Leibniz-Gemeinschaft

Final report

Title of the project: Advanced mass spectrometric tools to elucidate structures involved in viral infection

Leibniz-Institute: Heinrich Pette Institute, Leibniz Institute for Experimental Virology
Reference number: SAW-2014-HPI-4
Project period: 01-04-2014 to 30-06-2017
Contact partner: Charlotte Uetrecht

Stand: 04. Dezember 2015

Table of contents

Executive Summary

The project aimed at promoting Dr. Charlotte Uetrecht for academic leadership positions. Since the project start in 2014, 8 publications related to the project have been published, two of which with first/last authorship of Dr. Uetrecht and her group members. Currently, she is supervising 8 PhD students, three will graduate in 2018, and two postdocs. She established fruitful collaborations nationally and internationally. Significant third party funding was acquired including a prestigious ERC Starting Grant, coordination of a BMBF project, EU, DFG and local funding. Therefore, Dr. Uetrecht is well positioned to become an academic leader in the near future.

The scientific projects on viral protein complexes progress well and first results on coronaviral proteins and virus-host interactions will be published next year. New complexes of coronaviral proteins were identified. Several side projects as well as newly started projects were already published during the last years. All projects will be continued on acquired third party funding or from the basic group budget.

The European XFEL started operation in 2017 and Dr. Uetrecht participated in one of the first experiments. The prototype mass spectrometer for sample delivery will be available and tested in 2018/19. First publications related to testing and experiments at FLASH are in preparation. The guest status and collaboration with the European XFEL have been prolonged until at least 2023.

Research question and aim

The main aim was the promotion of the applicant for academic leadership positions and integrating her into the local research community with a special focus on the newly founded CSSB (Centre for Structural Systems Biology). The scientific aims focussed on studying viral protein complexes with native mass spectrometry (MS) and develop a new sample delivery system for the European XFEL.

The function of proteins or the complexes they form is closely related to their structure. Therefore, structural information is crucial to understand and fight diseases, whether they are caused by gene defects or contagious agents. During the last decade, native (MS) has become a valuable tool in structural biology. Together with available proteomics techniques, MS has also proved its importance for virology. The research in structural virology focussed so far on capsid assembly. However, viruses also encode additional proteins, which are not main constituents of the capsid or viral particle. Investigations of complexes formed primarily by these non-structural proteins within the life cycle of human pathogenic viruses using MS approaches were foreseen. Next to the general study of stoichiometry, topology, shape and binding affinities of ligands and inhibitors, new techniques in native MS should be developed. One project focussed on studying the formation of replication/transcription complexes by non-structural proteins (nsp). The main targets were coronaviruses. A second project aimed at deciphering the modulation of host cell responses by viral proteins. Here, mainly Hepatitis C virus (HCV) and adenoviruses were foreseen as model systems. The projects involved protein-nucleic acid and protein-lipid interactions requiring development of suitable protocols for the native MS measurements.

The structural resolution, which can be achieved in native MS, is limited to several nm. Usually, high resolution structures of proteins are obtained with X-ray crystallography, nuclear magnetic resonance (NMR) or electron microscopy (EM). These techniques require crystallisation or are limited to either small (NMR) or large samples (EM). Especially, X-ray crystallography is limited when studying dynamic processes. X-ray Free Electron Lasers (XFELs) offer the possibility to study the structure of biological molecules of any size. Even dynamics of a molecule or reactions between molecules can be monitored. Initial results indicated that such experiments will become possible at extremely high laser intensities, like those the European XFEL facility is delivering. So far, biological samples including cells, viruses and protein complexes have been introduced to XFEL beams using liquid jets or aerodynamic lenses. Liquid jets come along with a high background, which impedes studies on single molecules, whereas aerodynamic lenses suffer from high sample consumption and disturbing solvent as well as debris layers around small samples. Moreover, both techniques are incapable of selecting low populated or transient states from a reaction mixture; and of removing contaminants remaining after purification. In other words, the currently available techniques are best suited for biomolecules in a static state, which can be highly purified in large quantities. However, native MS allows the selection of ionic species from mixtures and to free them from impurities. At the same time, the sample consumption is extremely low. Therefore, the third project aimed at integrating native MS for sample delivery at XFELs to increase the achievable resolution in MS and the accessibility of dynamic processes to XFEL science. The instrument development required raising additional third party funding.

The application to the Senate Competition Committee (SAW 2014) within the funding line “Promoting women for academic leadership position” supported the restructuring process of the HPI to better promote junior researchers and to swiftly integrate new trends in virology. The project focussed on basic research of assembly processes within the life cycle of human viruses using new, state-of-the-art technology, which can further our understanding and guide drug development to the benefit of human health. This fitted well within the research field of the Heinrich Pette Institute - Leibniz Institute for Experimental Virology and had potential to strengthen its unique position as the only extramural research institution focussing on basic virology in Germany.

Project development

As the group investigates viral structures and their assembly, it easily integrated into the HPI programme area “Molecular Mechanisms of Viral Pathogenesis”, especially in the new thematic priority “Structure and Dynamics of viral Morphogenesis”, which is lead together with the “Structural Cell Biology of Viruses” department. The latter being located both at the HPI at the CSSB (Centre for Structural and Systems Biology). How valuable native MS can be for structural biology was quickly recognized in the Hamburg area highlighted by many new collaborations (see below). It is highly complementary to EM and crystallography as snapshots of the solution situation even of dynamic processes can be taken. Despite the benefits of native MS, it was not available in the Hamburg area. The HPI is now at the forefront of this quickly growing field in Germany. The applicant has also a strong expertise in virology and established collaborations with partner institutes of the HPI within the DZIF (German Center for Infection Research) as well as groups on the DESY (German Electron Synchrotron) campus. Additional persons have been hired and the technology portfolio was extended to hydrogen/deuterium exchange MS via third party funding from the DFG. The group by now runs three native MS systems (one at European XFEL) and another one will be purchased in 2018 within an ERC funded project to develop new instrumentation. The applicant aspires to a W2 or W3 professor at a university within the near future.

Non-structural proteins

Coronaviruses bring a machinery for replication and transcription for a genome of >30 kb, which is the largest amongst RNA viruses. Two thirds of this genome is translated into the polyprotein pp1.10 and via frameshift into pp1.16, these polyproteins are conserved among coronaviruses. Two viral proteases cleave the polyproteins into 10 or 16 domains, the non-structural proteins, which carry out replication and transcription of the CoV genome.

The pp7.10 regions of the CoV polyprotein undergo processing by the viral Chymotrypsin like protease nsp5. The released non-structural proteins nsp7, nsp8 nsp9 and nsp10 are regulatory subunits and control the activity of the coronaviral main enzymes, which are encoded upstream of the frameshift signal. Polymerase stimulating activity of nsp7 and nsp8 is not well understood yet. Nsp9 possesses RNA binding activity but its role during infection remains obscure. Hetero-dimerization of Nsp10 stimulates a 3’-5’ exoribonuclease and methyltransferase, Nsp 14 and Nsp16 respectively, and its role as activating cofactor is thoroughly studied.

Processing of the polyprotein confers activity to the non-structural proteins eventually leading to the transcription and replication of the genome. Cleavage sites between the nsp domains are separated by a protease recognition sequence that has a glutamic acid at position P’. Further, positons P6’ to P4’ are essential and also specific, but can vary in amino acid sequence. Specific protease activity of different coronaviruses was studied, *in situ* and *in vitro* with substrates representing the amino acid sequence at the protease recognition sites. The specific activity of the 3C like protease is the highest at its auto-cleavage sites, and can decrease to 2% on other sites. Activity of the cleavage sites was found to be essential for viral progeny, with exception to nsp9-10 cleavage site. Nsp7 and nsp8 domain switch in the polyprotein, lead to non-permissive virus, indicating that not only activity but also the correct order in processing of these proteins needs to be followed. During coronavirus infection, nsp8 is detected relatively late compared to other nsp domains, suggesting a possible role for nsp8 processing in regulating viral replication.

Nsp7-10 polyprotein was found to possess a polymerase activity, making it more than merely a byproduct of processing. Furthermore, Nsp7+8 was found to possess non-canonical *de novo* initiation and primer extension polymerase activity, which functions were discussed as primase for the main RNA dependent RNA polymerase (RdRp) nsp12. Interaction of nsp7 and nsp8 was found to be hexa-tetrameric (8:8) in SARS Coronavirus and trimeric (2:1) in feline Coronavirus by X-ray crystallography, other techniques indicate a dynamic oligomeri-

zation pattern. However, Nsp7 and nsp8 confer activity to nsp12, but yet, activity *in vitro* polymerase assays cannot be modulated to reach the fidelity required for replication of a 30 kb genome.

Up to date, the proteins and conditions for assembling a fully active polymerase complex need to be identified. The dynamics of the nsp7 and nsp8 interactions are not well understood, because molecular structures of nsp12 and its binding to nsp7-nsp8 are unresolved. Understanding the dynamics and functionality of the viral polymerase complex would allow insight into the sophisticated mechanisms which needs to be triggered for transcription of coronaviral (–)-strand, (+)-strand or mRNA.

To find out more about the mechanisms of polyprotein processing by nsp5, we chose nsp7-10 as a model. By analyzing a full sequence and folded nsp7-10 with native MS, would not only reveal the cleavage site activity, based on the recognition peptide, it would also allow us to consider further structural factors. Cleavage sites could be covered by the respective N- or C-terminus of the adjacent nsp domain, it could also be influenced by other nsp domains in the tertiary or quaternary structure.

The goal was to analyze polyprotein processing by native MS. First the polyproteins and the protease needed to be analyzed in single measurements to learn about their ionization efficiency and behavior in the gas phase. Then *in vitro* cleavage products of nsp7-10 were analyzed, allowing us to confirm the nsp products and their possible interactions. By comparing complete cleavage and cleavage after specific time points, we wanted to analyze nsp7-10 by products and order of cleavage site specificity. The kinetic parameters at the cleavage sites were further to be investigated, by time-dependent cleavage, in the test tube by sampling after specific time points and in the ESI capillary by observing the cleavage continuously. Because native MS retains non-structural interactions, it was also planned to analyze the interactions of the nsp products. Particularly, nsp7 and nsp8 analysis, was supposed to reveal interaction mechanisms, pathway of oligomerization and the respective conditions.

Overall, our goal was to understand polyprotein cleavage, regulation of nsp release and nsp complex formation as an early building block of the polymerase complex. To further classify our results and generate broader impact, our goal was to compare the different coronavirus families. This would allow us to discuss the level of amino acid conservation and functional conservation in the coronaviruses.

The project was planned in a close collaboration with University of Hamburg and University of Lübeck, which ran a joint project on structural analysis of coronaviral polyproteins and non-structural proteins. The project was supported by native MS measurements which were part of the dissertation “Coronaviral Polyprotein Nsp7-10: Proteolytic Processing and Dynamic Interactions within the Transcriptase/Replicase Complex” by Sven Falke and gave important preliminary data for this SAW project proposal.

The sample preparation, production and purification of the proteins, was planned to be outsourced to University of Hamburg, the analysis by native MS was performed at HPI. The required knowledge in sample purification was successfully transferred from collaborators to the PhD student. After the mass spectrometer was installed and functional, first analysis of the polyproteins was carried out in November 2014. The sample preparation requires intensive labour, but the protein had to be analyzed 1-5 days post-purification. Experiences of the collaborators and also own experiments in this early phase of the project in freezing the polyproteins were unsuccessful, due to unfolding of the domains. However, nsp5 protease remains active and detectable by native MS upon freezing and can be stored in small aliquots.

At the end of 2015, we had analyzed SARS and 229e Coronavirus nsp7-10 of different of length and from different expression plasmids. We found conditions for setting up successful native MS measurements and were able to record mass spectra of complete and time dependent cleavage. Our data showed delayed cleavage of nsp7-8, possibly regulated by tertiary structure, coherent with *in vivo* data but different than peptide cleavage assays sug-

gested. We refocussed the project to obtain supporting data, and to see if we can deduce the kinetic parameters at the cleavage sites.

The Coronavirus project was not pursued actively at University of Lübeck from December 2014 onwards. Therefore, protein production was carried by the PhD students initially during visits in Lübeck. However, the low protein stability required a more steady sample availability. We transferred the protein purification to the HPI in 2016 and recently established bacterial expression and cloning of the polyproteins in our lab to increase project progress. In collaboration with the University of Gießen, we compared our results on SARS and 229E coronavirus to those of other human and feline coronaviruses. Non-cleaved polyprotein and complexes after cleavage into nsp7+8 were analyzed. This collaboration will be continued.

In a side project with EMBL Hamburg, the already existing collaboration was intensified about the analysis of the bacteriophage endolysin CTPL1, which targets the cell wall of *Clostridium tyrobutyricum,* a human pathogen. The phage and the endolysin have potential applications as alternative to antibiotics. We used native MS to analyze the complex formation in regard of protein activity of full length endolysin and a c-terminal CTPL1, an alternative gene product. We found that the expression of the alternative gene product can regulate the activity of the full length endolysin and thus, trigger bacterial cell wall lysis eventually phage release. The work was published in December 2015 as: “Crystal Structure of the CTP1L Endolysin Reveals How Its Activity Is Regulated by a Secondary Translation Product”.

In another project with EMBL Hamburg, we aimed at identifying protein complexes guiding the growth of human axons during development. The guidance cues are complex and involve many players, which hampers structural characterization. Main focus was on the interactions of the guidance cues Netrin and Draxin and their interaction with the cellular receptors DCC and UNC5. At EMBL the protein samples are produced in large scale eukaryotic cell cultures. At HPI we performed mass spectrometric analysis of the guidance proteins. Challenges were to analyze the protein within 12 -48 hours, due to their stability in solution and the resolution of the mass spectrum due to heavy glycosylation. A student started with his bachelor thesis and finally graduated with his thesis: “Investigating Protein Interaction of Netrin-1 to Illuminate Axon Guidance”. We could not find significant support for the binding of the sample proteins in native MS. Sample loss during buffer exchange into ESI compatible buffer and low resolution due to glycosylation were the challenges by analysis with native MS.

We wanted to overcome these challenges by an alternative technique, high mass MALDI-XL. High mass MALDI XL allowed us to prevent sample loss by measuring out of the physiological buffer and by enhancing signal by stabilizing the protein interaction by chemical cross linking. Via a collaboration at ETH Zürich, the PhD student got training to perform High Mass MALDI XLsupported by a short term scientific mission grant from COST Action “BM1403” with the title “High mass XL-MALDI of guidance cues”. The proteins were analyzed as planned, however, sample instability and degradation were still challenging. However, the methodology was successfully applied to coronaviral protein complexes and can in the future be applied in collaboration with the EMBL Hamburg, which runs a MALDI system suitable for these experiments.

As nucleic acid-protein interaction are also of relevance to nsps, we aimed at gathering experience in handling RNAs and DNAs and analysing these by means of native MS. In 2016 we started a collaboration with University of Hamburg, analyzing G-quadruplexes (GQ) with native MS. G-quadruplexes are nucleaic acid structures, which dynamically appear and disappear, dependent on physiological conditions. Drugability of GQs is of great interest e.g. protection for chromosomal telomeric ends or as latency regulators in viral infection. A bachelor student helped in establishing analysis of GQs in our lab. In her thesis she investigated the formation of RNA G-quad structures of Eppstein-Barr-Virus, a human pathogen and graduated as a bachelor with a thesis title: “Nucleic acids in native mass spectrometry”. The work also resulted in a joined publication. Native mass spectrometry is an emerging tool for analyzing the binding sites, modes and flavours of cations to G-quadruplexes.

Analysis of HCV non-structural proteins NS2, NS3 and NS4a is ongoing. Also here, we had to face problems with poor stability of the samples. Therefore, protein production was transferred from Lübeck to the HPI. Studies on protease processing are ongoing and will be completed in 2018. Furthermore, initial experiments to analyse HBV X protein by native MS were performed.

Modulation of host cell responses

A research line focused on the question how Adenovirus nuclear proteins like the transcription factor Daxx are regulated by the adenoviral capsid protein VI. The interaction between pVI and Daxx should be confirmed and the role of ubiquitin modifications for the interaction was planned to be probed. In later stages of the project the study should be extended to other interaction partners and the methods and results should be transferred to the investigation of other viruses.

Furthermore, we planned to study the interaction of Hepatitis C virus (HCV) core proteins and lipid droplets that were shown to be located at HCV assembly sites. Specific lipids interacting with HCV core proteins should be identified and a putative trimeric protein complex containing HCV core, the triglyceride-synthesizing host enzyme DGAT1 and the HCV protein NS5A should be characterized. Also, the interaction of individual HCV core proteins or in vitro assembled core proteins to very-low-density lipoproteins (VLDL) should be studied.

In the beginning of the project, pVI was introduced into H1299 cells by transfection and also infection with human Adenovirus C serotype 5 and proteins were detected in cytoplasmic and nuclear fractions by Western Blots to establish standard tools and quality controls for the upcoming experiments. In order to confirm the interaction of Daxx and pVI and to possibly identify unknown interaction partners, the proteomics experiments that should verify the interaction of Daxx and pVI and possibly identify unknown interaction partners were set up. For this purpose, the SILAC (stable isotope labeling with amino acids in cell culture) methodology and tag-purification followed by in-gel digests and analysis of the generated peptides by LC-MS/MS were projected and started with control experiments, that proved a good ratio of proteins containing isotopically labelled amino acids. In the next steps, complex formation was thought to be tested specifically for nucleus and cytoplasm samples, which should be checked by native MS in parallel to prove complex formation.

This project was a collaboration with another group in the institute and the responsible researcher moved in early 2015 from the Heinrich-Pette-Institute to the TU Munich/Helmholtz Zentrum München. As a result, the supervision in cell biological methods was not ensured. Furthermore, some organizational aspects remained unclear for several weeks and in this period the major focus of the work was shifted to other projects. Recently, another project was started in this context, in which we aim at obtaining structural information on the early adenoviral protein E1b-55K. The aforementioned transcription factor Daxx has an antiviral function and is counteracted by E1b-55K, which submits Daxx for proteasomal degradation. So far, all crystallization trials were unsuccessful, which is not surprising considering the large unstructured regions that are predicted based on the sequence, especially at the N-terminus. The protein was produced in an insect cell/baculovirus sytem and is now available for structural analysis by covalent labelling MS. These experiments are carried out by a student assistant, who is currently establishing the labelling and analysis protocols using standard proteins.

Consequently, from 2015 on the major focus was put on the HCV project. Here, we started with establishing protocols to recombinantly produce HCV core protein as this was a prerequisite for all the planned *in vitro* interaction assays. As shown in previous studies, *E. coli* HCV core is expressed to inclusion bodies and thus requires solubilization with a chaotropic agent and refolding of the protein in the course of purification. HCV core protein consists of two domains, the hydrophobic RNA binding and oligomerization domain D1 and the more hydrophobic domain D2 that is said to be responsible for association to the endoplasmatic reticulum and lipid droplets. A signal sequence at the C-terminus is cleaved by the signal peptide

peptidase during protein maturation. Since the aim of the project was to study the protein oligomerization in presence of lipids, it was highly desirable to work with the full length HCV core protein. Furthermore, affinity tags at the termini could potentially hinder lipid binding or protein oligomerization, and accordingly we decided to work with untagged protein constructs.

This intention proved to be not straight forward. The PhD student successfully teamed up with the colleague working on nsps to establish all required systems and protocols for the purification. The expression yields in two tested *E. coli* strains (Rosetta (DE3) and BL21(DE3) pLysS) were low and also protocols of the solubilization of inclusion bodies did not show the desired efficacy. Moreover, protein handling turned out to be difficult and we were struggling with truncated degradation products and sample losses during purification. Nevertheless, a purification protocol with strong cation exchange chromatography followed by size exclusion chromatography could be established and core protein with good purity but unfortunately also degradation products could be produced. Protein yields were so low that we decided that further optimization of expression and solubilization of inclusion bodies is absolutely necessary before refolding and binding assays can be conducted.

Scientists from Hamburg, who are experts in recombinant expression and purification of proteins for X-ray crystallography advised us to clone, express and purify shorter constructs of the HCV core protein and also consider the addition of His-tags for purification for comparison. Cloning of untagged and His-tagged constructs of D1 domain only and D1-D2 (without signal sequence) was accomplished in 2017 and the constructs are ready for expression.

Recently, the role of RNA for assembly of HCV core proteins to nucleocapsid-like particles was highlighted (Holmstrom et al. JMB 2017) what will be considered in future experiments. Similarly, nucleoproteins of Ebolavirus assemble around the genomic RNA. In 2016/2017 a Master thesis was supervised that dealt with expression, purification and initial protein characterization by native MS.

Other collaborative side projects, mainly with the core theme of studying proteins in the context of lipids or lipid-mimicking detergents by native MS were pursued in the last three years. Thus we could get broad experience in sample handling, measurement and analysis, which will benefit the HCV and Ebola project. Another main benefit of the collaborative projects was a successful integration into the science landscape in Hamburg.

For instance, binding of two copies of a small protein to a plasma membrane localized ion pump could be shown by native MS and conformational changes upon protein binding were investigated by ion-mobility MS in collaboration with Prof. Frank Sobott at the University of Antwerp (now Astbury Center at the University of Leeds).

Also, innovative tools for solubilization of membrane proteins, like nanodiscs or Saposin A lipoprotein discs (“picodiscs”) were used and possible applications in native MS were tested. For example, in cooperation with Karim Fahmy’s group from the Helmholtz-Centre Dresden-Rossendorf different mutants of a copper-transporting bacterial ATPase were measured solubilized in detergent and compared to samples in nanodiscs. These samples were also used for experiments at the free-electron laser FLASH in early summer 2016, were we first tested the applicability of the integration of native MS and free-electron lasers (FELs).

In the course of a collaboration project with Rob Meijer’s group from the EMBL Hamburg, we could gain expertise in studying protein-lipid interactions and lipid-dependent protein oligomerization by native MS as required for the Hepatitis C project. Many viruses hijack the host’s endocytosis machinery and enter new host cells on this route. From the first measurements, it was obvious that this project could be a fruitful collaboration and consequently in 2015 and 2016 this project was intensively pursued. It recently resulted in a high-impact publication (Garcia-Alai, Heidemann et al., accepted in Nature Communications).

In another collaboration project with Michael Kolbe and his group from the recently opened CSSB, the sorting platform of the Type 3 secretion systems (T3SS) from Salmonella was

studied by native MS. The T3SS is a macromolecular protein complex utilized by many pathogenic Gram-negative bacteria to initiate infection by injecting virulence factors from the bacterial cytoplasm directly into eukaryotic host cells. Acting as a molecular syringe the T3SS needle complex spans both bacterial membranes. On the cytoplasmic side of the T3SS there is another multi-protein complex that is proposed to function as a sorting platform, which selects and targets substrates for translocation through the needle. Thus far, the precise molecular organization of the T3SS sorting platform is unknown. Four different proteins were recombinantly produced in *E. coli* and complex stoichiometries were determined by native MS. Together with clues from gas-phase dissociation experiments and results from SAXS measurements, a detailed model of the building blocks of the T3SS sorting platform could be generated. A manuscript presenting these data is submitted and currently undergoing the publishing process.

Thanks to the 2015 mobility grant of the HPI, the lab of Prof. Frank Sobott at the University of Antwerp could be visited regularly in the last 3 years and samples could be analyzed by MS techniques that are not available in our lab in Hamburg yet. Frank Sobott is a well-known specialist for studying protein conformations by ion-mobility MS, membrane proteins in nanodisc systems and also obtaining more structural information on non-covalent protein complexes using advanced gas-phase dissociation and fragmentation techniques (surface-induced dissocation, electron-transfer dissociation). As a consequence, a productive collaboration was established and we have access to a large number of valuable methods that are not available in-house at the HPI.

MS in combination with XFELs for single particle imaging

The work in this project largely focussed on determining feasibility. Therefore, the ion flux from ESI was monitored and showed that the amount of ions is sufficient. A publication is currently being prepared. Furthermore, different platforms and trapping systems were discussed to reach a robust design. The BMBF granted funding for the Visavix project in collaboration with the University in Greifswald in 2016. A QToF platform was purchased and installed at European XFEL. A postdoc started on this project and is constantly modifying and testing the new instrument modifications. Initial test at FLASH indicated that FELs could also be used for fragmentation of protein complexes next to single particle imaging. These results culminated in a recently granted ERC project, which will also ensure the continuation of the coronaviral nsp project.

Results and discussion

The group is well integrated at HPI and is involved in multiple collaborations within the institute. Furthermore, fruitful collaborations were established in Hamburg and the surrounding. This is exemplified by successful joint grant applications to the Joachim Herz Stiftung (main applicant Michael Kolbe) and the Landesforschungsförderung with the Deligrah project (main applicant Julia Kehr, UHH). The group is further running a project within the DFG research unit Virocarb, which includes partners from all over Germany. Internationally, we are project partner within the EU Horizon 2020 FETPROACT Viruscan project. The group leader has been invited to prestigious conferences, such as the GRC Physical Virology, to present her results.

Non-structural proteins

In order to perform native MS, proteins are required to be exchanged into a physiological solution that prevents adduct formation but preserves the non-covalent interaction of tertiary and quaternary protein structures. These requirements are met by ammonium acetate solutions of 150 to 300 mM at pH 7.4 to pH 7.9 for coronaviral nsps. The proteins are exchanged by different system depending on the experiment. The absence of reducing agents leads to aggregation and loss of soluble protein within a few days. Our data suggest that this effect is due to disulfide bonding of amino acid side chains located in nsp8. In order to stabilize the proteins, all ESI buffers are spiked with 1 mM DTT before buffer exchange. All polyproteins analyzed were found to be strictly monomeric at common native MS concentrations. The nsp5 protease was found to be mostly dimeric. A lower dimer-ratio at concentrations below 10 µM and a concentration dependent dimerization like described in several reports could not be observed. However, dimerization is linked to activity.

Interactions of nsp7 and nsp8, co-factors of a functional polymerase, were analyzed by native MS measurements of pp7.8 processing. Both products of the processing were detected as a mixture of monomer and homodimer. While nsp7 preferred to interact as a homodimer, nsp8 preferred to be monomeric. Processing of polyprotein 7.9 Polyprotein processing is governed by the specific activity of protease at the cleavage site. To observe release of non-structural proteins from the polyprotein, we performed native MS of pp7.9 incubated with nsp5 after different incubation times. Initially, we observe increasing signals for nsp9 and pp7.8, whereas later signals for nsp7 and nsp8 increase while the signals for pp7.8 decrease. Time dependent appearance of nsp signal suggests a successive order of processing of SARS-pp7.9, first nsp8-nsp9 and second nsp7-nsp8. Furthermore, the oligomeric states of released nsp's were analyzed. Nsp9 was exclusively observed as a monomer, while nsp7 was found as mixture of monomer and homodimer. Furthermore, heterotetramers (2:2) of nsp7 and nsp8 were found. These results support that heterotetramers form only after pp7.8 undergoes processing. Furhermore, in collaboration with University of Gießen, we obtained results indicating a distinct behavior of feline coronavirus pp7.9 after cleavage. Similar experiments were carried out with pp7.10, the results confirm lower activity of Mpro at the nsp7-8 cleavage site. Both, nsp9 and nsp10 could be assigned at earlier timepoints than nsp7.

In order to determine kinetic parameters of nsp5 on the cleavage sites of the folded polyproteins, following protein constructs with a non-cleavable tag were designed: His(6x)-pp78, pp8-9- His(6x) and pp9.10- His(6x). Currently, a cleavage assay with a native MS readout is carried out. Due to protein specific ionization intensities, ESI-MS can only be used as a semi-quantitative technique that requires a carefully selected internal control to convert intensities into relative concentrations and thus determine cleavage ratios. In parallel to the semi-quantitative assay with native MS readout, an absolute assay based on Förster Resonance Energy Transfer (FRET) will be carried out. HiLyte488 and Q250 labelled peptides of 10 amino acid length representing the cleavage sites of the pp4-5, pp7-8, 8-9,9-10 will be cleaved by nsp5. The cleavage of the peptides will dislocate the fluorophore and the quencher following an increase in fluorescence, proportional to the substrate turnover. By

measuring the increase of fluorescence at 488 nm the K_m and K_D values of the protease at the respective cleavage site can be determined. Comparison of the assay with folded polyprotein and assay with the peptide will allow us to determine the influence of polyprotein tertiary structure on the specific activity of protease and its possible structure dependent regulation. The experiments are being finalized and results prepared for publication.

A follow-up project has recently been granted by the ERC (Spock's MS, 2018-2022).

Modulation of host cell responses

In collaboration with Rob Meijers from the EMBL Hamburg, we studied how the plasma membrane lipids PI(4,5)P_2 bind to clathrin adaptor proteins, that oligomerize upon PI(4,5)P_2 binding. These complexes are physically linking the plasma membrane to the clathrin coat in clathrin-mediated endocytosis, which is a major trafficking pathway to the cytoplasm. Using native MS we could show that clathrin adaptor proteins with ENTH- and ANTH-domains possess two PI(4,5)P_2 binding sites and dissociation constants were estimated based on MS binding studies. *H.sapiens* ENTH-domains form a hexameric hub in presence of PI(4,5)P_2. Further oligomerization with ANTH proteins was also found to be PI(4,5)P_2 dependent and mature complexes contain more phospholipids than expected from previous binding studies.

Native MS was used to investigate complex stoichiometries for co-assembled ANTH/ENTH/PI(4,5)P_2-complexes from three different species and gas-phase dissociation of non-covalent complexes gave insights into complex topologies. Formation of chimeric assemblies between the three different species highlights the universal principle of clathrin adaptor protein assembly. Native MS time-course experiments reveal the conversion of smaller complexes, to a larger, more stable form. Our characterization of macromolecular protein-lipid complexes from *S. cerevisiae, C. thermophilum* and *H. sapiens* in terms of structure, stoichiometry and assembly pathway by native MS is currently being published (Garcia-Alai, Heidemann et al., accepted in Nature Communications). We envision that this project will be continued by looking into effects of viral proteins on clathrin adaptor assembly.

Insights into the molecular organization of Salmonella T3SS sorting platform were obtained from the collaboration with Michael Kolbe's group. We studied protein complexes, containing two C-ring proteins, a regulator protein and an ATPase by native mass spectrometry (MS). In-depth native MS and collision-induced (CID) MS/MS analysis of different combinations of proteins and different protein domains revealed protein stoichiometries and a protein interaction network. For different subcomplexes the overall complex shapes were obtained by small angle X-ray scattering (SAXS). Results from native MS and SAXS were combined into structural models of the T3SS sorting platform using fragment-based structural modeling and molecular dynamics (MD) simulations.

Our results show how native MS can be used in conjunction with SAXS and computational modeling to get insights into the molecular organization of a multi-protein complex. The largest characterized protein complex consists of 6 proteins that adopt an extended L-shaped conformation in solution, which can be assumed to be the minimal building block for the assembly of the fully functional T3SS sorting platform.

Results of the above mentioned projects, especially on the characterization of protein-lipid interactions of clathrin adaptor molecules, are not only being published in scientific journals, results were also presented in several talks at national and international conferences by the PhD student (Annual Conference of the American Society of Mass Spectrometry 2017, German Society for Mass Spectrometry 2015 and 2017, Structural Proteomics Symposium 2016, 1st international CSSB symposium 2015, CSSB opening symposium 2017).

MS in combination with XFELs for single particle imaging

As outlined above, the project successfully started and additional funding was secured (BMBF Visavix, coordinator). The results from FLASH will be investigated for additional ap-

plications in the ERC project Spock's MS. Two publications about ion flux and FLASH studies are in preparation.

Overall, it can be said that projects were initiated as foreseen. However, some delays and reshaping of the project delayed the publication output. This was successfully compensated by establishing side projects, which allowed developing protocols applicable to viral proteins as well. The group is well recognized in Hamburg as well as nationally and internationally. Discussions with universities to secure a professorship are ongoing. The two PhD students hired for the project will graduate in 2018.

Economic exploitation of results

The projects focus on basic research. However, the results could lead to identification of new drug targets. Furthermore, the instrumental developments will result in new MS components, which will be exploited in collaboration with commercial partners in the long run.

Fasmatech, Greece

- Involved in component development for native MS and European XFEL

MS Vision, the Netherlands

- Involved in component development for native MS and European XFEL

Contribution of collaborators

Cooperation partners:

John Ziebuhr, University of Gießen

- Production of coronaviral proteins and complementary functional analysis

Lars Redecke, University of Lübeck

- Production of coronaviral proteins and complementary crystallographic analysis

Renato Zenobi, ETH Zürich, Switzerland

- Training in high mass XL-MALDI MS

Rob Meijers, EMBL Hamburg

- Protein production, crystallographic and biophysical analysis of proteins

Ulrich Hahn, University of Hamburg

- RNA G quadruplexes

Zoya Ignatova, University of Hamburg

- G quadruplexes

Michael Kolbe, CSSB Hamburg

- Production of T3SS protein complexes

Frank Sobott, University of Antwerp/University of Leeds

- Access to ion mobility MS and advances fragmentation techniques, training in membrane protein handling

Dmitri Svergun, EMBL Hamburg

- Complementary SAXS analysis

Henning Tidow, University of Hamburg

- Membrane proteins for MS analysis

Christian Löw, EMBL Hamburg

- Membrane proteins for MS analysis

Karim Fahmy, Helmholtz-Zentrum Dresden-Rossendorf

- Membrane proteins for MS analysis and FLASH experiments

Norbert Tautz, University of Lübeck

- Production of HCV proteins and complementary functional analysis

Lutz Schweikhardt, EMA University Greifwald

- Digital ion traps and component development

Hartmut Schlüter, UKE Hamburg

- Access to state of the art proteomics equipment

European XFEL GmbH

- Collaboration on sample environment development

Graduation of students

Bachelor thesis:

Jonathan Zöller (2016) Investigating Protein Interaction of Netrin-1 to Illuminate Axon Guidance

Kira Schamoni (2017) Nucleic acids in native mass spectrometry

Master thesis:

Ronja Pogan (2017) Establishing native mass spectrometry for studying virus particle glycan interaction

Janine Denise Kopicki (2017) Selbstassemblierung des Ebola-Virus Nukleoproteins analysiert mit nativer Massenspektrometrie

Doctoral thesis:

Boris Krichel (expected mid 2018)

Johannes Heidemann (expected mid 2018)

Publications

von der Heyde A, Lockhauserbäumer J, **Uetrecht C**, Elleuche S. (2015). A hydrolase-based reporter system to uncover the protein splicing performance of an archaeal intein. Appl Microbiol Biotechnol 99 7613-7624.

Mallagaray A, Lockhauserbäumer J, Hansman G, **Uetrecht C**, Peters T. (2015). Attachment of Norovirus to Histo Blood Group Antigens: A Cooperative Multistep Process. Angew Chem Int Ed Engl 54(41):12014-9.

Dunne M, Leicht S, **Krichel B**, Mertens HD, Thompson A, Krijgsveld J, Svergun DI, Gómez-Torres N, Garde S, **Uetrecht C**, Narbad A, Mayer MJ, Meijers R. (2016) Crystal Structure of the CTP1L Endolysin Reveals How Its Activity Is Regulated by a Secondary Translation Product. J Biol Chem 291(10):4882-93.

Szameit K, Berg K, Kruspe S, Valentini E, Magbanua E, Kwiatkowski M, Chauvot de Beauchêne I, **Krichel B**, **Schamoni K**, **Uetrecht C**, Svergun DI, Schlüter H, Zacharias M, Hahn U. (2016) Structure and target interaction of a G-quadruplex RNA-aptamer. RNA Biol 29:1-15.

Wegener H, Mallagaray Á, Schöne T, Peters T, Lockhauserbäumer J, Yan H, **Uetrecht C**, Hansman GS, Taube S. (2017) Human norovirus GII.4(MI001) P dimer binds fucosylated and sialylated carbohydrates. Glycobiology 27: 1027-1037.

Garcia-Alai M*, **Heidemann J***, Skruzny M, Gieras A, Mertens H, Svergun D, Kaksonen M, **Uetrecht C#**, Meijers R#. (2017) Epsin and Sla2 form assemblies through phospholipid interfaces. Nat Comm accepted. *equal contribution, #corresponding authors

Shin HC, Deterra D, Park J, Kim H, Nishikiori M, **Uetrecht C**, Ahlquist PG, Arbulu M, Blick RH. (2017) Ultra-high mass multimer analysis of protein-1a capping domains by a silicon nanomembrane detector. J Proteomics doi: 10.1016/j.jprot.2017.11.024.

Pogan R, Schneider C, Reimer R, Hansman G, **Uetrecht C**. (2017) Norovirus-like VP1 particles exhibit isolate dependent stability profiles. J Phys Condens Matter accepted.

- Persons funded by the project in bold.

Data accessibility

All resulting publications with main authorship will be published open access. Data related to publications is archived at the HPI and all raw data and lab books are additionally stored on a local server maintained by the IT department of the HPI. Published data is available upon request from the authors.

Press releases

24th April 2014 – new department and junior research group at HPI:

http://www.hpi-hamburg.de/de/aktuelles/presse/einzelansicht/archive/2014/article/international-renommierter-strukturbiologe-prof-dr-kay-gruenewald-wechselt-von-oxford-nach-hamburg//?tx_ttnews%5Bmonth%5D=04&cHash=89538b6abe5e253ae66a2cef4b107285

30th June 2016 – BMBF funding for European XFEL project:

http://www.hpi-hamburg.de/de/aktuelles/presse/einzelansicht/archive/2016/article/eine-million-euro-fuer-die-erforschung-von-protein-strukturen//?tx_ttnews%5Bmonth%5D=06&cHash=cefe99378db5b963bf5e33406daadc3a

07th Sep 2017 – ERC starting grant for C Uetrecht (follow up of coronaviral studies and new method as result of SAW and BMBF project:

http://www.hpi-hamburg.de/de/aktuelles/presse/einzelansicht/archive/2017/article/hpi-nachwuchsgruppenleiterin-dr-charlotte-uetrecht-erhaelt-erc-starting-grant//?tx_ttnews%5Bmonth%5D=09&cHash=f51b0b1f4e13e560854f1742d854411e