* **2016**, *537*, 347-355.
[2] M. Mann, P. Hojrup, P. Roepstorff, *Biological mass spectrometry* **1993**, *22*, 338-345.
[3] B. T. Chait, *Science* **2006**, *314*, 65-66.
[4] aJ. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **1989**, *246*, 64-71; bM. Wilm, *Mol Cell Proteomics* **2011**, *10*, M111 009407.
[5] G. I. Taylor, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **1964**, *280*, 383-397.
[6] M. Karas, F. Hillenkamp, *Analytical chemistry* **1988**, *60*, 2299-2301.
[7] aM. J. Stump, R. C. Fleming, W. Gong, A. J. Jaber, J. J. Jones, S. C.W., C. L. Wilkins, *Applied spectroscopy reviews* **2002**, *37*, 275-2002; bI. C. Lu, C. Lee, Y. T. Lee, C. K. Ni, *Annu Rev Anal Chem (Palo Alto Calif)* **2015**, *8*, 21-39.
[8] E. Boeri Erba, C. Petosa, *Protein science : a publication of the Protein Society* **2015**, *24*, 1176-1192.
[9] J. Snijder, R. J. Rose, D. Veesler, J. E. Johnson, A. J. Heck, *Angew Chem Int Ed Engl* **2013**, *52*, 4020-4023.
[10] aM. Faini, F. Stengel, R. Aebersold, *J Am Soc Mass Spectrom* **2016**, *27*, 966-974; bE. Boeri Erba, L. Signor, M. F. Oliva, F. Hans, C. Petosa, *Methods Mol Biol* **2018**, *1764*, 133-151.
[11] F. Sobott, H. Hernandez, M. G. McCammon, M. A. Tito, C. V. Robinson, *Analytical chemistry* **2002**, *74*, 1402-1407.
[12] H. Hernandez, A. Dziembowski, T. Taverner, B. Seraphin, C. V. Robinson, *EMBO Rep* **2006**, *7*, 605-610.
[13] E. D. Levy, E. Boeri Erba, C. V. Robinson, S. A. Teichmann, *Nature* **2008**, *453*, 1262-1265.
[14] A. Politis, F. Stengel, Z. Hall, H. Hernandez, A. Leitner, T. Walzthoeni, C. V. Robinson, R. Aebersold, *Nature methods* **2014**, *11*, 403-406.
[15] E. Boeri Erba, K. Barylyuk, Y. Yang, R. Zenobi, *Analytical chemistry* **2011**, *83*, 9251-9259.
[16] A. W. Yee, M. Moulin, N. Breteau, M. Haertlein, E. P. Mitchell, J. B. Cooper, E. Boeri Erba, V. T. Forsyth, *Angew Chem Int Ed Engl* **2016**, *55*, 9292-9296.
[17] H. Zhang, W. Cui, M. L. Gross, R. E. Blankenship, *FEBS Lett* **2013**, *587*, 1012-1020.
[18] P. V. Sauer, J. Timm, D. Liu, D. Sitbon, E. Boeri-Erba, C. Velours, N. Mucke, J. Langowski, F. Ochsenbein, G. Almouzni, D. Panne, *eLife* **2017**, *6*.
[19] L. Signor, E. Boeri Erba, *Journal of visualized experiments : JoVE* **2013**.
[20] aD. Suckau, A. Resemann, *Analytical chemistry* **2003**, *75*, 5817-5824; bT. K. Toby, L. Fornelli, N. L. Kelleher, *Annu Rev Anal Chem (Palo Alto Calif)* **2016**, *9*, 499-519.
[21] aM. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, M. L. Roukes, *Nature nanotechnology* **2012**, *7*, 602-608; bE. Sage, A. Brenac, T. Alava, R. Morel, C. Dupre, M. S. Hanay, M. L. Roukes, L. Duraffourg, C. Masselon, S. Hentz, *Nature communications* **2015**, *6*, 6482; cE. Sage, M. Sansa, S. Fostner, M. Defoort, M. Gely, A. K. Naik, R. Morel, L. Duraffourg, M. L. Roukes, T. Alava, G. Jourdan, E. Colinet, C. Masselon, A. Brenac, S. Hentz, *Nature communications* **2018**, *9*, 3283.
[22] C. S. Hughes, M. K. McConechy, D. R. Cochrane, T. Nazeran, A. N. Karnezis, D. G. Huntsman, G. B. Morin, *Scientific reports* **2016**, *6*, 34949.
[23] C. Guijas, J. R. Montenegro-Burke, B. Warth, M. E. Spilker, G. Siuzdak, *Nature biotechnology* **2018**, *36*, 316-320.
[24] A. B. Ward, A. Sali, I. A. Wilson, *Science* **2013**, *339*, 913-915.
[25] aF. Herzog, *Methods Mol Biol* **2014**, *1091*, 259-273; bA. Sinz, C. Arlt, D. Chorev, M. Sharon, *Protein science : a publication of the Protein Society* **2015**, *24*, 1193-1209.
[26] H. Zhang, A. Berezov, Q. Wang, G. Zhang, J. Drebin, R. Murali, M. I. Greene, *The Journal of clinical investigation* **2007**, *117*, 2051-2058.
[27] E. Boeri Erba, R. Matthiesen, J. Bunkenborg, W. X. Schulze, P. Di Stefano, S. Cabodi, G. Tarone, P. Defilippi, O. N. Jensen, *J Proteome Res* **2007**, *6*, 2768-2785.
[28] S. Cabodi, M. del Pilar Camacho-Leal, P. Di Stefano, P. Defilippi, *Nature reviews. Cancer* **2010**, *10*, 858-870.

[29] L. Damiano, S. E. Le Devedec, P. Di Stefano, D. Repetto, R. Lalai, H. Truong, J. L. Xiong, E. H. Danen, K. Yan, F. J. Verbeek, E. De Luca, F. Attanasio, R. Buccione, E. Turco, B. van de Water, P. Defilippi, *Oncogene* **2012**, *31*, 624-633.
[30] S. Kennedy, M. Clynes, P. Doolan, J. P. Mehta, S. Rani, J. Crown, L. O'Driscoll, *British journal of cancer* **2008**, *98*, 1641-1645.
[31] D. Repetto, S. Aramu, E. Boeri Erba, N. Sharma, S. Grasso, I. Russo, O. N. Jensen, S. Cabodi, E. Turco, P. Di Stefano, P. Defilippi, *PloS one* **2013**, *8*, e54931.
[32] D. Repetto, P. Camera, R. Melani, N. Morello, I. Russo, E. Calcagno, R. Tomasoni, F. Bianchi, G. Berto, M. Giustetto, N. Berardi, T. Pizzorusso, M. Matteoli, P. Di Stefano, M. Missler, E. Turco, F. Di Cunto, P. Defilippi, *The Journal of neuroscience : the official journal of the Society for Neuroscience* **2014**, *34*, 1542-1553.
[33] P. Di Stefano, L. Damiano, S. Cabodi, S. Aramu, L. Tordella, A. Praduroux, R. Piva, F. Cavallo, G. Forni, L. Silengo, G. Tarone, E. Turco, P. Defilippi, *EMBO J* **2007**, *26*, 2843-2855.
[34] A. Alfieri, O. Sorokina, A. Adrait, C. Angelini, I. Russo, A. Morellato, M. Matteoli, E. Menna, E. Boeri Erba, C. McLean, J. D. Armstrong, U. Ala, J. D. Buxbaum, A. Brusco, Y. Coute, S. De Rubeis, E. Turco, P. Defilippi, *Frontiers in molecular neuroscience* **2017**, *10*, 212.
[35] aC. V. Robinson, *Biochemical Society transactions* **2017**, *45*, 251-260; bA. C. Leney, A. J. Heck, *J Am Soc Mass Spectrom* **2017**, *28*, 5-13.
[36] E. Boeri Erba, B. T. Ruotolo, D. Barsky, C. V. Robinson, *Analytical chemistry* **2010**, *82*, 9702-9710.
[37] N. Wei, *Structure* **2009**, *17*, 1-2.
[38] M. Sharon, H. Mao, E. Boeri Erba, E. Stephens, N. Zheng, C. V. Robinson, *Structure* **2009**, *17*, 31-40.
[39] C. E. Bulawa, S. Connelly, M. Devit, L. Wang, C. Weigel, J. A. Fleming, J. Packman, E. T. Powers, R. L. Wiseman, T. R. Foss, I. A. Wilson, J. W. Kelly, R. Labaudiniere, *Proc Natl Acad Sci U S A* **2012**, *109*, 9629-9634.
[40] E. Boeri Erba, P. A. Klein, L. Signor, *Journal of mass spectrometry : JMS* **2015**, *50*, 1114-1119.
[41] C. Amero, P. Schanda, M. A. Dura, I. Ayala, D. Marion, B. Franzetti, B. Brutscher, J. Boisbouvier, *Journal of the American Chemical Society* **2009**, *131*, 3448-3449.
[42] R. A. Rader, *Nat Biotechnol* **2008**, *26*, 743-751.
[43] aJ. L. Messerschmidt, G. C. Prendergast, G. L. Messerschmidt, *The oncologist* **2016**, *21*, 233-243; bH. Kaplon, J. M. Reichert, *mAbs* **2018**, *10*, 183-203.
[44] R. Kunert, D. Reinhart, *Applied microbiology and biotechnology* **2016**, *100*, 3451-3461.
[45] A. Beck, L. Goetsch, C. Dumontet, N. Corvaia, *Nature reviews. Drug discovery* **2017**, *16*, 315-337.
[46] G. Vidarsson, G. Dekkers, T. Rispens, *Frontiers in immunology* **2014**, *5*, 520.
[47] A. Beck, E. Wagner-Rousset, D. Ayoub, A. Van Dorsselaer, S. Sanglier-Cianferani, *Analytical chemistry* **2013**, *85*, 715-736.
[48] Leow C. H., Cheng Q., Fischer K., M. J., *The Development of Single Domain Antibodies for Diagnostic and Therapeutic Applications*, Intech, **2018**.
[49] aF. Higel, A. Seidl, F. Sorgel, W. Friess, *European journal of pharmaceutics and biopharmaceutics : official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik e.V* **2016**, *100*, 94-100; bR. J. Sola, K. Griebenow, *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy* **2010**, *24*, 9-21; cA. Beck, O. Cochet, T. Wurch, *Expert opinion on drug discovery* **2010**, *5*, 95-111.
[50] L. Fornelli, D. Ayoub, K. Aizikov, X. Liu, E. Damoc, P. A. Pevzner, A. Makarov, A. Beck, Y. O. Tsybin, *Journal of proteomics* **2017**, *159*, 67-76.
[51] G. D. Pipes, P. Campbell, P. V. Bondarenko, B. A. Kerwin, M. J. Treuheit, H. S. Gadgil, *Journal of pharmaceutical sciences* **2010**, *99*, 4469-4476.
[52] aC. L. Gaughan, *Molecular diversity* **2016**, *20*, 255-270; bA. A. Shukla, L. S. Wolfe, S. S. Mostafa, C. Norman, *Bioengineering & translational medicine* **2017**, *2*, 58-69; cS. K. Gupta, P. Shukla, *Frontiers in pharmacology* **2017**, *8*, 419.
[53] G. Kohler, C. Milstein, *Nature* **1975**, *256*, 495-497.
[54] J. R. Birch, A. J. Racher, *Advanced drug delivery reviews* **2006**, *58*, 671-685.
[55] B. T. Chait, *Annual review of biochemistry* **2011**, *80*, 239-246.
[56] U. von Pawel-Rammingen, B. P. Johansson, L. Bjorck, *EMBO J* **2002**, *21*, 1607-1615.
[57] A. L. Tarentino, T. H. Plummer, Jr., *Methods in enzymology* **1994**, *230*, 44-57.
[58] M. I. Iurascu, O. Marroquin Belaunzanar, C. Cozma, U. Petrausch, C. Renner, M. Przybylski, *J Am Soc Mass Spectrom* **2016**, *27*, 1105-1112.
[59] J. Pfalz, T. Pfannschmidt, *Trends in plant science* **2013**, *18*, 186-194.
[60] S. Steiner, Y. Schroter, J. Pfalz, T. Pfannschmidt, *Plant physiology* **2011**, *157*, 1043-1055.
[61] Y. Schroter, S. Steiner, K. Matthai, T. Pfannschmidt, *Proteomics* **2010**, *10*, 2191-2204.
[62] H. Brussow, *Virology* **2012**, *434*, 138-142.
[63] aA. Huet, J. F. Conway, L. Letellier, P. Boulanger, *Journal of virology* **2010**, *84*, 9350-9358; bO. Preux, D. Durand, A. Huet, J. F. Conway, A. Bertin, C. Boulogne, J. Drouin-Wahbi, D. Trevarin, J. Perez, P. Vachette, P. Boulanger, *J Mol Biol* **2013**, *425*, 1999-2014; cY. Zivanovic, F. Confalonieri, L. Ponchon, R. Lurz, M. Chami, A. Flayhan, M. Renouard, A. Huet, P. Decottignies, A. R. Davidson, C. Breyton, P. Boulanger, *Journal of virology* **2014**, *88*, 1162-1174; dA. Huet, R. L. Duda, R. W. Hendrix, P. Boulanger, J. F. Conway, *J Mol Biol* **2016**, *428*, 165-181; eE. Vernhes, M. Renouard, B. Gilquin, P. Cuniasse, D. Durand, P. England, S. Hoos, A. Huet, J. F. Conway, A. Glukhov, V. Ksenzenko, E. Jacquet, N. Nhiri, S. Zinn-Justin, P. Boulanger, *Scientific reports* **2017**, *7*, 41662.

[64] C. A. Arnaud, G. Effantin, C. Vives, S. Engilberge, M. Bacia, P. Boulanger, E. Girard, G. Schoehn, C. Breyton, *Nature communications* **2017**, *8*, 1953.
[65] P. Boulanger, P. Jacquot, L. Plancon, M. Chami, A. Engel, C. Parquet, C. Herbeuval, L. Letellier, *The Journal of biological chemistry* **2008**, *283*, 13556-13564.