Project title: Approach to structures of difficult protein targets using new mass spectrometry-based methods

Project leader: Michał Dadlez, prof.

Project start date: 01.06.2015

Project end: 30.05.2019

Aim: the aim of the project is structural characterization of biologically important protein assemblies selected from a class of protein systems difficult in structural studies, for which classic methods are not applicable or fail. A set of new, mainly mass spectrometry-based (MS) methods enabling the structural analysis of such targets will be used. Specifically, project targets include: Abeta peptide oligomers, involved in Alzheimer's disease, and their complexes with lipid membranes and membrane proteins (RAGE receptor, prion protein – task 1), histone pre-mRNA cleavage complex (task 2), procentriole complex (task 3), and intermediate filament proteins with their interactors (task 4). These targets represent selected general cellular processes, often involved in major human pathologies. Their structural analysis, important for design of therapeutically beneficial modifiers was often hampered by lack of appropriate analytical tools.

Methodology: Methodologically the project builds on the observed recently breakthrough in development of a set of new alternative methods providing access to structural properties in protein assemblies not accessible with use of crystallography or NMR. The breakthrough comes from increase in efficiency of methods known for a long time (MS-monitored hydrogen-deuterium exchange (HDex), cross-linking, oxidative footprinting) or from new concepts (ion mobility separation, electron transfer dissociation for HDex studies). New methods still mature and we plan to work out new procedures in frame of the project to improve their efficiency.

Requirements:

PhD student (3 positions): interest in biophysics, MSc in biology, physics or chemistry, good English communication skills.

Postdoc (2 positions): Interest in protein structural studies, PhD in life sciences with educational background in biology, physics, or chemistry, experience in protein expression and purification methods, good English communication skills.

Below the subprojects are presented to more detail along with basic literature. **The ability of the candidate to discuss one of the subprojects during the selection panel will be an important advantage so the lecture of the project description and the indicated literature is strongly advised.**

A. Aim of the project: The aim of the project are structural studies of selected biologically important protein assemblies with use of new mass spectrometry-based (MS) methods. Project focuses on a class of protein systems difficult in structural characterisation, for which classic methods are not applicable or fail. Specific targets include: Abeta peptide oligomers, involved in Alzheimer's disease, and their complexes with lipid membranes and membrane proteins (RAGE receptor, prion protein – *task 1A-C*), histone pre-mRNA cleavage complex (*task 2*), and procentriole complex (*task 3*), intermediate filament proteins with their interactors (*task 4*).

Project builds on recently observed switch in protein studies paradigm, in which protein function is no longer necessarily linked to stable, well defined 3D structures. This change of focus brings to light a large, but structurally unexplored group of intrinsically disordered proteins, which only now can be approached in terms of structural studies due to new developments in analytical techniques, mainly MS-based. This new methods are still *in statu nascendi*, so we also plan to work out new procedures which will make them more effective both at the step of data collection and analysis (*task 5A-C*), finally allowing to integrate the results from different analytical approaches into one procedure allowing to correlate structural molecular models with the set of obtained experimental constraints.

B. Background Classic methods of atomic-level protein structure analysis – crystallography and nuclear magnetic resonance – are the milestones of modern biology. During last 50 years these methods provided detailed insight and ultimate understanding of the basis of biological activity in case of plethora of molecular assemblies (multi-protein, protein-DNA/RNA, protein-lipid complexes). The application of these methods relies on several conditions imposed on the protein of interest which are not always easy to fulfill. For crystallography the protein must form high quality protein crystals. NMR requires relatively small, isotopically labeled objects, soluble at mg/ml concentration. It is estimated that only a small fraction of proteins can fulfill these conditions and a **larger part of protein structurome will never yield to these analytical tools**, constituting a vast group of difficult protein targets (DTs). **Classic methods are insufficient** to get access to these proteins and **often fail to reveal the structure of DTs**. Also, some proteins require oligomerisation for activity, like RAGE receptor ([1],[2] *see task 1 of the present project*); intermediate filament proteins - *task 4*) or gain pathologic activity upon oligomerization, like peptide Abeta [3] – *task 1,* or perfringolysin [4]**.** Oligomerization pathways include coexistence of numerous oligomeric forms of different order which are difficult/impossible to separate, so these protein enlarge lists of DTs even more.

A major group of DTs are proteins of high structural dynamics, so called **"intrinsically disordered" proteins** – IDP's. This is an abundant group of proteins, sometimes of primary biological importance, characterized by lack of unique three-dimensional fold in a substantial fraction of domains, which nevertheless retain functionality. It has **only recently been more widely appreciated how large and how important it is** [5]. It is estimated that **25–30% (!!) of eukaryotic proteins are mostly disordered**. **More than half of eukaryotic proteins (!!)** [6] **and more than 70% (!!) of signaling proteins** [7] **have long regions of disorder – IDRs.** Their abundance challenged the widespread structure-centric viewpoint, enforcing a **redefinition of an existing protein structure-function paradigm** [8] that dates back to "lock-and-key" hypothesis formulated in 1894 by Emil Fischer [9], which claims in simplified terms that a protein needs to fold to attain biological function. The paradigm, reinforced by amazing successes in solving thousands of crystallographic protein structures, imposed a static view of protein functionality. The paradigm still remains true as long as catalytic activity of proteins is involved, but **widespread presence of IDP's has shaken its validity** for other types of biological activities, which are many.

These results by no means stand in contradiction to abundance of crystallographic structures. In these structures IDRs are often seen as regions with missing electron densities or regions with high B-factor. Intriguingly, only ~7% of proteins in PDB are completely devoid of disorder and **only ~25% of PDB entries have >95% of their lengths observed in the corresponding PDB structures** [10]. Even in well-structured proteins the dynamics, assessed by measuring the protection of amide protons against exchange with solvent, can differ by at least four orders of magnitude [11]. Interestingly, the

fraction of proteins with IDRs is even 10-fold smaller in archaea or eubacteria, with evolution towards a higher ordered kingdom paralleled by a jump in number of IDPs. This suggests their advantageous role and **involvement of IDPs in biological function** [8].

Indeed, numerous subsequent studies proved the **functional role of IDRs** [12] **complementary to the catalytic and transport activities of ordered proteins**. Moreover, it was argued that many disorder-related functions (e.g. signaling, regulation, and recognition) are incompatible with stable well-defined structures and that the disordered status is functionally advantageous [13]. Many IDRs undergo a disorder-to-order transition upon binding to their partners. This **sub-class of order-disorder IDRs can be exemplified by coiled-coil domains** (for instance building blocks of intermediate filaments - *task 4*) which are a common and abundant structural motif in proteins. Coiled-coils are often unfolded when deprived of their intermolecular partners, while in complexes they form stable structures. Disorder-order transition facilitates binding reversibility and thus increases the effectiveness of signaling. This is achieved due to their **ability to fulfill the requirement of weak binding accompanied by high specificity**, as the free energy required for disorder to order transition decreases contact free energy, while large contact areas enforce high specificity. IDPs also are overrepresented among other major disease-related proteins, 79% of cancer-associated and 66% of cell-signaling proteins contain IDRs [7]. In particular hub/scaffold proteins [14] and transcription factors [15] commonly use disordered regions to multiplex interactions. **In result IDPs were found to orchestrate major cellular processes**, as it has been exemplified by the analysis of Wnt signaling pathway [16], regulation and execution of different modes of **programmed death of the cell** [17] or **organization of cellular division process** (Richter M. et al., in preparation).

It is thus not astonishing that **of abundant and biologically important IDPs many are involved in major pathologies**. The most classic are amyloidoses or conformational diseases, where structural disorder leads to pathological loss or gain of function through aberrant oligomerization and aggregation into fibrils. In this group tauopathies, synucleinopathies and prion-like phenomena can be listed – with the best known **Alzheimer's disease – AD (*task 1*). In AD the identity of the neurotoxic molecule is known for years (Abeta peptide) but the structural form of the peptide which is the most neurotoxic still remains unknown,** in spite of enormous scale of connected research (PubMed returns 16084 papers when searched for Abeta). The reason for this is that a 40-42 amino acid, strongly aggregating, Abeta peptide is a DT. Recent studies have shown that main neurotoxic form of this peptide has to be searched among an enormous plethora of different-order oligomers, which may coexist and equilibrate fast in solution making them inaccessible to classic tools of structure analysis [3],[18]. In addition, its activity is revealed in the interaction with biological membranes of different composition and membrane-embedded proteins, complicating analysis even further (*task 1*).

Changes in perspective on structure-function relationship also enforce **adjustment of drug design strategies which must take into account the dynamic nature of the protein target** [19]. Successful designs, taking into account the disordered character of the target protein, include inhibitors of c-Myc oncoprotein [20], or a corrector against the mal-folded form of ΔF508CFTR protein in cystic fibrosis [21] (*task 4*). **In conclusion – a vast class of biologically important proteins may remain completely unexplored on structural level if new analytical tools are not worked out and popularized.** This caused a call for new approaches that would enable to characterize the "unfoldome" in structural terms [22]. Unfortunately, it had to be repeated after 11 years [12] as the progress is slow. There is thus an **even more pressing need for alternative and effective new ways to approach the structural properties of difficult protein targets** or protein assemblies.

Most promising for a real breakthrough in this respect are recent developments of various MS-based approaches [23]. In the last 2-3 years several procedures have been worked out which either make methods known for a long time more efficient (MS-monitored hydrogen-deuterium exchange (HDex), cross-linking, oxidative footprinting) or build on new concepts (ion mobility separation, electron transfer dissociation for HDex studies) – *task 5*. These improvements streamline data collection and analysis and give hope for their routine use not only in specialized laboratories, with the timespan between the plan and the final result counted in days and not months. **In the opinion of the applicant, when integrated into one toolbox, enabling the structural characterization of difficult protein assemblies, these methods may in near future routinely complement classic methods for unfoldome studies**, allowing to solve many problems which cannot be treated by classic

methods. **Applicant's group takes active part in this process** since identification of new proteins or new protein-protein interactions by proteomic approaches rises interest in structural follow-up. This sequence of events is a frequent motif in **our collaborative efforts** and can be exemplified by numerous cases in which **identification of new proteins prompted their structural studies** (Table 1). Many of our objects belong to class of DTs and MS-based approaches serve as
of choice. In other instance new methods **complement crystallographic or NMR-based studies** [24] for instance by applying a minimization strategy in which **HDex identifies well-structured regions, allowing for rational design of minimized constructs** suitable for classic tools. [25][26][27][28][29]

Table 1
- D. Glover, Cambridge Univ., UK, *Drosophila* kinetochore & centriole complex (Richter, M., in preparation), [56]
- A. Edelman, Necker Inst., France & H. Herrmann, DKFZ, Heidelberg, Germany, K8-NBD1-CFTR complex (in preparation)
- S. Harrison, Harvard Med. School, USA, Yeast kinetochore COMA complex (Richter, M. in preparation)
- Z. Dominski, B. Marzluff, North Carolina Univ., USA, Histone pre-mRNA cleavage complex [25],[26]
- N. Sonnenberg, McGill University, Montreal, Canada, miRNA-mediated gene silencing complex proteins CNOT1, GW18 [27],[28]
- G. Dobrowolska, IBB - SnRK2 plant-specific stress response Ser/Thr kinases complex [29]
- M. Bochtler, IIMCB, Warsaw, Cross-talk of DNA-binding domains in DpnI endonuclease [24]

C. Project outline. The project builds on the new developments in MS-based approaches to study protein structure. It involves the studies of selected protein assemblies and further improvements of the methodologies on experimental and bioinformatic levels. Applicant's group has gained substantial experience in the application of MS-based methods for protein structural studies participating actively in the field for many years. An important aspect of the project is its **interdisciplinary** character, linking **new advances in physics** (measurements of collisional cross-section of a molecule during MS experiment, application of new types of fragmentation like ETD or new way to generate free radicals for oxidative footprinting) and **informatics** (new data analysis tools) to provide new data on **protein function for biology and medicine**. Deep **involvement of above-mentioned collaborating teams** is planned (see letters of intent at http://mslab-ibb.pl/en/collaborators), allowing for efficient bi-lateral networking of our laboratory with European and American laboratories. Since the present project deals with structural aspects of protein complexes under study the experimental part of work will be carried out nearly exclusively in the applicant's lab. The general aims of the project are to: **1.** Answer new questions on selected objects currently under study in the laboratory; **2.** Expand these studies to incorporate new partners of protein classes under study, not known at the time of writing. We plan to allocate the resources in a flexible way focusing on the most biologically important protein assemblies for which classic methods failed in spite of considerable effort invested by collaborating laboratories; **3.** Develop new data collection and analysis methods and tools with the ultimate goal for the integrated software for automated selection of molecular models, best explaining structural constraints derived from MS-based analyses.

Task 1. Peptide Aβ oligomers and their interactions with lipid membranes and selected membrane proteins. A postdoc ½ time and one PhD student assigned.

Oligomers of **Abeta peptide are widely believed to be the main synaptotoxic and neurotoxic agent in Alzheimer's disease (AD).** Causative link between the non-monomeric forms of Aβ peptide, derived by proteolysis of APP protein, is supported by several lines of evidence [30], leading to the amyloid cascade hypothesis. The oligomer hypothesis is central for AD studies at present and the subject of frequent reviews [31],[32]. Basic hurdle in the search for effective drug against AD is that the identity of the **major neurotoxic form of Abeta has not been identified yet** [32]. Studies using different methods discovered a stunning variety of oligomeric forms of Abeta which display different levels of toxicity, also in the absence of fibrils [31]. Moreover, possible interconversion of different forms precludes typical structure-function studies, as the species cannot be preparatively separated and studied in isolation. Some of the various forms co-existing in solution may further assemble into mature fibrils, while other intermediates may evolve differently, leading to alternative forms the so called off-pathway species. Recent work indicates that the most neurotoxic oligomers indeed can originate from alternate pathways [33].

Since mature fibrils do not seem to represent the major neurotoxic form, **off-pathway oligomers gain special attention, as potential drug targets**. This provoked a call for systematic definition,

classification and characterization of all species that can be formed by oligomeric Abeta [31], so that the design of the beneficial modifiers of their activity could be more rational. The studies of such a complex network of interactions of the oligomerising molecule requires tools which allow to separate signals of different oligomeric forms coexisting in solution. Recently, a limited spectrum of methods available has been expanded by ion mobility separation (IM) coupled with mass spectrometry (IM-MS). It allows to resolve signals from co-existing species not only according to their molecular mass but also according to their collisional cross section (CCS), and to measure the CCS values characteristic for each resolved form. In such a way basic structural characterization of each of coexisting species can be obtained in spite of the complexity of the starting mixture. IMS-MS, thought relatively new, became an established analytical tool especially for peptide/protein aggregation studies [34]. In previous work, **we have provided the first experimental evidence for the co-existence of two structural conformers of Abeta oligomers of the same order** [18]. For a given n-meric oligomer IMS drift time distribution is bimodal, indicating the presence of the species of different CCS and thus families of more compact and more extended alternate forms. These forms may represent two pathways of the oligomer evolution, one leading towards fibrils and the second towards off-pathway oligomers, potential candidates for the most neurotoxic species.

We have also identified several factors, including metal ion binding, either stabilizing on-pathway species and fibril formation, or destabilizing this pathway and promoting off-pathway species [3], [35]. Though our work is a gas-phase study it provided experimental evidence that oligomeric structures can be retained during/after electrospray ionization and that "structural memory" can last longer than the time of IMS experiment [3]. Structural conversion between compact and extended oligomers may be crucial for the development of the disease directing Abeta monomers either to relatively benign fibrils or to more aggressive off-pathway species. **Understanding the factors that influence the relative population of structural forms of oligomers seems thus to be crucial for understanding the disease and modifying its progression in a therapeutically beneficial way. To fully understand the code directing the stability of different forms a more systematic studies of mutation series is planned in the frame of the present project** (*task 1A*). IM-MS experiments will be carried out for oligomers assembled in different starting conditions, metal ion (copper, zinc) presence, and for different point mutants scanned along known structural element like central hydrophobic core (CRH), metal binding region, or Met35. Also, pathologic mutations like H6R, D7N, A21G, E22G,Q,K, D23N will be studied.

Several other factors add to the complexity of the problem since oligomer's toxicity is mediated by interaction with metal ions [36], lipid membranes [37] and numerous membrane proteins (α7 nicotinic acetylcholine-, NMDA, AMPA, insulin-, RAGE, EphB2 receptors, prion protein etc.). Since so many proteins were undoubtedly shown to be affected**, it has been speculated that the toxic effect might be indirect, mediated by membrane-binding event and/or changes in membrane properties caused by binding of oligomers.** In addition Abeta has been found to form ionic channels in model biological membranes [38] which might directly lead to the observed calcium dyshomeostasis. This underscores the need for the studies of the membrane binding properties of oligomers [39]. In a pilot study, using HDex we have mapped the regions protected in Abeta upon binding to liposomes (Fig. 1). In the course of the project (*task 1A,B*) systematic correlation of peptide sequence (point mutants and peptide libraries) and the ability to structurize in contact with membranes of different composition, with cholesterol as a major component, in the presence or absence of metal ions (copper, zinc) will be carried out. Peptide libraries at different positions will be tested for their membrane-binding properties by their ability to co-sediment with liposomes of different composition and the structures of resulting ion channels/membrane bound assemblies will be analyzed by new methods.

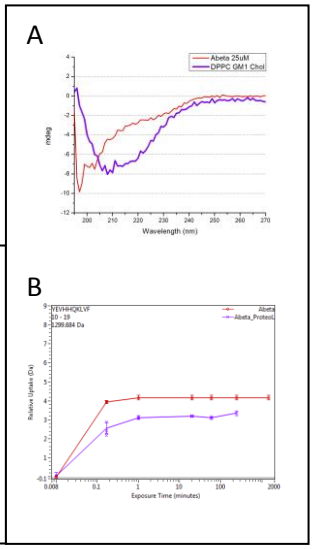

Fig. 1 (A) Circular dichroism of Abeta and (B) deuterium uptake (in Daltons) at different incubation times (horizontal axis) for Abeta fragment 10-19 either in solution (red) or in liposomes (blue). HDex indicates a peptide region with decreased exchange caused by secondary structure revealed in A.

Abeta also interacts with many membrane-embedded proteins. Since its impact on these proteins may be indirect, caused by peptide-membrane interactions, the structural studies should be carried out in the presence of model lipid membranes. MS-based methodologies are well suited for this purpose; the analytical step can be carried out also for protein-lipid membrane assemblies, as we have shown in a study of the conformational changes of a bacterial toxin, upon transition the liposomes [4]. One of these proteins is receptor RAGE (recently reviewed in [40]), responsible for the transport of Abeta across blood brain barrier and mediating the direct interaction of Abeta with neurons. **RAGE receptor plays crucial role in maintaining balance between inflammation and tissue repair and is involved in variety of pathologies, including neurodegenerative diseases (including AD** [41]), diabetes, cancer, stroke, septic shock, atherosclerosis, chronic renal inflammation, etc. Its pathologic activation leads to chronic inflammation, a common denominator of these diseases and make it a potential field of therapeutic intervention in numerous cases.

RAGE has been classified as a pattern recognition receptor by analogy with other receptor systems that recognize common structural patterns shared by different classes of ligands. Many different ligands bind to RAGE and the binding event is transmitted in a ligand-dependent manner into the cells [42], in spite of the fact that monomeric RAGE is an IDP with binding domain highly dynamic [1] and structurally uncoupled from the intracellular short peptide. This makes simple allosteric signal transduction unlikely and leads to the requirement for receptor oligomerization. Ligand introduction is presumed to lead to a shift in the oligomeric form distribution to higher order oligomerization states [43]. Though finally the structure of extracellular part of RAGE (exRAGE) has been solved [44] its **oligomeric structure remains highly speculative as many different groups show different and mutually exclusive modes of oligomerization, involving each of RAGE domains** [44],[45]. Majority of model structures of oligomers come from in solution studies of RAGE fragments, or crystal structure packing modes, so it is possible that RAGE oligomerises differently when embedded in the membrane. Using a di-tyrosine cross-link, we have created a C-terminally covalently linked exRAGE dimer mimicking anchoring in the membrane. Using HDex, we have identified the oligomerisation interface to C1-C2 linker region [3] unlike in the models of other groups. **Final proof, however, for the oligomerization mode must come from the structural studies of the membrane form of RAGE.** In a pilot set of experiments we have established a protocol to obtain membrane-bound form of recombinant full length flRAGE receptor, which retains binding of Abeta (Fig. 2). The protocol has to be scaled up (*task 1C*) to obtain quantity sufficient for MS-based structural studies. When done, it will open a wide field of activity. We believe that **the structural studies of this construct will bring definite answers to many important questions**: **what are the differences in structure between known solution structure of RAGE domains and their membrane bound form? what are the oligomerization interfaces? what is the dependence from lipid membrane composition? what is the role of metal ions (calcium and zinc)? how this structure changes upon binding different ligands?** (we have working protocols for overexpression and purification of Abeta peptide and S100 family of proteins, RAGE interactors [46]), **does the structure of short cytoplasmic tails change when ligands bind? what is the status of cysteines?** etc. There are six cysteine residues in RAGE sequence, and these were believed to stabilize each of the three RAGE domains by a single disulfide each [47]. But recent data [45] indicate that such situation is an exception, pertaining only for lung tissue, whereas in the other tissues disulfide in domain C2 rearranges intermolecularly to stabilize RAGE dimer by C2-C2 disulfide, lack of this disulfide directs RAGE to degradation. MS is especially well suited to map cross-links, including cysteines. Using an in-house LC-MS data analysis software [48] we have worked out an efficient procedure which allows to identify and relatively quantitate cysteine status in the protein sample of interest (M. Walczak, M.Sc. Thesis, Warsaw. Univ. 2014). **It will be thus of**

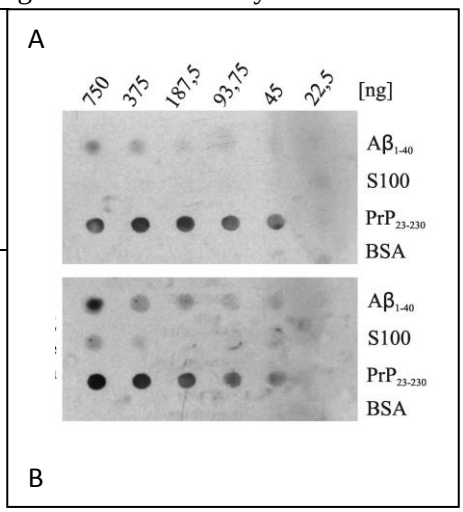

Fig. 2 flRAGE (B), cosedimenting with liposomes, retains binding to its ligands (Abeta, S100B protein, prion) in a dot blot assay, similarly as observed before for its extracellular part (A).

interest to monitor cysteine status in our membrane bound RAGE in different conditions, and binding of different ligands.

It was recently found that the cellular **prion protein, PrPc binds Aβ oligomers with high affinity acting as a putative receptor that mediates at least some of their neurotoxic effects** [49]. The question of prion involvement in AD has become an important and controversial issue [50]. Cellular prion protein, one of the most intriguing DTs known, is abundantly expressed in the nervous system. The binding site of Aβ oligomers was mapped to a cluster of basic residues at N terminus of PrP and the region within the unstructured central domain of PrPc (amino acids 95-134), but the structure of prion-Abeta complex is not known yet.

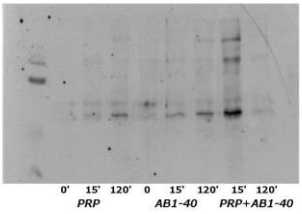

Fig. 3 Aβ40-PrP complex stimulates tyrosine phosphorylation in b.End3 endothelial cell line.

Because PrP is known to transduce signals through tyrosine kinase Fyn, and overexpression of Fyn has been reported to exacerbate AD phenotypes, Fyn is a natural candidate for mediating Aβ oligomers/PrP signaling [51]. Fyn signaling through other receptors was preserved in Prnp-/- neurons but no Aβ oligomers-induced Fyn activation was detected, implicating PrP as an indispensable intermediate in all Aβ oligomers/Fyn signaling. Another group demonstrated that Aβ oligomers/PrP-induced Fyn activation leads to Tau phosphorylation, potentially linking together the two major histopathological phenotypes of AD [52]. These data make prion-Abeta complex a new promising therapeutic target. By combining *in vitro* overexpressed prion and Abeta we have reconstituted a biologically active prion-Abeta complex (Fig. 3). In the course of the present project **we plan to minimize the complex to its well-folded core using new MS-based methods and characterize its structure by NMR** (in collaboration with dr I. Zhoukov, IBB)

Task 2. Histone pre-mRNA cleavage complex (collaboration with B. Marzluff, Z. Domiński North Carolina Univ. USA). A postdoc ½ time and one PhD student assigned.

In eukaryotes, expression of histone mRNAs is restricted to S-phase and is highly coordinated with DNA replication. Their concentration reaches peak values during S-phase to meet the massive demand for histones required for chromatin formation. Levels of histone mRNAs rapidly decline in G2-phase to gradually increase in G1-phase of the following cell cycle. **In most animal cells, the tightly controlled up-regulation of histone mRNA levels is achieved by combination of two S-phase specific events**: a nearly 5-fold increase in the rate of transcription of histone genes and the concomitant activation of specific 3' end processing that converts histone transcripts (pre-mRNAs) into mature and translationally active histone mRNAs.

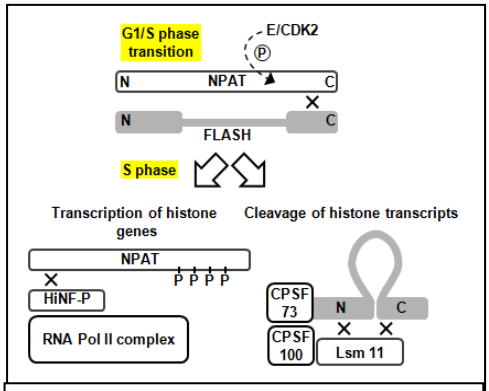

Fig. 4. The putative cycle of FLASH and NPAT interactions. Outside S-phase, FLASH and NPAT tightly interact through their C-terminal regions. The NPAT-binding site on FLASH overlaps with the Lsm11-binding site, which consists of both the N-terminal and C-terminal sequences. During the G1-S-phase transition, NPAT becomes phosphorylated at multiple sites within the C-terminal half by Cyclin E/CDK2 complex, causing its dissociation from FLASH. The released NPAT is predicted to engage into interactions with components of the transcriptional machinery to activate transcription of histone genes, whereas FLASH is now free to interact with Lsm11 and proteins involved in 3' end cleavage of histone transcripts

The enhanced **transcription of histone genes during S phase critically depends on Nuclear Protein, Ataxia-Telangiectasia locus (NPAT)**, a universal transcriptional co-activator of all 5 classes of histone genes. During the G1-S phase transition, NPAT becomes phosphorylated at multiple sites within the C-terminal half by the Cyclin E/Cyclin Dependent Kinase 2 (CDK2) complex (Fig. 4). This hyperphosphorylation ultimately results in enhanced transcription of histone genes but the underlying mechanism remains unknown [53]. 3' end processing of histone pre-mRNAs involves a single step endonucleolytic cleavage immediately downstream of a highly conserved stem-loop. The reaction critically depends the U7 snRNP and its integral component Lsm11 and is inactive throughout the majority of the cell cycle to be turned on as cells enter S phase, i.e. simultaneously with the activation

of histone gene transcription by NPAT. In this unique processing reaction, which gives rise to histone mRNAs, terminated with the stem-loop rather than a polyA tail, Lsm11 tightly binds the N-terminal region (amino acids 100-140) of a large protein, FLASH [54]. Together, these two proteins form a platform that recruits the 3' endonuclease, CPSF73, to U7 snRNP and histone pre-mRNA for cleavage. Changes in the composition of mammalian and *Drosophila* U7 snRNP that result from the interaction between Lsm11 and FLASH were identified in a common effort of the applicant's and dr Domiński's lab [25],[26]. **The interaction between Lsm11, FLASH and CPSF73 is a critical step in expression of replication-dependent histone genes** and one of the key events in S phase cells in all animals.

Recently, dr Dominski's group demonstrated that the C-terminal region of FLASH tightly interacts with the C-terminal region of NPAT, directly linking these two key regulators in histone gene expression (Fig. 4). More importantly, A. Skrajna, a Ph.D. student co-directed by the applicant and dr. Dominski, made an **intriguing discovery that the interaction of Lsm11 is not limited to the N-terminus of FLASH but extends to its C-terminal region** located nearly 1800 amino acids further downstream **and includes the region that interacts with NPAT**. This observation, providing a **pilot dataset for the present project** (Fig. 5), suggests that in full length FLASH both ends are directly juxtaposed, forming a continuous binding platform for Lsm11. Moreover, the interaction of Lsm11 with this platform might be allosterically regulated by the interaction of NPAT with the C-terminus of FLASH. Interestingly, human embryonal stem cells express a splice variant of FLASH, called MiniFLASH, in which the two opposite ends are directly linked, with the entire long central part being deleted due to exon skipping.

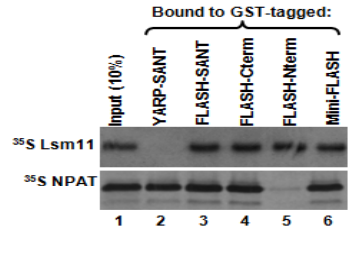

Fig. 5. FLASH C-terminus efficiently pulls down both NPAT and Lsm11 C-terminal parts (lane 4), whereas FLASH N-terminal part pulls down only Lsm11 C-terminus. GST pull down of ^{35}S-labeled N-terminal Lsm11 (top panel) and C-terminal NPAT (bottom panel) by GST-tagged proteins, as indicated. SANT domain is a smaller fragment of the C-terminal FLASH

We envision a model in which NPAT and FLASH are functionally latent in the heterodimer and inactive in their respective functions, i.e. transcription and processing, respectively. Phosphorylation of NPAT by the Cyclin E/ CDK2 complex at the onset of the S-phase would loosen the interaction with FLASH thus liberating each component of the heterodimer for a timely synthesis of histone mRNAs. A sequence highly similar to the C-terminal region of FLASH exists at the end of an unrelated protein, YARP, predicted to function as repressor of histone gene transcription. This sequence is also capable of interacting with NPAT. This suggests that the **interaction of NPAT with either FLASH or YARP may provide an on/off switch that regulates expression of histone genes during cell cycle and development.** Strikingly, as few as 16 last amino acids of NPAT are sufficient to strongly and specifically interact with the C-terminal domain of FLASH and YARP. This short sequence is highly conserved in all vertebrate orthologues of NPAT but does not resemble any known domain in the database. The C-terminal end of *Drosophila* NPAT shares no recognizable similarity with the vertebrate NPAT yet it strongly interacts with the C-terminus of *Drosophila* YARP. This indicates that the C-terminal region of NPAT in all animals likely adopts a unique fold that is capable of interacting with the structural domain shared by FLASH and YARP.

The **aim of this task is to explore in structural terms the network of interactions** that was identified during the on-going collaborative effort of the applicant group and dr Domiński's group. These studies should be instrumental in gaining **important insights into the mechanism that coordinates transcription of histone genes with processing of histone pre-mRNAs**. MS-based methods will be used to identify the structured domains of the new proteins of interest, and their regions of contact. This will allow to minimize the constructs to crucial regions, small enough that high resolution NMR structural studies can be applied. Such "minimization" strategy has already been implemented by us (with participation of dr I. Zhoukov, IBB PAN) in the NMR-based studies the structure of Lsm11 alone and in complex with the interacting domain of FLASH (on-going research). We would like to **extend our collaboration to investigate the new and potentially very important interactions involving NPAT, FLASH and YARP**. **Of particular interest for us is the structure of the C-terminal region of NPAT and the mechanism of its interaction with the C-terminal regions of FLASH and YARP** in both vertebrates and invertebrates. We would also like to study the

possibility that the interaction between FLASH and Lsm11 is mutually exclusive with the interaction between FLASH and NPAT and regulated by phosphorylation of NPAT by the Cyclin E/CDK2 complex. This part would require mostly *in vivo* experiments and would be conducted by dr Dominski's group at UNC in the frame of NIH grant.

We constructed a number of clones that express large amounts of the C-terminal regions of FLASH, YARP and NPAT in bacteria and established conditions of purifying these proteins in amounts suitable for structural studies. In addition, we chemically synthesized large amounts of a minimal region of NPAT capable of tightly interacting with FLASH and YARP (last 31 amino acids) and tested suitability of this peptide for structural studies using pull down assays. We also bacterially expressed several variants of MiniFLASH in which the N- and C-terminal regions are linked either directed or separated by the central FLASH region of various length. These extended MiniFLASH variants will be used in competition experiments with Lsm11 and various C-terminal fragments of NPAT in either native state or hyperphosphorylated in vitro by recombinant E/CDK2 complex.

Task 3. Pro-Centriole complex (a collaborative effort with prof. D. Glover, Cambridge Univ, UK). A postdoc ½ time and one PhD student assigned.

The centriole is the 9-fold symmetrical structure found at the core of centrosomes and at the base of cilia. Defects in centrosomes are seen in cancer and genetically inherited disease [55]. **Animal cells leave mitosis with two disengaged centrioles that are partially duplicated during G1, enabling a cell to enter the next mitosis with two mother-daughter pairs of centrioles, each pair found at the spindle poles. The daughters mature during mitosis and only then disengage from the mothers during telophase.** This permits the recruitment of Ana2 so that it occupies a single site on both mother and daughter centriole. This in turn results in recruitment of Sas6 that we have shown binds to Ana2 that has been phosphorylated by Plk4 ([56] – supported by a common "Harmonia" grant). The new procentriole is established by 9 dimers of Sas6, the N-termini of which interact at the core to establish symmetry and the C-terminal part of which is largely intertwined coiled coils. The terminal part of these tails interact with a ring of Sas4 at the wall that binds the centriolar microtubules (reviewed in [57]). In *Drosophila* cells, Sas6 remains in the lumen of the centriole throughout the cell cycle whereas in human cells, Sas6 is only found at the most distal part of the centriole and is released in a Plk4 dependent step [58]. This has led to the proposal that the centriole wall templates the formation of the Sas6-based cartwheel structure. So **9-fold symmetry could arise in two ways – either established by the centriole wall that can guide Sas6 assembly – or as an inherent property of Sas6 itself**, a property that is of particular importance in the *de novo* assembly of centrioles. In human cells, Cep135 binds to both Sas6 and Sas4 [59] leading to the notion that this complex is key for linking the cartwheel to the microtubule wall. Cep135/Bld10 is essential for cartwheel formation and centriole duplication in *Chlamydomonas* and *Paramecium* and is required for excessive centriole duplication in Plk4-overexpressing human cells [57]. When Cep135 is absent in *Drosophila*, centrioles are able to form but are shorter in length.

In addition to the structure formed by Sas6, Ana2 and Sas4, **we have recently identified another complex between three centriolar proteins, Cep135, Ana1 and Asterless (Asl)**. Ana1 is a poorly characterized protein essential for centriole duplication in *Drosophila* [60]. Asl binds to the cryptic Polo Box of Plk4 and is required to recruit Plk4 to the centriole [61]. Several lines of **our pilot**

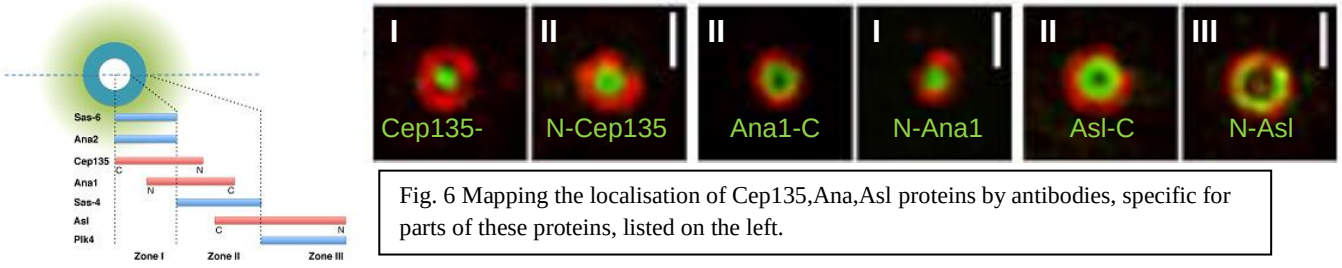

Fig. 6 Mapping the localisation of Cep135,Ana,Asl proteins by antibodies, specific for parts of these proteins, listed on the left.

data support the idea of the Cep135-Ana1-Asl complex extending from core to periphery. First, studies with antibodies specific to the terminal parts of these proteins or GFP tags at either termini indicate that their N- and C-termini to lie at different locations and in fact span the four zones defined by previous analyses of the 3D localisation of centriolar proteins and PCM (Fig. 6) by structured

illumination microscopy [62]. Thus Cep135 has its C-terminus at the centre of the cylinder in zone I whilst its N-terminus extends into zone II. The N-terminus of Ana1 is also found within zone I while its C-terminus extends even further into zone II. Finally the C-terminus of Asl is found in zone II with its N-terminus extending into zone III.

The above observation led us to ask whether these **particular three proteins showed interactions**. We showed they do in three ways: by direct *in vitro* binding of one ^{35}S labelled protein expressed by IVTT to its bacterially expressed partner protein on beads (Fig. 7A); by showing Flag-tagged Ask and Cep135 could be pulled down in a GFP-trap when Ana1 is GFP-tagged (Fig. 7B); and using a new assay we have developed to screen for complex formation *in vivo*. In this last approach, we transiently express pairs of exogenous centriolar proteins in cultured cells (at levels too high for incorporation into centrioles) and ask whether they become incorporated into common aggregates. For example, if exogenous GFP-tagged Cep135 and RFP-tagged Asl are co-expressed, they form independent complexes of difference appearance in wide-field fluorescence microscopy. If however also exogenous Ana1 is co-expressed, then Cep135 and Asl co-associate in a common aggregate (Fig. 7C)

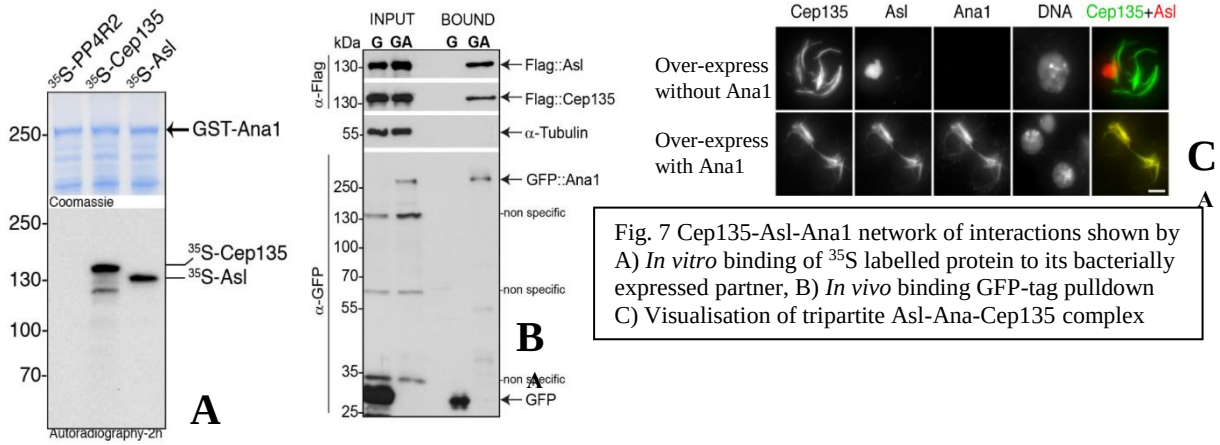

Fig. 7 Cep135-Asl-Ana1 network of interactions shown by A) *In vitro* binding of ^{35}S labelled protein to its bacterially expressed partner, B) *In vivo* binding GFP-tag pulldown C) Visualisation of tripartite Asl-Ana-Cep135 complex

Taking this assay one step further by expressing each terminally tagged "half-protein", we find that the N-terminal part of Cep135 binds to the N-terminal part of Ana1 and the C-terminal part of Ana1 binds the C-terminal part of Asl. Thus **Ana1 provides a molecular link between Cep135 in the very interior of the centriole and Asl at the junction with pericentrolar material** (Fig. 8).

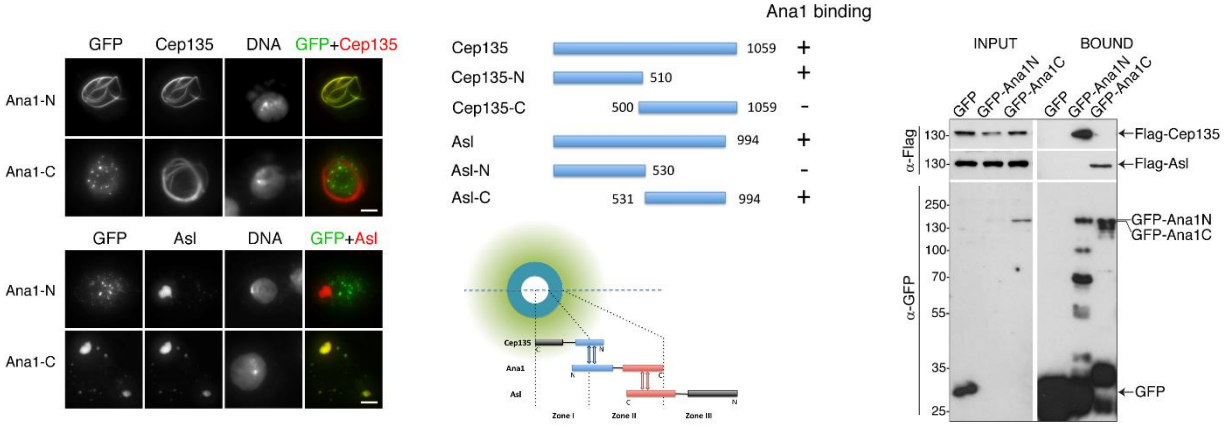

Fig. 8 Assays as described in fig. 7 carried out for N- and C-terminal domains of Asl and Cep135 proteins show that N-terminal part of Cep135 binds to the N-terminal part of Ana1 and the C-terminal part of Ana1 binds the C-terminal part of Asl.

Consistent with these physical interactions, Asl is recruited to the centriole later than Ana1 and moreover, depletion of Ana1 leads to a failure to recruit Asterless to the maturing daughter centriole in cultured cells. We also know that the C-terminal region of Asl will also interact with the N-terminal region of Sas4 on the centriole wall. Indeed, antibodies that bind to Sas4 prevent recruitment of Asl and *vice versa* [63]. Pilot data also shows that Asl also interacts with Ana2 in a manner dependent upon Plk4 phosphorylation. Thus **Asl makes complex interactions with at least three other key centriolar proteins**.

Why is this important? Asterless is a critical centriolar component – as a physical partner of Sas4, it required to recruit peri-centriolar material and as a binding partner of Plk4, it also plays key roles in recruiting Plk4 to initiate centriole duplication. We therefore need to understand more about precisely how Asl interacts with its partners, how these interactions are regulated and how they relate to centriole duplication and function.

To decipher the precise nature of the physical interactions between these proteins required to build this molecular machine, protein crystallography generally delivers the best quality description. In some cases, however, obtaining crystals or high resolution diffraction on crystals proves to be very difficult or impossible. **Our experience in studying complexes formed between coiled-coil-protein complexes between the centromeric protein CENP-C and its partners in the Mis12 complex of the *Drosophila* kinetochore** (Richter, M. et al, in preparation), indicated that we can **overcome some of the difficulties inherent in protein crystallization using HDex.** The task then is to study the points of contact between the centriolar proteins, Cep135, Ana1, Asl and Sas4. In each case, the interacting segments have already been identified and these domains can be expressed in a variety of expression systems. **HDX will define points of interact** and so allow mutations to be made at these sites and so study the consequences in either cultured *Drosophila* cells or in transgenic flies.

Besides the HDex approach to elucidate the structural features of these centriolar complexes, the reconstituted protein assemblies will be subjected to other biophysical assays, which may provide more new and valuable data. Whenever possible we will minimize the complexes to make them tractable by classis atomic level methods. Overall, using those parallel approaches, we will be able to characterize procentriole complex and its close interactors with high precision.

We fully anticipate that these interacts we study will be in part regulated by the phosphorylation state of the proteins. Already in our common "Harmonia" funded project we have shown that Ana2 is a major substrate of Plk4 and mapped the phosphorylation sites. This has led to the finding that phosphorylation of Ana2 is essential for recruitment of Sas6 to the procentriole [56]. Our pilot data indicates that in addition to the Plk4 phosphorylation sites that we have recently characterized on Ana2, **Plk4 also phosphorylates other sites on Ana2 in addition to sites on Asl itself. Thus, we need to use other MS approaches to map these new sites** in order to be able to study their importance for complex formation.

The projects outlined above, are embedded into the larger research programme being developed in the laboratory of Professor Glover. There is extensive cross-feeding both of methodologies and research output between different projects within that lab. It is therefore likely that the experimental design presented in our proposal can be improved or modified as our knowledge base develops. From its side, the proteomics laboratory in Warsaw constantly implements new MS-based techniques that have potential to add new value to the projects. This emphasizes a need for regular and close communication between our groups. We expect that members of our laboratories should and will visit collaborators at their places of work. This is very important to have a full understanding of each other's goals and priorities, as well as to develop new ways of problem solving and tackling new challenges.

Task 4. Intermediate Filament proteins and their interactors. A postdoc ½ time and MSc student.

Intermediate filaments (IFs) are assembled from a large family of structurally related but sequence-wise very different fibrous proteins. They constitute the **principal cytoplasmic and nuclear filament system** that is responsible for the plasticity of the cell [64]. In addition, IFs are also engaged in multiple signaling pathways [65]. At present, mutations in IF protein–encoding genes have been indicated as causative for nearly 100 human inherited diseases [66]. Therefore, IF proteins are **considered as potential drug targets** and structural studies may provide important insight for drug design process. However, due to their intrinsic oligomerisation propensity, IF proteins are DTs, which makes them poorly accessible for classic tools. The three-dimensional architecture of IF protein oligomers and fibrils has been extensively studied by a number of low-resolution techniques. At the atomic level, the structural knowledge of IF protein dimers

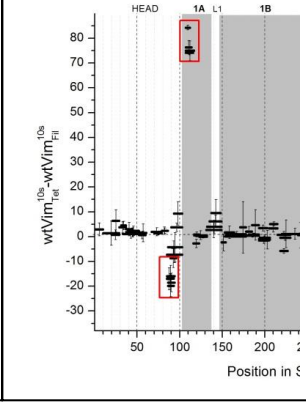

Fig. 9 Difference in % of deuter uptake (vertical axis) after 10 sec. of incubation in vimentin tetramers and filaments (N-terminal region up to position 250 is shown). Upper red rectangle marks a region, spanning positions 105-120 in the sequence, strongly stabilized in filament and lower red rectangle a nearby region slightly destabilized in filament, as compared to tetramers.

has been obtained by a "divide and conquer" X-ray crystallization approach [67], i.e. by combining the available structures of their partially overlapping fragments, whereas **high resolution data on authentic, full-length IFs is not available. Using HDex allowed us to overcome this difficulty.** We have compared HDex patterns for the full length homooligomers of vimentin and heterooligomers of keratins 8/18 (K8/K18) in conditions where low-order (dimers, tetramers) oligomers are preferred and in other condition with prevalence of higher order oligomers and filaments. **By monitoring HDex on authentic IF proteins we could for instance precisely map regions engaged in stabilization of different oligomeric forms, from monomers to filaments** in vimentin (Fig. 9) directly proving what previously could only be hypothesized. HDex thus to be an efficient tool for structural studies of this class of proteins. Vimentin is a major constituent of IF in normal mesenchymal cells, an important marker of epithelial to mesenchymal transition in cancer. Vimentin, abundant in early myoblasts becomes substituted by desmin differentiated in mature muscle cells and is known to coassemble with desmin. In the present project we **plan to extend these studies to vimentin-desmin** system [68], a non-sarcomeric component of muscle cell cytoskeleton and their known pathologic mutants (selected from 67 known), leading to severe myopathies, including cardiomyopathies [69], with the aim to better understand the rules of IF assembly and its derangements. The project will be consulted by prof. H. Herrmann, DKFZ, Heidelberg, Germany, expert in the field, who has a collection of overexpression plasmids of numerous IF proteins and their mutants. We will study coassembly phenomena with their IF partners forming mixed filaments with desmin, like nestin, synemin, and other proteins like plakin repeat of plectin. In collaboration with prof. A, Edelman, Institut Necker, Paris, France we have also mapped the region of contact in keratin 8 responsible for interaction with ΔF508NBD1 domain of chloride channel CFTR (A. Kupniewska et al., in preparation). This interaction, discovered in our previous common effort [70], is thought to be responsible for titrating out the ΔF508CFTR for degradation and thus causing cystic fibrosis. A competing K8 complex with α1-antitrypsin has been detected recently and will also be studied in the present project to elucidate the structural basis for K8 unexpected scavenging function.

D. Project methodology Task 5 - Improving the toolbox. **5A**. Hydrogen deuterium exchange of backbone amide protons monitored by mass spectrometry (HDexMS). Informatician assigned.

The idea of application of this approach to protein structure studies is as old as the very idea that proteins might have a 3D structure. As early as 1954 [71] Kai Linderstrom-Lang exploited HDex to verify Linus Pauling's speculation that a network of intra-chain hydrogen-bonds might restrict chain entropy and impose stable structure on proteins. Over the years the method developed slowly, being rather a bystander of an enormous success of X-ray crystallography and NMR. Different strategies to follow HDex in proteins were used with the most modern being NMR and MS [72]. **At present HDexMS is an established alternative method of protein structure analysis**, **allowing to overcome the above mentioned limitations of the classic methods, being especially well suited to study IDP's** [73]. HDexMS probes the susceptibility of main-chain amide protons to exchange with bulk solvent and thus allows to map their burial and stability of hydrogen bonded networks in different regions of the protein or in other words to quantitate their levels of dynamics**. A protein or its complex, also with non-protein partners are studied in native buffer conditions in low micromolar concentration, in principle without mass limit.** The method, however, is characterized at present by medium level resolution, since the exchange is measured for proteolytic peptides of the protein of interest and not for single amino acids. **Very recent advancements** in data acquisition (high resolution MS instruments, with ion mobility (IM) separation device, ETD fragmentation capability, automated HDex control units) and data processing (multiple software solution for reliable automation of mass-shift extraction from Ion-Mobility separated MS spectra) **allow to characterize complex protein assemblies in structural terms in a time of days** [74]. This developments for instance enable an exhaustive mapping of contact sites within a complex protein assembly for a large number of constructs in a number of buffer conditions, all in reasonable time. Applicant's lab is experienced in application of this method for a variety of protein complexes [4],[1],[24],[21],[28],[2], [11], operating on two MS (Synapt G2) systems equipped with HDex control units and IM option.

To **improve resolution of the method to single amide protons,** application of electron transfer dissociation (ETD) was proposed [75], to eliminate deuterium position scrambling, unavoidable with low-energy collision induced dissociation. Usage of ETD was shown for a few proteins [76], however

some limitations slow down its wider application. At present the most successful procedure includes classic low-resolution HDex followed by ETD fragmentation of selected peptides covering regions of special interest. **Data analysis of ETD results is also a limiting factor**, as at present it is restricted to manual feature extraction and recalculation of deuterium content at each amide without the aid of automated software tools. In the project **we plan to fill this gap working out a tool for automated mass-shift extraction procedure for each amide proton from ETD fragmentation spectra** measured after HDex. ETD data analysis requires taking into account the fact that each fragmentation ion mass shift includes information on exchange of several amides. Thus, the least-squares fitting procedure will be applied to select the most probable distribution of deuteria along the peptide sequence which best explains a set of observed mass shifts in the entire ETD fragmentation spectrum.

5B. Oxidative footprinting. Application of oxidative protein footprinting for protein structure studies is based on the observation, that the efficiency of modification of residue side chains of peptides and proteins by reactive oxygen species depends of accessibility of these residues to solvent [77]. In this method, hydroxyl radicals react with proteins to yield stable oxidative modifications of solvent accessible amino acid side chains. After proteolysis, liquid chromatography coupled to tandem MS (LC-MS-MS/MS) analysis is performed to identify and quantify the modification sites [78]. This approach brings information complementary to HDex as it tests the status of selected side-chains, and is especially valuable when the macromolecular complexes are of interest. Several approaches to generate hydroxyl radicals for footprinting are currently in use. The synchrotron X-ray and ^{137}Cs γ-ray methods generate hydroxyl radicals as the major reactive products from water radiolysis by high-energy photons [79]. Other authors reported similar methods of fast photochemical oxidation of proteins (FPOP) that generate ·OH by photolysis of mM hydrogen peroxide with a pulsed laser (either 248-nm KrF excimer laser or 266-nm frequency quadrupled Nb YAG) and react them with protein in a flow system [80]. Due to the limited access to the high energy beam line on a synchrotron or excimer laser, other labs are focusing on reactions that form hydroxyl radicals by chemical processes, including the Fenton chemistry, where hydroxyl radicals are generated by the reaction of a redox-active metal ion complex with hydrogen peroxide [81]. The use of this method is limited by numerous side reactions making the data analysis difficult. **We plan to test a new method of generation of hydroxyl radicals, recently worked out in the Institute of Physics, Pol. Acad. Sci.** [82] **which might overcome these difficulties**

5C. Covalent cross-links. Cross-linkers usually molecules with a defined length and reactive group at each end of the cross-linker, usually of high specificity towards selected amino acid type (primary amines like lysine). During the chemical cross-linking, a covalent linkage (the cross-link) between two distinct protein sites, either intramolecular or intermolecular is introduced. In principle the reaction is expected only for proximal protein sites sterically compatible with the dimension of the cross-linking agent. Subsequently cross-linked proteins are digested with protease and analyzed by LC/MS, resulting in identification of cross-linked peptides. When identified, cross-linked peptides deliver information which amino acid residues are in a close proximity (dependent from the cross-linker length), hence providing low-resolution structural information, that might support protein structure modeling, or identify proteins forming protein complexes and indicate inter-protein contact sites [83]. Cross-linking experiments consume relatively low amount of proteins (usually few micrograms per experiment) and the whole workflow can be accomplished within a week.

Extraction and identification of cross-linked peptides (XLs) in LC-MS datasets presents considerable challenge. LC-MS data usually contain numerous signals and combinatorial space of possible masses of XLs (increasing strongly for larger protein complexes) is enormous so the probability of false positives is very high in case of cross-linking experiment. Number of false positives cannot be decreased by the analysis of fragmentation spectra as these are usually of poor quality for XLs. Instead, to decrease the false positive rate a pair of two cross-linkers can be used, one regular (D0), paired with its isotopically labelled counterpart (D4) of different mass. Two dimensional representation of LC-MS data

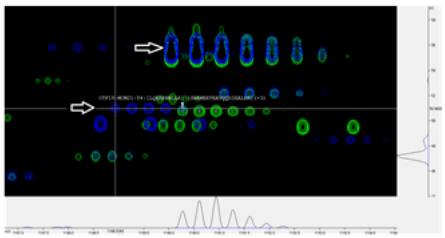

Figure 10. Heat map from cross-linker analysis using MSparky program. Peptide isotopic envelopes are marked with white arrows. Blue isotopic envelopes originate from the D0 cross-linker labelled peptides and green from D4 cross-linked peptides. Non-labeled peptides are overlapping (upper envelope), whereas cross-linked peptides present shift proportional to charge and mass increase (lower envelope).

facilitates unequivocal identification D0-D4 pair of XLs as can be exemplified by use of an in-house LC-MS data analysis program MSparky [48] (Fig. 10) which is still in development. Such analysis ensures selection of true XLs for further analysis, but still requires automation which is planned in the frame of the present project in the form of additional procedure in MSparky, as manual analysis is time consuming and error-prone.

All these methods bring complementary information in the form of experimental structural constraints characterizing involvement of side chains and main chain amide protons in the structure. Obtained data can be integrated into one set of constraints modifying for instance force fields used in molecular modelling. In collaboration with prof. J. Poznański, IBB we plan to work out such procedures. For this purpose the original force field of Yasara Structure package [84] will be extended to incorporate MS-based structural constraints. Residue-specific protection of HDex can be easily incorporated as additional terms weighting contribution of either solvent accessible surface area (SASA) or H-bonding interactions. The procedures will be tested using datasets for proteins of known structure, like for instance Rack1 [11].

Numerous classic methods of structure analysis and all molecular biology techniques will also be available for the project including, NMR, AUC, EM, AFM, CD, fluorescence, SEC, The group is experienced in all molecular biology tools of heterologous protein overexpression and purification necessary for structural studies.

Literature:

[1] Kupniewska-Kozak A, Gospodarska E, Dadlez M. Intertwined structured and unstructured regions of exRAGE identified by monitoring hydrogen-deuterium exchange. J Mol Biol 2010;403:52–65.

[2] Sitkiewicz E, Tarnowski K, Poznański J, Kulma M, Dadlez M. Oligomerization interface of RAGE receptor revealed by MS-monitored hydrogen deuterium exchange. PloS One 2013;8:e76353.

[3] Sitkiewicz E, Kłoniecki M, Poznański J, Bal W, Dadlez M. Factors Influencing Compact-Extended Structure Equilibrium in Oligomers of Aβ1-40 Peptide. J Mol Biol 2014, 426,2871

[4] Kacprzyk, O., et al. Crucial Role of Perfringolysin O D1 Domain in Orchestrating Structural Transitions Leading to Membrane-Perforating Pores. J.Biol.Chem. pii: jbc.M114.577981. [Epub ahead of print]

[5] Peng Z, et al. Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. Cell Mol Life Sci CMLS 2014..

[6] Oldfield CJ,et al. Comparing and combining predictors of mostly disordered proteins. Biochemistry (Mosc) 2005;44:1989–2000.

[7] Iakoucheva LM, et al. Intrinsic disorder in cell-signaling and cancer-associated proteins. J Mol Biol 2002;323:573.

[8] Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 2005;6:197–208.

[9] Fischer, E. Einfluss der configuration auf die wirkung der enzyme. Ber Dt Chem Ges 27:2985–2993 n.d.

[10] Le Gall T, et al. Intrinsic disorder in the Protein Data Bank. J Biomol Struct Dyn 2007;24:325–42.

[11] Tarnowski K, Dadlez M, Kaus-Drobek M. Patterns of structural dynamics in RACK1 protein retained throughout evolution: a HDex study of three orthologs. Protein Sci 2014;23:639–51.

[12] Uversky VN. A decade and a half of protein intrinsic disorder.Protein Sci 2013;22:693

[13] Xie H, et al. Functional anthology of intrinsic disorder. J Proteome Res 2007;6:1917–32.

[14] Cortese MS, et al. Intrinsic disorder in scaffold proteins. Prog Biophys Mol Biol 2008;98:85.

[15] Minezaki Y, Homma K, Kinjo AR, Nishikawa K. Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. J Mol Biol 2006;359:1137–49.

[16] Xue B, Dunker AK, Uversky VN. The roles of intrinsic disorder in orchestrating the Wnt-pathway. J Biomol Struct Dyn 2012;29:843.

[17] Peng Z, Xue B, Kurgan L, Uversky VN. Resilience of death: intrinsic disorder in proteins involved in the programmed cell death. Cell Death Differ 2013;20:1257–67.

[18] Kłoniecki M, Jabłonowska A, …Dadlez, M. Ion mobility separation coupled with MS detects two structural states of Alzheimer's disease Aβ1-40 peptide oligomers. J Mol Biol 2011;407:110.

[19] Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys 2008;37:215–46.

[20] Hammoudeh DI, Follis AV, Prochownik EV, Metallo SJ. Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. J Am Chem Soc 2009;131:7390–401.

[21] Odolczyk N, et al. Discovery of novel potent ΔF508-CFTR correctors that target the nucleotide binding domain. EMBO Mol Med 2013;5:1484–501.

[22] Uversky VN. Natively unfolded proteins:biology waits for physics. ProteinSciPublProteinSoc 2002;11:739

[23] Walzthoeni T, et al. MS supported determination of protein complex structure. Curr Opin Struct Biol 2013;23:252
[24] Mierzejewska K, et al. Structural basis of the methylation specificity of R.DpnI. Nucleic Acids Res 2014.
[25] Sabath I, Skrajna A, Yang X-C, Dadlez M, Marzluff WF, Dominski Z. 3'-End processing of histone pre-mRNAs in Drosophila: U7 snRNP is associated with FLASH and polyA factors. RNA 2013;19:1726–44..
[26] Yang X-C, et al. A complex containing the CPSF73 endonuclease and other polyA factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. Mol Cell Biol 2013;33:28–37.
[27] Niedzwiecka A, et al. Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. J Mol Biol 2002;319:615–35.
[28] Rutkowska-Wlodarczyk I et al. Structural changes of eIF4E upon binding to the mRNA 5' monomethylguanosine and trimethylguanosine Cap. Biochemistry (Mosc) 2008;47:2710–20.
[29] Wawer I, et al. Regulation of Nicotiana tabacum osmotic stress-activated protein kinase and its cellular partner GAPDH by nitric oxide in response to salinity. Biochem J 2010;429:73–83..
[30] Walsh DM, Selkoe DJ. A beta oligomers - a decade of discovery. J Neurochem 2007;101:1172–84.
[31] Benilova I, et al. The toxic Aβ oligomer and Alzheimer's disease: Nat Neurosci 2012;15:349.
[32] Mucke L, Selkoe DJ. Neurotoxicity of amyloid β-protein: synaptic and network dysfunction. Cold Spring Harb Perspect Med 2012;2:a006338.
[33] Huang TH, Yang DS, Fraser PE, Chakrabartty A. Alternate aggregation pathways of the Alzheimer beta-amyloid peptide. An in vitro model of preamyloid. J Biol Chem 2000;275:36436–40.
[34] Woods LA, Radford SE, Ashcroft AE. Advances in ion mobility spectrometry-mass spectrometry reveal key insights into amyloid assembly. Biochim Biophys Acta 2013;1834:1257–68.
[35] Sitkiewicz E, Olędzki J, Poznański J, Dadlez M. Di-tyrosine cross-link decreases the collisional cross-section of aβ Peptide dimers and trimers in the gas phase: an ion mobility study. PloS One 2014;9:e100200.
[36] Watt AD, Villemagne VL, Barnham KJ. Metals, membranes, and amyloid-β oligomers: key pieces in the Alzheimer's disease puzzle? J Alzheimers Dis JAD 2013;33 Suppl 1:S283–93.
[37] Lau T-L, Ambroggio EE, Tew DJ, Cappai R, Masters CL, Fidelio GD, et al. Amyloid-beta peptide disruption of lipid membranes and the effect of metal ions. J Mol Biol 2006;356:759–70.
[38] Arispe N, Rojas E, Pollard HB. Alzheimer disease amyloid beta protein forms calcium channels in bilayer membranes: blockade by tromethamine and aluminum. Proc Natl Acad Sci U S A 1993;90:567–71.
[39] Kumar A, et al. Specific soluble oligomers of amyloid-β peptide undergo replication and form non-fibrillar aggregates in interfacial environments. J Biol Chem 2012;287:21253–64.
[40] Sorci G, et al. RAGE in tissue homeostasis, repair and regeneration. BiochimBiophys Acta 2013;1833:101.
[41] Matrone C, Djelloul M, Taglialatela G, Perrone L. Inflammatory risk factors and pathologies promoting Alzheimer`s disease progression: is RAGE the key? Histol Histopathol 2014.
[42] Donato R. RAGE: a single receptor for several ligands and different cellular responses: the case of certain S100 proteins. Curr Mol Med 2007;7:711–24.
[43] Koch M, Chitayat S, Dattilo BM, Schiefner A, Diez J, Chazin WJ, et al. Structural basis for ligand recognition and activation of RAGE. Struct Lond Engl 1993 2010;18:1342–52.
[44] Yatime L, Andersen GR. Structural insights into the oligomerization mode of the human receptor for advanced glycation end-products. FEBS J 2013;280:6556–68.
[45] Wei W, Lampe L, Park S, Vangara BS, Waldo GS, Cabantous S, et al. Disulfide bonds within the C2 domain of RAGE play key roles in its dimerization and biogenesis. PloS One 2012;7:e50736.
[46] Lenarčič Živković M, et al. Post-translational S-nitrosylation is an endogenous factor fine tuning the properties of human S100A1 protein. J Biol Chem 2012;287:40457–70..
[47] Hanford LE, et al. Purification and characterization of mouse soluble receptor for advanced glycation end products (sRAGE). J Biol Chem 2004;279:50019–24.
[48] Bakun M, et al. An integrated LC-ESI-MS platform for quantitation of serum peptide ladders. Application for colon carcinoma study. Proteomics Clin Appl 2009
[49] Benilova I, De Strooper B. Prion protein in Alzheimer's pathogenesis: EMBO Mol Med 2010;2:289–90.
[50] Kessels HW, et al. The prion protein as a receptor for amyloid-beta. Nature 2010;466:E3–4;
[51] Um JW, et al. Alzheimer amyloid-β oligomer bound to postsynaptic prion protein activates Fyn to impair neurons. Nat Neurosci 2012;15:1227–35.
[52] Larson M, et al. The complex PrP(c)-Fyn couples human oligomeric Aβ with pathological tau changes in Alzheimer's disease. J Neurosci 2012;32:16857
[53] Ma T, et al. Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. Genes Dev 2000;14:2298–313.
[54] Yang X-C, et al. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. Mol Cell 2009;36:267–78.
[55] Bettencourt-Dias M, et al. Centrosomes and cilia in human disease. Trends Genet TIG 2011;27:307

[56] Dzhindzhev, N. et al. Plk4 phosphorylates Ana2 to trigger Sas6 recruitment and procentriole formation. Curr. Biol. in press
[59] Lin Y-C, Chang C-W, Hsu W-B, Tang C-JC, Lin Y-N, Chou E-J, et al. Human microcephaly protein CEP135 binds to hSAS-6 and CPAP, and is required for centriole assembly. EMBO J 2013;32:1141–54.
[60] Dobbelaere J, Josué F, Suijkerbuijk S, Baum B, Tapon N, Raff J. A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in Drosophila. PLoS Biol 2008;6:e224.
[61] Dzhindzhev NS, et al. Asterless is a scaffold for the onset of centriole assembly. Nature 2010;467:714.
[62] Fu J, Glover DM. Structured illumination of the interface between centriole and peri-centriolar material. Open Biol 2012;2:120104.
[63] Novak ZA, Conduit PT, Wainman A, Raff JW. Asterless licenses daughter centrioles to duplicate for the first time in Drosophila embryos. Curr Biol CB 2014;24:1276–82.
[64] Herrmann H, et al. Intermediate filaments: primary determinants of cell architecture and plasticity. J Clin Invest 2009;119:1772–83.
[65] Colas J, Faure G, Saussereau E, Trudel S, Rabeh WM, Bitam S, et al. Disruption of cytokeratin-8 interaction with F508del-CFTR corrects its functional defect. Hum Mol Genet 2012;21:623–34.
[66] Szeverenyi I, et al. The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. Hum Mutat 2008;29:351–60.
[67] Strelkov SV, et al. Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. J Mol Biol 2001;306:773
[68] Clemen CS, et al. Desminopathies: pathology and mechanisms. Acta Neuropathol (Berl) 2013;125:47–75.
[69] Bär H, et al Impact of disease mutations on the desmin filament assembly. J Mol Biol 2006, 360, 1031
[70] Davezac N, et al. Global proteomic approach unmasks involvement of keratins 8 and 18 in the delivery of CFTR/deltaF508-CFTR to the plasma membrane. Proteomics 2004;4:3833–44.
[71] A. Hvidt, K. Linderstrom-Lang. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. Biochim. Biophys. Acta 14 (1954) 574-575.
[72] Katta V, Chait BT. Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. Rapid Commun Mass Spectrom RCM 1991;5:214–7.
[73] Balasubramanian D, Komives EA. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. Biochim Biophys Acta 2013;1834:1202–9..
[74] Iacob RE, Engen JR. Hydrogen Exchange Mass Spectrometry. J Am Soc Mass Spectrom 2012;23:1003.
[75] Rand KD, Zehl M, Jensen ON, Jørgensen TJD. Protein hydrogen exchange measured at single-residue resolution by electron transfer dissociation mass spectrometry. Anal Chem 2009;81:5577–84.
[76] Landgraf RR, et al. Automated HDex Electron Transfer Dissociation High Resolution Mass Spectrometry Measured at Single-Amide Resolution. J Am Soc Mass Spectrom 2012;23:301–9.
[77] Garrison WM, Radiation-induced oxidation of protein in aqueous solution. Radiat Res 1962;16:483–502.
[78] Guan J-Q, Chance MR. Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry. Trends Biochem Sci 2005;30:583–92.
[79] Kiselar JG, Maleknia SD, Sullivan M, Downard KM, Chance MR. Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. Int J Radiat Biol 2002;78:101–14.
[80] Hambly DM, Gross ML. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. J Am Soc Mass Spectrom 2005;16:2057–63.
[81] Sharp JS, Becker JM, Hettich RL. Protein surface mapping by chemical oxidation: structural analysis by mass spectrometry. Anal Biochem 2003;313:216–25.
[82] Morawski O, Izdebska K, Karpiuk E, Nowacki J, Suchocki A, Sobolewski AL. Photoinduced water splitting with oxotitanium tetraphenylporphyrin. Phys Chem Chem Phys PCCP 2014;16:15256–62.
[83] Merkley ED, Cort JR, Adkins JN. Cross-linking and mass spectrometry methodologies to facilitate structural biology: finding a path through the maze. J Struct Funct Genomics 2013;14:77–90.
[84] Krieger E, Joo K, Lee J, Lee J, Raman S, et al. (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins: Structure, Function, and Bioinformatics 77: 114–122