Enhancing top-down proteomics of brain tissue with FAIMS

Authors

James M. Fulcher^a, Aman Makaju^b, Ronald J. Moore^a, Mowei Zhou^c, David A. Bennett^d, Philip L. De Jager^e, Wei-Jun Qian^a, Ljiljana Paša-Tolić^c, Vladislav A. Petyuk^{a,*}

*To whom correspondence should be addressed:

vladislav.petyuk@pnnl.gov ,+1 (509) 371-6223

Affiliations

^a Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA

^b Life Sciences Mass Spectrometry Unit, Thermo Fisher Scientific, San Jose, CA 95134, USA

^c Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, USA

^d Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL 60612, USA

^e Department of Neurology, Center for Translational & Computational Neuroimmunology, Columbia University Medical Center, New York, NY 10032, USA

Keywords

FAIMS, differential mobility spectrometry, top-down proteomics, Alzheimer's, brain tissue, ion mobility

Abstract

Proteomic investigations of Alzheimer's and Parkinson's disease have provided valuable insights into neurodegenerative disorders. Thus far, these investigations have largely been restricted to bottom-up approaches, hindering the degree to which one can characterize a protein's "intact" state. Top-down proteomics (TDP) overcomes this limitation, however it is typically limited to observing only the most abundant proteoforms and of a relatively small size. Therefore, offline fractionation techniques are commonly used to reduce sample complexity, limiting throughput. A higher throughput alternative is online fractionation, such as gas phase high-field asymmetric waveform ion mobility spectrometry (FAIMS). Utilizing a high complexity sample derived from Alzheimer's disease brain tissue, we describe how the addition of FAIMS to TDP can robustly improve the depth of proteome coverage. For example, implementation of FAIMS at -50 compensation voltage (CV) more than doubled the mean number of non-redundant proteoforms observed ($1,833 \pm 17$, $n = 3$), compared to without (754 ± 35 proteoforms). We also found FAIMS can influence the transmission of proteoforms and their charge envelopes based on their size. Importantly, FAIMS enabled the identification of intact amyloid beta ($A\beta$) proteoforms, including the aggregation-prone $A\beta_{1-42}$ variant which is strongly linked to Alzheimer's disease.

Introduction

Over the last two decades, the incidence of neurodegenerative diseases has more than doubled worldwide, with Alzheimer's disease (AD) as the most prevalent form.¹ Two protein species, amyloid beta (A β) peptides and phosphorylated microtubule-associated protein tau (tau),²⁻⁵ have been found to be strongly associated with AD. Therefore, detailed proteome characterization of AD has been of importance.^{6, 7} Mass spectrometry has played a central role in these investigations,⁸ particularly bottom-up approaches.⁹⁻¹⁶ However, the protease digestion required for bottom-up analyses impedes the capturing of a protein's complete state, which can vary due to genetic alleles, alternative splicing, proteolytic processing, and post-translational modifications (referred to as "proteoforms").¹⁷⁻²⁰ Since top-down proteomic (TDP) approaches analyze proteins in an intact state, the likelihood of capturing proteoforms associated with certain pathologies is greater and allows for a stronger, more direct connection between genotype and phenotype.²⁰⁻²² For example, TDP is particularly well suited for capturing endogenous proteolytic fragments derived from proteins such as tau, which have been linked to Alzheimer's pathology.²³⁻²⁶ However, TDP of complex samples requires offline fractionation techniques to reduce sample complexity, as the human proteome spans several orders of magnitude in size and abundance, and proteins are generally not well-resolved with reverse-phase high-performance liquid chromatography (RP-HPLC).²⁷ This offline fractionation can be accomplished with gel-eluted liquid fraction entrapment electrophoresis (GELFrEE), size-exclusion chromatography (SEC), ion exchange chromatography, or affinity purification, to name a few.^{22, 28-31} Unfortunately, offline fractionation typically reduces throughput.³² An

attractive online fractionation alternative that can be introduced between the LC and MS dimensions with no impact on throughput is gas-phase ion mobility separation. High field asymmetric waveform ion mobility spectrometry (FAIMS) is particularly well suited to this task with the recent introduction of the FAIMS Pro device, whose modular design allows facile incorporation of ion mobility onto several current Orbitrap instruments.³³ FAIMS, also referred to as differential mobility spectrometry (DMS), separates ions in a carrier gas based on combinations of factors such as size, charge, or shape through the introduction of an asymmetric waveform with high and low electric fields.³⁴ To prevent collision of the ions with the electrode, a deviation in the ion's path is introduced through the application of a DC compensation voltage (CV),³⁵ allowing selective transmission of that ion.³⁴ The application of FAIMS to intact protein analysis has largely been restricted to the separation of conformers of individual proteins or small combinations of proteins,³⁶⁻⁴¹ however, FAIMS has recently been applied to intact and native protein analyses using liquid extraction surface analysis (LESA).⁴²⁻⁴⁶ Here we apply FAIMS-TDP analysis to a whole-tissue sample from the medial frontal cortex (MFC) of an Alzheimer's patient. Scanning across a CV range of -50 to -20 CV with external stepping allowed us to determine how modulation of FAIMS CV influences the characteristics of the ions transmitted through the cylindrical FAIMS unit, and how this relationship can be exploited to target proteoforms based on size. FAIMS-TDP more than doubled proteoform identifications at a single CV compared to without FAIMS and enabled deeper interrogation of proteoforms relevant to neurodegenerative diseases, including α -, β -, and γ -synucleins, PARK7, tau splice isoforms, and several intact A β proteoforms. Taken together, this work describes how greater identifications of

proteoforms and proteome coverage can be achieved reproducibly via gas-phase fractionation with FAIMS in TDP.

Methods

Sample Preparation

Medial frontal cortex (MFC) sample was received from Rush University Alzheimer's Disease Center. The patient's overall cognitive diagnostic category was Alzheimer's dementia with no other causes of cognitive impairment,⁴⁷ while the post-mortem interval was 230 minutes. 33 mg of MFC tissue stored at -80°C was transferred to 1.5 mL LoBind Eppendorf tubes (Eppendorf, Cat# 022431081) and combined with 0.5 mL homogenization buffer consisting of 8 M urea, 10 mM ammonium bicarbonate (ABC) and 10 mM tris(2-carboxyethyl)phosphine (TCEP). Homogenization was performed with a handheld pellet pestle (BioVortexer with SpiralPestle; Biospec 1083 and 1017) before combination with another 0.5 mL homogenization buffer. The sample was then incubated at 37°C for 60 minutes with 1,200 RPM mixing using a ThermoMixer to allow for extraction and denaturation of cellular proteins. Pelleting of urea-insoluble material was accomplished with centrifugation at 18,000 *g* for 20 minutes at 22°C and the resulting supernatant was then added to 3 mL wash buffer (WB; 8 M Urea, 10 mM ABC) within a 4 mL 100 kDa MWCO filter to remove large MW species. After 1 hr of centrifugation at 22°C and 5,000 *g*, the retentate volume was 200 µL and filtrate 3.8 mL. The 3.8 mL filtrate was transferred to a 3 kDa MWCO filter to remove small molecule contaminants and low MW peptides. The 3 kDa MWCO filter was centrifuged at 22°C and 7,300 *g* for 1 hr, giving a retentate volume of 200 µL. The 200 µL retentate from the 100 kDa MWCO filter was washed once more by diluting to 4 mL with WB and filtering it

again through the same filter. This flowthrough was added to the same 3 kDa MWCO filter as above and concentrated one more time. The final 200 μ L retentate from the 3 kDa MWCO filter wash was transferred to a 1.5 mL LoBind Eppendorf tube. To acidify the sample, 10 μ L of 10% formic acid (FA) was added, resulting in 0.5% FA concentration in the final sample. The sample was centrifuged at 18,000g for 15 minutes to remove precipitated material before measuring the protein concentration using BCA assay. To account for any potential buffer contributions to the BCA assay, BSA standards were prepared in 8 M urea, 10 mM ABC, 0.5% FA. Protein concentration by BCA assay was determined to be 1.1 mg/mL, corresponding to a final yield of 0.22 mg total recovered protein in the ~200 μ L sample. Sample was diluted to 0.5 mg/mL with WB containing 0.5% FA and stored at -80°C until LC-MS analysis.

LC-MS/MS Analysis and FAIMS Settings

Samples were analyzed using a Waters NanoACQUITY UPLC system with mobile phases consisting of 0.2% FA in H₂O (Mobile Phase A) and 0.2% FA in ACN (Mobile Phase B). Both trapping-precolumn (100 μ m i.d., 5-cm length) and analytical column (75 μ m i.d., 50-cm length) were slurry-packed with C2 packing material (5 μ m and 3 μ m for trap/analytical respectively, 300 Å, Separation Methods Technology). Samples were loaded into a 5 μ L loop, corresponding to 2.5 μ g of loaded material amount, and injected onto the trapping column with an isocratic flow of 5% B at 3 μ L/min over 20 min. Separation was performed with a 5% to 50% B gradient over 180 min at 300 nL/min. For MS/MS analysis of proteins, the NanoACQUITY system was coupled to a Thermo Scientific™ Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer equipped with the FAIMS Pro interface. Source parameters included

electrospray voltage of 2.2 kV, transfer capillary temperature of 275°C, and ion funnel RF amplitude of 60%. FAIMS was set to standard resolution without supplementary user-controlled carrier gas flow and a dispersion voltage (DV) of -5 kV (equivalent to a dispersion field of -33.3 kV/cm),⁴⁶ while the compensation voltage (CV) was varied depending on the experiment. The Fusion Lumos was set to “intact protein” application mode, which lowers the HCD cell N₂ pressure to 2 mTorr, and data was collected as full profile. MS¹ and MS² data were acquired at a resolution of 120k and 60k, microscans of 3 and 2, across a 500-2,000 *m/z* and 400-2,000 *m/z* range, and with AGC targets of 5E6 and 5E5, respectively. MS¹ and MS² were acquired with a maximum inject time of 400 ms as well. Data dependent settings included selection of top 6 most intense ions, exclusion of ions lower than charge state 5+, inclusion of undetermined charge states, and dynamic exclusion after 1 observation for 30 s. Ions selected for MS² were isolated over a +/- 1.5 *m/z* window and fragmented through collision-induced dissociation (CID) with a collision energy of 35%. Datasets for each FAIMS CV were collected in triplicate. All raw files have been deposited into the MassIVE data repository, and can be accessed via accession MSV000086696, or alternatively with PXD023607 on ProteomeXchange.

Data Analysis

Proteoform identification was performed with TopPIC version 1.3.⁴⁸ Settings for TopPIC included a precursor window of 3 *m/z* (to account for isotopic envelope), mass error tolerance of 15 ppm, a max/min unknown mass shift of 500 Da and a maximum number of allowed unknown modifications of 1. MS² spectra were searched against a database concatenated with entries from *Homo sapiens* Swiss-Prot (20,352), Swiss-

Prot splice variants (22,000), and TrEMBL (54,436), as well as common contaminants. Identified proteoforms were filtered to an FDR of 1% through TopPIC. Downstream data analysis was performed in the R environment for statistical computing and figure generation.⁴⁹ In the downstream analysis, proteoforms from the same gene with the same starting and ending amino acids that were found to be within +/- 5 Da were combined as a single proteoform in order to increase the stringency of our assignments. Although the 5 Da threshold is arbitrary, this approach compensated for incorrectly assigned monoisotopic peaks, ambiguity in unknown mass shifts, and other artifactual deviations. In order to determine the relative standard deviation of proteoforms within replicates, we utilized the 'feature intensity' output from TopPIC. Fragmentation sequence coverage maps and associated spectra were generated using LCMsSpectator.⁵⁰ Graphical abstract and image assets within Figure 1 were created using Biorender.com.

Results and Discussion

Addition of FAIMS Robustly Increases Proteome Coverage

The cortex tissue sample from a patient diagnosed with Alzheimer's was processed following the workflow in **Figure 1**.

Figure 1: Workflow demonstrating sample preparation, LC-MS, FAIMS-TDP, and data analysis with TopPic/R.

We first sought to analyze the performance of FAIMS at various CVs in comparison to without FAIMS (referred to as “No FAIMS”). It is worth remarking that the FAIMS interface produces a modest loss of ion transmission, presumably due to the longer ion path through the source.⁵¹ Therefore, to better represent typical conditions and ensure equitable comparisons, the “No FAIMS” data were collected using the same sample without the FAIMS unit installed. These “No FAIMS” data were collected in

triplicate and analyzed according to the workflow in **Figure 1** with a top 6 DDA MS² analysis.

Figure 2: Bar charts demonstrating several metrics used in the comparison of FAIMS CVs in steps of 5V with “No FAIMS” datasets. **(A)** Mean number of proteoform identifications per replicate (n = 3). Error bars represent standard deviation from the mean. **(B)** Total unique proteoforms found across all three replicates for each FAIMS CV or “No FAIMS”. **(C)** Total unique genes found across all three replicates for each FAIMS CV or “No FAIMS”. **(D)** Total proteome coverage, or non-redundant amino acids covered by every proteoform’s sequence, found across all three replicates for each FAIMS CV or “No FAIMS”.

The “No FAIMS” datasets on average identified 754 ± 35 proteoforms (**Figure 2A** and **Table S1**) and collectively identified 1,073 unique (non-redundant) proteoforms (**Figure 2B**) derived from 293 unique genes (**Figure 2C**), covering 29,359 amino acids across the proteome (**Figure 2D**). The metric “proteome coverage” (**Figure 2D**) is defined as non-redundant amino acids covered by each proteoform’s sequence. Essentially, this

metric accounts for both the length and diversity of the identified proteoforms. We consider this to be a more balanced metric compared to the raw number of proteoforms and genes, as it avoids biases towards smaller proteoforms. Although simple counts of proteoforms or genes are intuitive, these metrics appear to be strongly influenced by short proteolytic fragments.

The FAIMS datasets were collected from -50 to -20 CV scanned in steps of 5 V. Each of the resulting 7 CVs were collected in triplicate. With respect to proteoforms and genes, FAIMS outperformed “No FAIMS” for all the CVs tested within the -50 to -30 CV range (**Figure 2A-C**). For example, the 3 datasets at -50 CV identified $1,833 \pm 17$ proteoforms, an average increase of ~140% per run compared to “No FAIMS”. Collectively, these three datasets at -50 CV identified 2,564 unique proteoforms from 530 unique genes covering a total of 43,437 amino acids, an increase of 95%, 69%, and 69% over the three “No FAIMS” datasets, respectively (**Figure 2B-D** and **Table S1**). Also, although FAIMS at -25 V observes fewer unique proteoforms and genes compared to “No FAIMS”, the proteoforms themselves are much longer on average. Thus, when considering proteoform length, (**Figure 2D**), the -25 V datasets covers nearly as many amino acids from the proteome (28,398) relative to “No FAIMS” (29,359), doing so with only half as many proteoforms (**Figure 2B**).

We also investigated if FAIMS could provide as robust and reproducible an analysis compared to without FAIMS. We assessed quantification reproducibility by calculating the relative standard deviation (RSD, equivalent to coefficient of variation) of each proteoform’s feature intensity, provided the proteoform met the criterion of being observed across all three replicate datasets for each CV setting. Because the FAIMS

and “No FAIMS” datasets were all collected from a single biological sample, variance among the replicates should be exclusively attributed to instrumental factors. The boxplot in **Figure 3** (as well as individual histograms in **Figure S1**) demonstrates how the distributions of RSDs compare between the FAIMS and “No FAIMS” datasets.

Figure 3: Boxplot displaying the distributions of the relative standard deviations (RSDs) determined from the feature intensities of proteoforms found in FAIMS and “No FAIMS” replicates. The median RSD for each condition is written within the upper box, above the line indicating its position. The upper and lower hinges correspond to the first and third quartiles (25th and 75th percentiles), while the upper and lower whiskers represent 1.5 times the interquartile range (IQR). Outliers beyond the whiskers are not plotted.

Based on median RSD, addition of FAIMS resulted in similar or slightly improved quantification quality relative to “No FAIMS”. The improvement in RSDs among lower voltages (-20 to -30 CV) may be related to the observation that fewer proteoforms are transmitted within that range, providing higher quality MS¹ spectra and therefore better estimations of feature intensity. Furthermore, it is worth noting that the RSDs determined from these replicates are similar to previous top-down⁵⁰ and bottom-up^{52, 53} experiments performed with similar instrumentation. Taken together, this data

demonstrates that the addition of FAIMS to TDP robustly increases proteoform identifications, without any sacrifice of quantification quality.

We next investigated the overlap in identifications between the CVs, and which combination of CVs could provide the largest non-redundant set of observations. Using overlap coefficients, defined as the intersection between two sets divided by the smaller of the two sets, we determined the similarity between each dataset. The overall mean overlap coefficient for both proteoform and gene identifications within each FAIMS CV was 0.76. This is comparable to overlap coefficient from “No FAIMS” datasets (0.77), suggesting that the technical replicates show similar reproducibility. A heatmap displaying the FAIMS overlap coefficients of proteoforms and genes is shown in **Figure 4** and **Figure S2**, respectively.

Figure 4: Heatmap generated in R comparing the overlap coefficients of each CV replicate based on proteoforms identified with TopPIC. White lines separate replicates while black lines separate different CVs. Mean overlap coefficients of each CV's replicates are shown in the middle block of the CVs being compared. White text color is used on overlap coefficients ≥ 0.5 and black text color < 0.5 to improve visibility.

Notably, the overlap of gene identifications across the CV space is much greater relative to proteoforms, although it is particularly noticeable when comparing the longest CV distances. For example, the overlap between -50 and -20 CV is 0.64 at the gene level and 0.06 at the proteoform level. This was expected, as each gene can potentially be represented by multiple proteoforms. **Figure 4** demonstrates that while a 5 to 10 V difference in CV produces overlap coefficients mostly above 0.5, distances greater than 15 V produce larger degrees of dissimilarity and are therefore better spaced for capturing sufficiently different sets of proteoforms. With these datasets, we next pursued determining which combinations of CVs are optimal for achieving maximal gene identifications, proteoform identifications, and sequence coverage with constraints

on the number of CVs for each combination. First, we established the optimal combinations for any number of CVs from 1 to 7 using the metrics unique proteoforms, genes, and proteome sequence coverage (**Figure 5**).

Figure 5: Combinatorial analysis of FAIMS CVs with regards to (**A**) number of proteoforms, (**B**) number of genes, and (**C**) sequence coverage. For each combination, the highest number plotted is highlighted in red and labeled with the CVs for that specific combination.

Using combinations limited to 3 CVs as an example, **Figure 5A** and **5B** detail how -35, -40, and -50 V are ideal if one desires to maximize the number of proteoforms or genes. The proteoforms and genes identified in these 3 CVs (9 datasets) cover 84% of the total unique proteoforms and 92% of the total genes over the entire set of 21 datasets. However, when weighing a proteoform's length with sequence coverage as a

metric, the ideal combination is at -30, -40 and -50 CV, covering 88% of the unique amino acids from all 21 FAIMS datasets with only a modest loss of proteoform and gene identifications (4221 proteoforms and 723 genes, **Figure 5C**). Considering these results with the overlap analysis (**Figure 4**), 15 V or greater distances between CVs

provide the least amount of overlap between proteoforms, while 5-10 V separation within the -50 to -30 CV maximizes genes, proteoforms, and proteome sequence coverage.

To further determine the impact of FAIMS on the depth of proteome coverage, we compared the genes identified from our top-down data sets to a typical bottom-up analysis of human brain. Comprehensive bottom-up datasets of human brain tissue allowed us to estimate abundances for 8,528 proteins using weighted spectral counting.⁵⁴⁻⁵⁶ Proteins were binned into 10 different abundance percentiles based on spectral counts compiled from several bottom-up datasets of the human brain. By cross-referencing with our top-down datasets, we were able to determine where the genes, and the proteoforms they are derived from, rank in terms of estimated protein abundance (using dataset-normalized spectral counts) in the human brain (**Figure 6**).

Figure 6: Bottom-up datasets from human brain tissue were used to estimate protein abundance by weighted spectral counting. These proteins were then binned into 10 different abundance percentiles. **(A)** Comparison between the genes found in FAIMS and “No FAIMS” datasets across the 10 different abundance percentiles. **(B)** The number of genes for each FAIMS CV were normalized to the number of genes found in the “No FAIMS” datasets for each abundance percentile.

One immediate and apparent observation is that the higher the abundance percentile, the higher the probability of identification with TDP. About half of the genes identified from the top-down datasets (with or without FAIMS) were contained within the top 80th percentile of abundance (**Figure 6A**). As anticipated, we found that FAIMS could improve identifications in lower-abundant percentiles over “No FAIMS”, increasing the depth of the observable proteome. By determining the ratio of FAIMS to “No FAIMS” across all percentiles, an appreciable increase in identifications below the 40th percentile is apparent within CVs in the -40 to -50 range (**Figure 6B**). However, when

considering the absolute number of genes, the major contribution of FAIMS' advantage is through broadening the spectrum of identifications across all abundance percentiles.

Relationship between FAIMS Compensation Voltage and Transmission of Proteoforms

With FAIMS, we also noted a trend between FAIMS CV and proteoform molecular weight (**Figure 7**).

Figure 7: Boxplot displaying distributions of proteoform masses identified with TopPic against tested FAIMS CVs. Each distribution includes only non-redundant proteoforms found within all three replicates. The median proteoform mass for each condition is written within the upper box, above the line indicating its position. The upper and lower hinges correspond to the first and third quartiles (25th and 75th percentiles), while the upper and lower whiskers represent 1.5 times the interquartile range (IQR). Outliers beyond the whiskers are not plotted.

At -50 CV the median proteoform mass is ~5 kDa and increases to ~15 kDa at -20 CV. Based on these mass distributions (**Figure 7**), CVs less than -50 V appear well suited for top-down or middle-down proteomic experiments while CVs greater than -50 V may best benefit peptidomic or bottom-up experiments. Interestingly, although the majority of spectra identifying a single proteoform were limited to being found within a 10 V range

(3,599 out of 5,165), there were 7 proteoforms observed across the entire CV range from -50 to -20. Presumably, this can be attributed to their high abundance in the sample, particularly in the case of ubiquitin (UBB), myelin basic protein (MBP), and acyl-CoA-binding protein (ACBP). However, it is also likely that the different charge states of these proteoforms adopt several gas-phase conformations depending on their proton-isomerization, impacting their mobilities.³⁸ These 7 proteoforms allowed us to investigate how the charge state envelope is differentially transmitted through the modulation of CV. As a proxy for the charge state distribution at each CV, we used the median charge state based on the PrSMs identified at each CV and tracked how this value changed as a function of CV starting at -50 CV. The MS¹ scans in **Figure S3** demonstrate how the median PrSM charge state tracks the charge state distribution of UBB. Of these 7 proteoforms, 4 followed an inverse relationship, with median observed charge state increasing as CV was decreased (**Figure 8**).

Figure 8: Heatmap of the median PrSM charge states from 7 proteoforms identified across the -50 to -20 CV range, filtered by ≥ 4 PrSMs in order to increase the

confidence in identifications and provide a reasonable estimate of median charge state. The median PrSM charge state is shown within each block, while the deviation of the PrSM charge state is indicated in red (increasing) or blue (decreasing) relative to -50 CV. Proteoforms on the x-axis are written as the gene name, starting amino acid, and ending amino acid.

It's worth noting this trend has generally been observed with peptides and small proteins.^{51, 57-59} Surprisingly, 3 proteoforms showed the opposite relationship where decreasing CV decreased the median charge states observed (**Figure 8**). A cursory examination of the mean precursor mass between the two groups suggests that larger precursors are more likely to favor higher charge states as CV is decreased. To validate these relationships further, we expanded this analysis to include proteoforms that were identified within a more modest, but still wide, 20-30 CV range (n = 256 proteoforms). Here, proteoforms whose median charge shifted greater than 1 charge across the entire CV range were binned into the two different groups depending on the direction of that shift. Those that shifted less than one charge were considered “neutral” with respect to changing CV. Similar to the previous results, larger precursors were significantly more likely to be associated with an inverse relationship between observed charge states and CV (**Table S2**). With decreasing CV, the majority of proteoforms appeared to transmit at lower charge states (n = 177) compared to those that favored higher charge states (n = 39). The remainder (n=40) were considered “neutral” with respect to changes in CV. Proteoforms binned within the neutral group demonstrated an average precursor mass between the inverse and direct groups, once again suggesting the mass of the proteoform is intrinsically linked to this behavior. Other primary sequence-based parameters such as basic/acidic amino acid composition and aliphatic index were not significantly correlated however (**Table S2**). Although the proteins being introduced into

the gas-phase are presumably denatured, factors typically associated with native proteins such as dipole moment or collisional-cross section may have better predictive value towards this phenomenon.⁵⁹⁻⁶¹ Our results demonstrate how CV can be used to filter proteins of different masses, and how a protein's charge state envelope can be differentially transmitted through FAIMS as well.

Utility of Protein Fragments in Top-Down Proteomic Experiments

Surprisingly, a significant number of proteoforms we identified were fragments of larger proteins. We found that only 25% of the unique proteoforms identified covered greater than half of the protein's sequence from which they were derived. This 25% had an average mass of 11.4 kDa, compared to the remaining 75% which had an average mass of 5.7 kDa. Several factors may contribute to this observation, some of which are independent of FAIMS. For example, these fragments themselves may be proteolytic cleavage products produced under normal homeostatic conditions as part of the cellular “degradome”,⁶² or despite the various precautions taken, introduced during the post-mortem interval and sample handling. Central nervous system tissue is a rich source of signaling peptides known as “neuropeptides” that are commonly derived from much larger precursor proteins as well.⁶³ By cross-referencing our proteoform identifications with an established neuropeptide database (NeuroPedia)⁶⁴ we were able to identify several such neuropeptides including vasostatin-1, secretoneurin, cholecystokinin-58 desnonopeptide, neuropeptide Y, and many non-canonical sequence variants derived from known neuropeptide-producing genes. Outside of biological factors, certain instrument parameters can influence the observation of protein fragments as well. For example, low in-source “fragmentation” voltages (between 10-20 V) can typically be

used to remove adducts and desolvate protein ions, while higher voltages can produce source-induced dissociation. However, the susceptibility of amide bonds to dissociation can vary widely, and as the size of a protein increases so does the likelihood of it containing labile amide bonds such as Xaa-Pro.⁶⁵⁻⁶⁸ As *b*- and *y*-ions may be produced inadvertently through this mechanism, even at low source voltages, we decided to eschew applying source voltage to reduce the chance of introducing fragment ions into the instrument. It is also possible proteins are susceptible to fragmentation events as they pass through the electric fields created in FAIMS. However, these fragments would not be expected to have the same mobility as the parent ion, with the caveat that this is likely dependent on the size of the fragment relative to the precursor ion as well.

Finally, it is worth pointing out a few factors that can bias towards the observation of smaller proteoforms (< ~15 kDa). First and foremost is the signal spreading that occurs as the size of a proteoform increases.⁶⁹ This signal spreading can largely be attributed to the charge state envelope and isotopic distributions of each charge state. Specific to our experimental setup was the use of a size 3K MWCO filter as the final filtration step in our sample prep, which in this case was chosen to ensure smaller amyloid beta proteoforms could be efficiently captured if they were present. General instrumental factors that can create bias towards small proteoforms including the tuning of the quadrupole, electrodynamic capture, and collisions with residual gas molecules in the Orbitrap cell that lead to quicker decay of the transient for larger molecules.⁷⁰ Irrespective of these biases or the exact origin of the fragments, they can still provide insight into the proteome. These fragments are particularly useful in the context of neurodegenerative disease where disruptions in proteostasis are commonly linked to

pathology.^{71, 72} Indeed, fragments of tau are often found to be neurotoxic and play a role in the progression of tauopathies such as Alzheimer's disease.²³⁻²⁶ Generally, the fragments we observe are much larger than tryptic peptides and in several cases where there are many fragments for a particular protein, we observe considerable sequence coverage.

Figure 9: Sequence maps of (A) tubulin alpha-1B chain (TUBA1B) and (B) synapsin-1 (SYN1). Each colored segment represents a proteoform with a unique amino acid start and end site. Color scale represents the number of observed PrSMs that can be mapped to the corresponding segment. Unknown mass shifts were not utilized in this analysis.

This is exemplified by examining the fragments providing coverage for the ~50 kDa tubulin alpha-1B chain (TUBA1B) and the ~70 kDa synapsin-1 (SYN1) in **Figure 9A** and **9B**, both large proteins that would otherwise be difficult to observe in their full-length form in TDP of a complex sample.

Identification of Proteoforms with Relevance to Neurodegenerative Disease

The utilization of FAIMS with our top-down analysis enabled identification of a number of unique Swiss-Prot splice variants and TrEMBL entries, with 267 unique splice variants and 96 TrEMBL entries (**Table S3**). Notably, we observe multiple PrSMs unambiguously identifying an alternative ORF isoform (A0A0D9SF30) for neural cell adhesion molecule 1 (NCAM1). We also observed proteoforms derived from genes that have known roles in several neurodegenerative diseases, in particular α -synuclein and PARK7. In both of these cases, the predominant proteoforms contain the full-length sequence. Interestingly, the majority of α -synuclein, β -synuclein, and (to a lesser extent) γ -synuclein PrSMs were found to contain a ~177 Da unknown mass shift near the C-terminus (full length α -synuclein spectrum with unknown modification shown in **Figure S4A**). This mass shift has been previously observed in open database searches and is generally found at Asp or Glu residues.⁷³ A potential explanation that closely matches the average and monoisotopic mass of the modification consists of one oxygen with three iron atoms as well as the loss of seven hydrogen atoms (based on Unimod accession #1971). Comparison of the isotopic peaks of a y_{24}^{2+} fragment ion from α -synuclein (**Figure S4B**) with a simulated spectrum containing the aforementioned elements thought to belong to the unknown modification (**Figure S4C**) demonstrates the similarity between the two isotopic distributions, with many of the isotopic peaks aligning within 2 ppm of each other. It should also be pointed out that the left most isotopic peaks are unique to naturally occurring iron isotopes, and the absence of these peaks is readily apparent when the spectrum is simulated without containing the three iron atoms (**Figure S4D**), strongly suggesting this unknown modification is very likely composed of

the proposed elements. Prior studies have also described the high affinity of α -synuclein for various metal ions, and the region we observe to contain this modification overlaps with residues known to be involved in binding – specifically the $^{119}\text{DPDNEA}^{124}$ motif (**Figure S5**).⁷⁴⁻⁷⁶ Our data also suggests that the C-terminal regions of β -synuclein and γ -synuclein have similar roles in metal binding as multiple spectra with the same unknown mass shift were matched to similar sequences within these proteins. With regards to the mitochondrial protein PARK7, we noted a ~116.0 Da mass shift on its single active site Cys residue, C¹⁰⁶ (**Figure S6**). One potential explanation with a similar delta mass is succinylation, a modification that is attributed to mitochondrial stress and forms due to Michael addition of fumarate onto a Cys thiol group.⁷⁷

We also identified numerous fragments derived from non-canonical splice isoforms of tau protein known to be expressed in the human brain.⁷⁸ Briefly, tau isoforms are defined by the number of N-terminal inserts due to alternative splicing of exon 2 and/or 3 (referred to as 0N, 1N, and 2N), and the number of microtubule-binding repeats due to alternative splicing of exon 10 (referred to as 3R or 4R).⁷⁹ It's also worth noting that endogenous fragments of tau are a common finding in human brain tissue and CSF fluid, and differential fragmentation of tau may play a role in Alzheimer's disease progression.^{80, 81} We observed several fragments that could unambiguously distinguish 0N and 1N tau (**Figure 10A, S7, and S8**). The fact that we observe high spectral counts for 0N (86 PrSMs) and 1N (91 PrSMs), but do not observe any PrSMs for 2N, is aligned with previous quantitative immunoblotting that found 0N and 1N to be the dominant forms and make up ~91% of total tau while 2N comprised only ~9%.⁸² With regards to unambiguous assignment of the microtubule-binding repeats we

observed spectra that could be assigned to both 3R (**Figure 10A** and **S9**) and 4R tau (**Figure 10A** and **S10**) as well.

Figure 10: (A) Schematic representation of selected tau fragments with the number of PrSMs that unambiguously identify 0/1N or 3/4R tau. 0N and 1N fragments are being shown as sharing the same KKVAV residues at their C-terminus, while 3R and 4R proteoforms share VRTPP at their N-terminus. Residues underlined are the KKVAVR hexapeptide sequence in the second-proline rich region of tau which is cleaved at proteolytic cleavage site #1. The 3R and 4R hexapeptide motifs are not shown but are both cleaved at proteolytic cleavage site #2. (B) Sequence coverage map of MAPT 1N3R (Tau-B) showing all proteoforms that could be mapped to this isoform with a unique amino acid start and end site. Red lines at the bottom represent the positions of proteolytic cleavage sites #1 and #2. Color scale represents the number of observed PrSMs that can be mapped to the corresponding segment. Unknown mass shifts were not utilized in this analysis.

Parsimonious inference allowed us to conclude that brain tau is represented mostly by mixture of 1N3R (Tau-B) followed by 0N4R (Tau-D) splice variants, which is in

agreement with previous observations.^{82, 83} While other splice isoforms also had unique fragments, the low number of spectral matches (<4) did not allow us to confidently conclude about their existence.

Surprisingly, N-termini of the predominant R-domain containing fragments are a continuation from the C-termini of the predominant N-domain containing fragments, or in other words they appear to be linked fragments produced from a possible proteolytic cleavage event (referred to as site #1 in **Figure 10A**). Furthermore, the C-termini of the 3R and 4R domain-containing proteoforms share the same cleavage site (referred to as site #2 in **Figure 10A**) even though the sequences themselves are different due to splice variation of exon 10. Even more surprising, all of these cleavage sites lie within three well described hexapeptide motifs.⁸⁴⁻⁹⁰ **Figure 10B** demonstrates all of the unique fragments we identified from our datasets that could be mapped to the 1N3R tau isoform, as well as the location of the cleavage sites. Cleavage site #1 lies within the second proline-rich region of tau and divides the KVAVVR hexapeptide sequence which is known to provide one of the strongest binding sites for microtubules (**Figure 10B**).⁸⁴⁻⁸⁶ Cleavage site #2 lies within the 3R and 4R hexapeptide motifs VQIVYK and VQIINK, which have been demonstrated to drive aggregation of tau (**Figure 10B**).⁸⁷⁻⁹⁰ Interestingly, many of the fragments shown in **Figure 10B** begin or end within proximity of these two cleavage sites. Each of these hexapeptide motifs are known to form β -structure, often in the form of β -hairpins which underlie the aggregation involved in various neurodegenerative diseases.^{84, 88, 91} Intriguingly, these observations highlight the possibility of a shared proteolytic degradation pathway among tau isoforms that is potentially capable of disrupting their aggregation-seeding regions.^{25, 80, 81}

Lastly, we found numerous spectral matches for A β that could only be observed exclusively by using FAIMS. These A β proteoforms, which typically require special fractionation or handling techniques in order to improve recovery due to their hydrophobic and aggregation prone-properties,^{9, 16, 92, 93} were observed intact using FAIMS within the -40 to -50 CV range. This includes the canonical A β ₁₋₄₂ and A β ₁₋₄₀ (Figure 11A and 11B, respectively), as well as several N-terminally truncated forms (specifically A β ₂₋₄₂ and A β ₄₋₄₂, Figure 11C and 11D).

866.4374 m/z precursor with 48.7% sequence coverage (6 PrSMs). (C) A β_{2-42} from 880.6570 m/z precursor with 40.0% sequence coverage (3 PrSMs). (D) A β_{4-42} from 840.6410 m/z precursor with 44.7% sequence coverage (4 PrSMs).

It is worth noting that A β proteoforms with various N- and C-terminal cleavages are notoriously difficult to identify in bottom-up analyses due to the extreme hydrophobicity of the tryptic products, which highlights an advantage of intact protein analysis. We believe the FAIMS-TDP methodology offers a robust way for achieving intact identification of these A β proteoforms, allowing for facile determination of the different combinations of N- and C-terminal cleavage products that exist.

Conclusions

We have described how nano-LC reverse-phase separation of a highly complex sample containing proteins over a wide mass range, can benefit from the implementation of gas-phase fractionation through FAIMS in the context of top-down mass spectrometry. FAIMS was demonstrated to impact the transmission of proteoforms by size and/or charge, reducing MS¹ complexity and allowing greater depth of coverage of the proteome. FAIMS at a single CV (-50) enabled identification of $1,833 \pm 17$ unique proteoforms on average, more than double compared to without FAIMS (754 ± 35). Addition of FAIMS did not result in deterioration of quantification reproducibility and remained near 20% RSD, which is comparable with label-free bottom-up approaches. Decreasing CV of FAIMS was noted to increase the molecular mass of ion species being transmitted to the instrument (median MW of ~5 kDa at -50 CV vs ~15 kDa at -20 CV), and modulation of CV was also observed to differentially

influence the transmission a proteoform's charge state envelope. We also defined optimal combinations of CVs that could produce the largest number of proteoforms/genes (-50, -40, -35 CV) or proteome sequence coverage (-50, -40, and -30 CV). It is worth pointing out that our work only explored a relatively small CV range which we felt best benefited TDP of a complex sample. Based on our results here, future mass analyzer instrumentation compatible with FAIMS that improve on the acquisition of larger proteoforms may benefit from exploring CVs below -20, extending into the positive voltage range.

Our TDP workflow allowed us to identify and characterize unique proteoforms derived from genes with known roles in neurodegenerative diseases. This included determination of the composition of an unknown mass shift present near the iron-binding domain of α -synuclein proteins, which was used to pinpoint the locations of potential iron-binding domains in β - and γ -synuclein as well. PARK7 was also found to be modified with a succinyl group on its active site Cys residue. We were able to unambiguously distinguish tau fragments corresponding to 0N, 1N, 3R, and 4R splice variant isoforms individually, and describe new proteolytic cleavage sites located within or near several aggregation-seeding hexapeptide repeats. Finally, FAIMS enabled identification of several intact A β proteoforms, including the aggregation-prone A β ₁₋₄₂, without the need for complex fractionation or purification techniques. In summary, we believe that addition of FAIMS to discovery top-down proteomics will provide a robust and reproducible method for increasing the proteome available for detection and characterization.

Acknowledgments

The authors would like to thank Tao Liu and Richard Smith for advice and helpful discussions. This work was supported by U01 AG061356 (P.L.D.J.) and R01 AG015819 (D.A.B.). A portion of the research was performed using EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program.

References

1. Nichols, E.; Szeke, C. E. I.; Vollset, S. E.; Abbasi, N.; Abd-Allah, F.; Abdela, J.; Aichour, M. T. E.; Akinyemi, R. O.; Alahdab, F.; Asgedom, S. W.; Awasthi, A.; Barker-Collo, S. L.; Baune, B. T.; Bejot, Y.; Belachew, A. B.; Bennett, D. A.; Bijani, B. A.; Bin Sayeed, M. S.; Brayne, C.; Carpenter, D. O.; Carvalho, F.; Catala-Lopez, F.; Cerin, E.; Choi, J. Y. J.; Dang, A. K.; Degefa, M. G.; Djalalinia, S.; Dubey, M.; Duken, E. E.; Edvardsson, D.; Endres, M.; Eskandarieh, S.; Faro, A.; Farzadfar, F.; Fereshtehnejad, S. M.; Fernandes, E.; Filip, I.; Fischer, F.; Gebre, A. K.; Geremew, D.; Ghasemi-Kasman, M.; Gnedovskaya, E. V.; Gupta, R.; Hachinski, V.; Hagos, T. B.; Hamidi, S.; Hankey, G. J.; Haro, J. M.; Hay, S. I.; Irvani, S. S. N.; Jha, R. P.; Jonas, J. B.; Kalani, R.; Karch, A.; Kasaeian, A.; Khader, Y. S.; Khalil, I. A.; Khan, E. A.; Khanna, T.; Khoja, T. A. M.; Khubchandani, J.; Kisa, A.; Kissimova-Skarbek, K.; Kivimaki, M.; Koyanagi, A. K.; Krohn, K. J.; Logroscino, G.; Lorkowski, S.; Majdan, M.; Malekzadeh, R.; Marz, W.; Massano, J.; Mengistu, G.; Meretoja, A.; Mohammadi, M.; Mohammadi-Khanaposhtani, M.; Mokdad, A. H.; Mondello, S.; Moradi, G.; Nagel, G.; Naghavi, M.; Naik, G.; Nguyen, L. H.; Nguyen, T. H.; Nirayo, Y. L.; Nixon, M. R.; Ofori-Asenso, R.; Ogbo, F. A.; Olagunju, A. T.; Owolabi, M. O.; Panda-Jonas, S.; Passos, V. M. D.; Pereira, D. M.; Pinilla-Monsalve, G. D.; Piradov, M. A.; Pond, C. D.; Poustchi, H.; Qorbani, M.; Radfar, A.; Reiner, R. C.; Robinson, S. R.; Roshandel, G.; Rostami, A.; Russ, T. C.; Sachdev, P. S.; Safari, H.; Safiri, S.; Sahathevan, R.; Salimi, Y.; Satpathy, M.; Sawhney, M.; Saylan, M.; Sepanlou, S. G.; Shafieesabet, A.; Shaikh, M. A.; Sahraian, M. A.; Shigematsu, M.; Shiri, R.; Shiue, I.; Silva, J. P.; Smith, M.; Sobhani, S.; Stein, D. J.; Tabares-Seisdedos, R.; Tovani-Palone, M. R.; Tran, B. X.; Tran, T. T.; Tsegay, A. T.; Ullah, I.; Venketasubramanian, N.; Vlassov, V.; Wang, Y. P.; Weiss, J.; Westerman, R.; Wijeratne, T.; Wyper, G. M. A.; Yano, Y.; Yimer, E. M.; Yonemoto, N.; Yousefifard, M.; Zaidi, Z.; Zare, Z.; Feigin, V. L.; Vos, T.; Murray, C. J. L.; Collaborators, G. D., Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* **2019**, *18* (1), 88-106.

2. Beyreuther, K.; Masters, C. L., Amyloid Precursor Protein (APP) and BZA4 Amyloid in the Etiology of Alzheimer's Disease: Precursor-Product Relationships in the Derangement of Neuronal Function. *Brain Pathology* **1991**, 1 (4), 241-251.
3. Hardy, J.; Higgins, G., Alzheimer's disease: the amyloid cascade hypothesis. *Science* **1992**, 256 (5054), 184-185.
4. Goedert, M., Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Trends in Neurosciences* **1993**, 16 (11), 460-465.
5. Iqbal, K.; Liu, F.; Gong, C.-X., Tau and neurodegenerative disease: the story so far. *Nature Reviews Neurology* **2016**, 12 (1), 15-27.
6. Musiek, E. S.; Holtzman, D. M., Three dimensions of the amyloid hypothesis: time, space and 'wingmen'. *Nature Neuroscience* **2015**, 18 (6), 800-806.
7. Takeda, S., Tau Propagation as a Diagnostic and Therapeutic Target for Dementia: Potentials and Unanswered Questions. *Frontiers in Neuroscience* **2019**, 13 (1274).
8. Gregorich, Z. R.; Ge, Y., Top-down proteomics in health and disease: challenges and opportunities. *Proteomics* **2014**, 14 (10), 1195-1210.
9. Zakharova, N. V.; Bugrova, A. E.; Kononikhin, A. S.; Indeykina, M. I.; Popov, I. A.; Nikolaev, E. N., Mass spectrometry analysis of the diversity of A β peptides: difficulties and future perspectives for AD biomarker discovery. *Expert Review of Proteomics* **2018**, 15 (10), 773-775.
10. Grasso, G., Mass spectrometry is a multifaceted weapon to be used in the battle against Alzheimer's disease: Amyloid beta peptides and beyond. *Mass Spectrometry Reviews* **2019**, 38 (1), 34-48.
11. Andreev, V. P.; Petyuk, V. A.; Brewer, H. M.; Karpievitch, Y. V.; Xie, F.; Clarke, J.; Camp, D.; Smith, R. D.; Lieberman, A. P.; Albin, R. L.; Nawaz, Z.; El Hokayem, J.; Myers, A. J., Label-Free Quantitative LC-MS Proteomics of Alzheimer's Disease and Normally Aged Human Brains. *Journal of Proteome Research* **2012**, 11 (6), 3053-3067.
12. Zhang, Q.; Ma, C.; Gearing, M.; Wang, P. G.; Chin, L.-S.; Li, L., Integrated proteomics and network analysis identifies protein hubs and network alterations in Alzheimer's disease. *Acta Neuropathologica Communications* **2018**, 6 (1), 19.
13. Ping, L.; Duong, D. M.; Yin, L.; Gearing, M.; Lah, J. J.; Levey, A. I.; Seyfried, N. T., Global quantitative analysis of the human brain proteome in Alzheimer's and Parkinson's Disease. *Scientific Data* **2018**, 5 (1), 180036.
14. Johnson, E. C. B.; Dammer, E. B.; Duong, D. M.; Yin, L.; Thambisetty, M.; Troncoso, J. C.; Lah, J. J.; Levey, A. I.; Seyfried, N. T., Deep proteomic network analysis of Alzheimer's disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. *Molecular Neurodegeneration* **2018**, 13 (1), 52.
15. Ovod, V.; Ramsey, K. N.; Mawuenyega, K. G.; Bollinger, J. G.; Hicks, T.; Schneider, T.; Sullivan, M.; Paumier, K.; Holtzman, D. M.; Morris, J. C.; Benzinger, T.; Fagan, A. M.; Patterson, B. W.; Bateman, R. J., Amyloid β concentrations and stable isotope labeling kinetics of human plasma specific to central nervous system amyloidosis. *Alzheimer's & Dementia* **2017**, 13 (8), 841-849.
16. Schindler, S. E.; Bollinger, J. G.; Ovod, V.; Mawuenyega, K. G.; Li, Y.; Gordon, B. A.; Holtzman, D. M.; Morris, J. C.; Benzinger, T. L. S.; Xiong, C.; Fagan,

- A. M.; Bateman, R. J., High-precision plasma β -amyloid 42/40 predicts current and future brain amyloidosis. *Neurology* **2019**, 93 (17), e1647-e1659.
17. Smith, L. M.; Kelleher, N. L.; Linial, M.; Goodlett, D.; Langridge-Smith, P.; Ah Goo, Y.; Safford, G.; Bonilla*, L.; Kruppa, G.; Zubarev, R.; Rontree, J.; Chamot-Rooke, J.; Garavelli, J.; Heck, A.; Loo, J.; Penque, D.; Hornshaw, M.; Hendrickson, C.; Pasa-Tolic, L.; Borchers, C.; Chan, D.; Young*, N.; Agar, J.; Masselon, C.; Gross*, M.; McLafferty, F.; Tsybin, Y.; Ge, Y.; Sanders*, I.; Langridge, J.; Whitelegge*, J.; Marshall, A.; The Consortium for Top Down, P., Proteoform: a single term describing protein complexity. *Nature Methods* **2013**, 10 (3), 186-187.
18. Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.; Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Ogorzalek Loo, R. R.; Lundberg, E.; MacCoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.; Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Van Eyk, J. E.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschläger, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B., How many human proteoforms are there? *Nat Chem Biol* **2018**, 14 (3), 206-214.
19. Smith, L. M.; Kelleher, N. L., Proteoforms as the next proteomics currency. *Science* **2018**, 359 (6380), 1106-1107.
20. Savaryn, J. P.; Catherman, A. D.; Thomas, P. M.; Abecassis, M. M.; Kelleher, N. L., The emergence of top-down proteomics in clinical research. *Genome Medicine* **2013**, 5 (6), 53.
21. Toby, T. K.; Fornelli, L.; Kelleher, N. L., Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annual Review of Analytical Chemistry* **2016**, 9 (1), 499-519.
22. Catherman, A. D.; Skinner, O. S.; Kelleher, N. L., Top Down proteomics: facts and perspectives. *Biochem Biophys Res Commun* **2014**, 445 (4), 683-693.
23. Chessner, A.; Pritchard, S.; Johnson, G., Tau Clearance Mechanisms and Their Possible Role in the Pathogenesis of Alzheimer Disease. *Frontiers in Neurology* **2013**, 4 (122).
24. Derisbourg, M.; Leghay, C.; Chiappetta, G.; Fernandez-Gomez, F.-J.; Laurent, C.; Demeyer, D.; Carrier, S.; Buée-Scherrer, V.; Blum, D.; Vinh, J.; Sergeant, N.; Verdier, Y.; Buée, L.; Hamdane, M., Role of the Tau N-terminal region in microtubule stabilization revealed by new endogenous truncated forms. *Sci Rep* **2015**, 5 (1), 9659.
25. Hanger, Diane P.; Wray, S., Tau cleavage and tau aggregation in neurodegenerative disease. *Biochemical Society Transactions* **2010**, 38 (4), 1016-1020.
26. Quinn, J. P.; Corbett, N. J.; Kellett, K. A. B.; Hooper, N. M., Tau Proteolysis in the Pathogenesis of Tauopathies: Neurotoxic Fragments and Novel Biomarkers. *J Alzheimers Dis* **2018**, 63 (1), 13-33.
27. Shen, Y.; Tolić, N.; Piehowski, P. D.; Shukla, A. K.; Kim, S.; Zhao, R.; Qu, Y.; Robinson, E.; Smith, R. D.; Paša-Tolić, L., High-resolution ultrahigh-pressure long column reversed-phase liquid chromatography for top-down proteomics. *Journal of Chromatography A* **2017**, 1498, 99-110.

28. Doucette, A. A.; Tran, J. C.; Wall, M. J.; Fitzsimmons, S., Intact proteome fractionation strategies compatible with mass spectrometry. *Expert Review of Proteomics* **2011**, 8 (6), 787-800.
29. Ayaz-Guner, S.; Zhang, J.; Li, L.; Walker, J. W.; Ge, Y., In Vivo Phosphorylation Site Mapping in Mouse Cardiac Troponin I by High Resolution Top-Down Electron Capture Dissociation Mass Spectrometry: Ser22/23 Are the Only Sites Basally Phosphorylated. *Biochemistry* **2009**, 48 (34), 8161-8170.
30. Parks, B. A.; Jiang, L.; Thomas, P. M.; Wenger, C. D.; Roth, M. J.; Boyne, M. T.; Burke, P. V.; Kwast, K. E.; Kelleher, N. L., Top-Down Proteomics on a Chromatographic Time Scale Using Linear Ion Trap Fourier Transform Hybrid Mass Spectrometers. *Analytical Chemistry* **2007**, 79 (21), 7984-7991.
31. Sharma, S.; Simpson, D. C.; Tolić, N.; Jaitly, N.; Mayampurath, A. M.; Smith, R. D.; Paša-Tolić, L., Proteomic Profiling of Intact Proteins Using WAX-RPLC 2-D Separations and FTICR Mass Spectrometry. *Journal of Proteome Research* **2007**, 6 (2), 602-610.
32. Smoluch, M.; Mielczarek, P.; Drabik, A.; Silberring, J., 5 - Online and Offline Sample Fractionation. In *Proteomic Profiling and Analytical Chemistry (Second Edition)*, Ciborowski, P.; Silberring, J., Eds. Elsevier: Boston, 2016; pp 63-99.
33. Eiceman, G. A.; Karpas, Z.; Jr, H., *Ion mobility spectrometry, third edition*. 2013; p 1-400.
34. Guevremont, R., High-field asymmetric waveform ion mobility spectrometry: A new tool for mass spectrometry. *Journal of Chromatography A* **2004**, 1058 (1), 3-19.
35. Shvartsburg, A. A.; Noskov, S. Y.; Purves, R. W.; Smith, R. D., Pendular proteins in gases and new avenues for characterization of macromolecules by ion mobility spectrometry. *Proceedings of the National Academy of Sciences* **2009**, 106 (16), 6495-6500.
36. Purves, R. W.; Guevremont, R., Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry–Mass Spectrometry. *Analytical Chemistry* **1999**, 71 (13), 2346-2357.
37. Robinson, E. W.; Leib, R. D.; Williams, E. R., The role of conformation on electron capture dissociation of ubiquitin. *Journal of the American Society for Mass Spectrometry* **2006**, 17 (10), 1470-1479.
38. Shvartsburg, A. A.; Smith, R. D., High-resolution differential ion mobility spectrometry of a protein. *Analytical chemistry* **2013**, 85 (1), 10-13.
39. Shvartsburg, A. A.; Li, F.; Tang, K.; Smith, R. D., Characterizing the Structures and Folding of Free Proteins Using 2-D Gas-Phase Separations: Observation of Multiple Unfolded Conformers. *Analytical Chemistry* **2006**, 78 (10), 3304-3315.
40. Borysik, A. J. H.; Read, P.; Little, D. R.; Bateman, R. H.; Radford, S. E.; Ashcroft, A. E., Separation of β 2-microglobulin conformers by high-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled to electrospray ionisation mass spectrometry. *Rapid Communications in Mass Spectrometry* **2004**, 18 (19), 2229-2234.
41. Melani, R. D.; Srzentić, K.; Gerbasi, V. R.; McGee, J. P.; Huguet, R.; Fornelli, L.; Kelleher, N. L., Direct measurement of light and heavy antibody chains using ion mobility and middle-down mass spectrometry. *MAbs* **2019**, 11 (8), 1351-1357.

42. Griffiths, R. L.; Hughes, J. W.; Abbatiello, S. E.; Belford, M. W.; Styles, I. B.; Cooper, H. J., Comprehensive LESA Mass Spectrometry Imaging of Intact Proteins by Integration of Cylindrical FAIMS. *Analytical Chemistry* **2020**, 92 (4), 2885-2890.
43. Griffiths, R. L.; Dexter, A.; Creese, A. J.; Cooper, H. J., Liquid extraction surface analysis field asymmetric waveform ion mobility spectrometry mass spectrometry for the analysis of dried blood spots. *Analyst* **2015**, 140 (20), 6879-6885.
44. Sarsby, J.; Griffiths, R. L.; Race, A. M.; Bunch, J.; Randall, E. C.; Creese, A. J.; Cooper, H. J., Liquid Extraction Surface Analysis Mass Spectrometry Coupled with Field Asymmetric Waveform Ion Mobility Spectrometry for Analysis of Intact Proteins from Biological Substrates. *Analytical Chemistry* **2015**, 87 (13), 6794-6800.
45. Griffiths, R. L.; Creese, A. J.; Race, A. M.; Bunch, J.; Cooper, H. J., LESA FAIMS Mass Spectrometry for the Spatial Profiling of Proteins from Tissue. *Analytical Chemistry* **2016**, 88 (13), 6758-6766.
46. Hale, O. J.; Illes-Toth, E.; Mize, T. H.; Cooper, H. J., High-Field Asymmetric Waveform Ion Mobility Spectrometry and Native Mass Spectrometry: Analysis of Intact Protein Assemblies and Protein Complexes. *Analytical Chemistry* **2020**, 92 (10), 6811-6816.
47. Schneider, J. A.; Arvanitakis, Z.; Bang, W.; Bennett, D. A., Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology* **2007**, 69 (24), 2197-2204.
48. Kou, Q.; Xun, L.; Liu, X., TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. *Bioinformatics* **2016**, 32 (22), 3495-3497.
49. Team, R. C., R: A Language and Environment for Statistical Computing. 2019.
50. Park, J.; Piehowski, P. D.; Wilkins, C.; Zhou, M.; Mendoza, J.; Fujimoto, G. M.; Gibbons, B. C.; Shaw, J. B.; Shen, Y.; Shukla, A. K.; Moore, R. J.; Liu, T.; Petyuk, V. A.; Tolić, N.; Paša-Tolić, L.; Smith, R. D.; Payne, S. H.; Kim, S., Informed-Proteomics: open-source software package for top-down proteomics. *Nature methods* **2017**, 14 (9), 909-914.
51. Pfammatter, S.; Bonneil, E.; McManus, F. P.; Prasad, S.; Bailey, D. J.; Belford, M.; Dunyach, J.-J.; Thibault, P., A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements. *Molecular & Cellular Proteomics* **2018**, 17 (10), 2051-2067.
52. Piehowski, P. D.; Petyuk, V. A.; Orton, D. J.; Xie, F.; Moore, R. J.; Ramirez-Restrepo, M.; Engel, A.; Lieberman, A. P.; Albin, R. L.; Camp, D. G.; Smith, R. D.; Myers, A. J., Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis. *Journal of proteome research* **2013**, 12 (5), 2128-2137.
53. Bekker-Jensen, D. B.; Martinez-Val, A.; Steigerwald, S.; Rütger, P. L.; Fort, K. L.; Arrey, T. N.; Harder, A.; Makarov, A. A.; Olsen, J. V., A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. *Molecular & Cellular Proteomics* **2020**, mcp.TIR119.001906.
54. Petyuk, V. A.; Chang, R.; Ramirez-Restrepo, M.; Beckmann, N. D.; Henrion, M. Y. R.; Piehowski, P. D.; Zhu, K.; Wang, S.; Clarke, J.; Huentelman, M. J.; Xie, F.; Andreev, V.; Engel, A.; Guettoche, T.; Navarro, L.; De Jager, P.; Schneider, J. A.; Morris, C. M.; McKeith, I. G.; Perry, R. H.; Lovestone, S.; Woltjer, R. L.; Beach, T. G.; Sue, L. I.; Serrano, G. E.; Lieberman, A. P.; Albin, R. L.; Ferrer, I.; Mash, D. C.;

- Hulette, C. M.; Ervin, J. F.; Reiman, E. M.; Hardy, J. A.; Bennett, D. A.; Schadt, E.; Smith, R. D.; Myers, A. J., The human brainome: network analysis identifies HSPA2 as a novel Alzheimer's disease target. *Brain* **2018**, *141* (9), 2721-2739.
55. Johnson, E. C. B.; Dammer, E. B.; Duong, D. M.; Ping, L.; Zhou, M.; Yin, L.; Higginbotham, L. A.; Guajardo, A.; White, B.; Troncoso, J. C.; Thambisetty, M.; Montine, T. J.; Lee, E. B.; Trojanowski, J. Q.; Beach, T. G.; Reiman, E. M.; Haroutunian, V.; Wang, M.; Schadt, E.; Zhang, B.; Dickson, D. W.; Ertekin-Taner, N.; Golde, T. E.; Petyuk, V. A.; De Jager, P. L.; Bennett, D. A.; Wingo, T. S.; Rangaraju, S.; Hajjar, I.; Shulman, J. M.; Lah, J. J.; Levey, A. I.; Seyfried, N. T., Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat Med* **2020**, *26* (5), 769-780.
56. Kim, M. S.; Pinto, S. M.; Getnet, D.; Nirujogi, R. S.; Manda, S. S.; Chaerkady, R.; Madugundu, A. K.; Kelkar, D. S.; Isserlin, R.; Jain, S.; Thomas, J. K.; Muthusamy, B.; Leal-Rojas, P.; Kumar, P.; Sahasrabudhe, N. A.; Balakrishnan, L.; Advani, J.; George, B.; Renuse, S.; Selvan, L. D.; Patil, A. H.; Nanjappa, V.; Radhakrishnan, A.; Prasad, S.; Subbannayya, T.; Raju, R.; Kumar, M.; Sreenivasamurthy, S. K.; Marimuthu, A.; Sathe, G. J.; Chavan, S.; Datta, K. K.; Subbannayya, Y.; Sahu, A.; Yelamanchi, S. D.; Jayaram, S.; Rajagopalan, P.; Sharma, J.; Murthy, K. R.; Syed, N.; Goel, R.; Khan, A. A.; Ahmad, S.; Dey, G.; Mudgal, K.; Chatterjee, A.; Huang, T. C.; Zhong, J.; Wu, X.; Shaw, P. G.; Freed, D.; Zahari, M. S.; Mukherjee, K. K.; Shankar, S.; Mahadevan, A.; Lam, H.; Mitchell, C. J.; Shankar, S. K.; Satishchandra, P.; Schroeder, J. T.; Sirdeshmukh, R.; Maitra, A.; Leach, S. D.; Drake, C. G.; Halushka, M. K.; Prasad, T. S.; Hruban, R. H.; Kerr, C. L.; Bader, G. D.; Iacobuzio-Donahue, C. A.; Gowda, H.; Pandey, A., A draft map of the human proteome. *Nature* **2014**, *509* (7502), 575-81.
57. Schweppe, D. K.; Prasad, S.; Belford, M. W.; Navarrete-Perea, J.; Bailey, D. J.; Huguet, R.; Jedrychowski, M. P.; Rad, R.; McAlister, G.; Abbatiello, S. E.; Woulters, E. R.; Zabrouskov, V.; Dunyach, J.-J.; Paulo, J. A.; Gygi, S. P., Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry. *Analytical Chemistry* **2019**, *91* (6), 4010-4016.
58. Pfammatter, S.; Bonnell, E.; McManus, F. P.; Thibault, P., Gas-Phase Enrichment of Multiply Charged Peptide Ions by Differential Ion Mobility Extend the Comprehensiveness of SUMO Proteome Analyses. *Journal of the American Society for Mass Spectrometry* **2018**, *29* (6), 1111-1124.
59. Shvartsburg, A. A., Ultrahigh-Resolution Differential Ion Mobility Separations of Conformers for Proteins above 10 kDa: Onset of Dipole Alignment? *Analytical Chemistry* **2014**, *86* (21), 10608-10615.
60. Shvartsburg, A. A.; Bryskiewicz, T.; Purves, R. W.; Tang, K.; Guevremont, R.; Smith, R. D., Field Asymmetric Waveform Ion Mobility Spectrometry Studies of Proteins: Dipole Alignment in Ion Mobility Spectrometry? *The Journal of Physical Chemistry B* **2006**, *110* (43), 21966-21980.
61. Shvartsburg, A. A.; Andrzejewski, R.; Entwistle, A.; Giles, R., Ion Mobility Spectrometry of Macromolecules with Dipole Alignment Switchable by Varying the Gas Pressure. *Analytical Chemistry* **2019**, *91* (13), 8176-8183.

62. Pérez-Silva, J. G.; Español, Y.; Velasco, G.; Quesada, V., The Degradome database: expanding roles of mammalian proteases in life and disease. *Nucleic Acids Research* **2015**, *44* (D1), D351-D355.
63. Hökfelt, T.; Broberger, C.; Xu, Z.-Q. D.; Sergeyev, V.; Ubink, R.; Diez, M., Neuropeptides — an overview. *Neuropharmacology* **2000**, *39* (8), 1337-1356.
64. Kim, Y.; Bark, S.; Hook, V.; Bandeira, N., NeuroPedia: neuropeptide database and spectral library. *Bioinformatics* **2011**, *27* (19), 2772-2773.
65. Grewal, R. N.; El Aribi, H.; Harrison, A. G.; Siu, K. W. M.; Hopkinson, A. C., Fragmentation of Protonated Tripeptides: The Proline Effect Revisited. *The Journal of Physical Chemistry B* **2004**, *108* (15), 4899-4908.
66. Bleiholder, C.; Suhai, S.; Harrison, A. G.; Paizs, B., Towards Understanding the Tandem Mass Spectra of Protonated Oligopeptides. 2: The Proline Effect in Collision-Induced Dissociation of Protonated Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe, and Trp). *Journal of The American Society for Mass Spectrometry* **2011**, *22* (6), 1032-1039.
67. Brechi, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H., Cleavage N-Terminal to Proline: Analysis of a Database of Peptide Tandem Mass Spectra. *Analytical Chemistry* **2003**, *75* (9), 1963-1971.
68. Haverland, N. A.; Skinner, O. S.; Fellers, R. T.; Tariq, A. A.; Early, B. P.; LeDuc, R. D.; Fornelli, L.; Compton, P. D.; Kelleher, N. L., Defining Gas-Phase Fragmentation Propensities of Intact Proteins During Native Top-Down Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2017**, *28* (6), 1203-1215.
69. Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; Lantz, C.; Jung, W.; Johnson, K. R.; Koller, A.; Wolff, J. J.; Campuzano, I. D. G.; Auclair, J. R.; Ivanov, A. R.; Whitelegge, J. P.; Paša-Tolić, L.; Chamot-Rooke, J.; Danis, P. O.; Smith, L. M.; Tsybin, Y. O.; Loo, J. A.; Ge, Y.; Kelleher, N. L.; Agar, J. N., Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nature Methods* **2019**, *16* (7), 587-594.
70. Makarov, A.; Denisov, E., Dynamics of Ions of Intact Proteins in the Orbitrap Mass Analyzer. *Journal of the American Society for Mass Spectrometry* **2009**, *20* (8), 1486-1495.
71. Thibaudeau, T. A.; Anderson, R. T.; Smith, D. M., A common mechanism of proteasome impairment by neurodegenerative disease-associated oligomers. *Nature Communications* **2018**, *9* (1), 1097.
72. Kurtishi, A.; Rosen, B.; Patil, K. S.; Alves, G. W.; Møller, S. G., Cellular Proteostasis in Neurodegeneration. *Molecular Neurobiology* **2019**, *56* (5), 3676-3689.
73. Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I., MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nature Methods* **2017**, *14* (5), 513-520.
74. Binolfi, A.; Rasia, R. M.; Bertoncini, C. W.; Ceolin, M.; Zweckstetter, M.; Griesinger, C.; Jovin, T. M.; Fernández, C. O., Interaction of α -Synuclein with Divalent Metal Ions Reveals Key Differences: A Link between Structure, Binding Specificity and Fibrillation Enhancement. *Journal of the American Chemical Society* **2006**, *128* (30), 9893-9901.

75. Lu, Y.; Prudent, M.; Fauvet, B.; Lashuel, H. A.; Girault, H. H., Phosphorylation of α -Synuclein at Y125 and S129 Alters Its Metal Binding Properties: Implications for Understanding the Role of α -Synuclein in the Pathogenesis of Parkinson's Disease and Related Disorders. *ACS Chemical Neuroscience* **2011**, 2 (11), 667-675.
76. Duce, J. A.; Wong, B. X.; Durham, H.; Devedjian, J.-C.; Smith, D. P.; Devos, D., Post translational changes to α -synuclein control iron and dopamine trafficking; a concept for neuron vulnerability in Parkinson's disease. *Molecular Neurodegeneration* **2017**, 12 (1), 45.
77. Blatnik, M.; Thorpe, S. R.; Baynes, J. W., Succination of Proteins by Fumarate. *Annals of the New York Academy of Sciences* **2008**, 1126 (1), 272-275.
78. Goedert, M.; Spillantini, M. G.; Jakes, R.; Rutherford, D.; Crowther, R. A., Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron* **1989**, 3 (4), 519-526.
79. Liu, F.; Gong, C.-X., Tau exon 10 alternative splicing and tauopathies. *Molecular Neurodegeneration* **2008**, 3 (1), 8.
80. Chen, H.-H.; Liu, P.; Auger, P.; Lee, S.-H.; Adolfsson, O.; Rey-Bellet, L.; Lafrance-Vanasse, J.; Friedman, B. A.; Pihlgren, M.; Muhs, A.; Pfeifer, A.; Ernst, J.; Ayalon, G.; Wildsmith, K. R.; Beach, T. G.; van der Brug, M. P., Calpain-mediated tau fragmentation is altered in Alzheimer's disease progression. *Sci Rep* **2018**, 8 (1), 16725.
81. Cicognola, C.; Brinkmalm, G.; Wahlgren, J.; Portelius, E.; Gobom, J.; Cullen, N. C.; Hansson, O.; Parnetti, L.; Constantinescu, R.; Wildsmith, K.; Chen, H.-H.; Beach, T. G.; Lashley, T.; Zetterberg, H.; Blennow, K.; Höglund, K., Novel tau fragments in cerebrospinal fluid: relation to tangle pathology and cognitive decline in Alzheimer's disease. *Acta Neuropathol* **2019**, 137 (2), 279-296.
82. Hong, M.; Zhukareva, V.; Vogelsberg-Ragaglia, V.; Wszolek, Z.; Reed, L.; Miller, B. I.; Geschwind, D. H.; Bird, T. D.; McKeel, D.; Goate, A.; Morris, J. C.; Wilhelmsen, K. C.; Schellenberg, G. D.; Trojanowski, J. Q.; Lee, V. M.-Y., Mutation-Specific Functional Impairments in Distinct Tau Isoforms of Hereditary FTDP-17. *Science* **1998**, 282 (5395), 1914-1917.
83. Goedert, M.; Jakes, R., Expression of separate isoforms of human tau protein: Correlation with the tau pattern in brain and effects on tubulin polymerization. *EMBO Journal* **1990**, 9 (13), 4225-4230.
84. Kadavath, H.; Jaremko, M.; Jaremko, Ł.; Biernat, J.; Mandelkow, E.; Zweckstetter, M., Folding of the Tau Protein on Microtubules. *Angewandte Chemie International Edition* **2015**, 54 (35), 10347-10351.
85. Kadavath, H.; Hofele, R. V.; Biernat, J.; Kumar, S.; Tepper, K.; Urlaub, H.; Mandelkow, E.; Zweckstetter, M., Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. *Proceedings of the National Academy of Sciences* **2015**, 112 (24), 7501-7506.
86. Schwalbe, M.; Kadavath, H.; Biernat, J.; Ozenne, V.; Blackledge, M.; Mandelkow, E.; Zweckstetter, M., Structural Impact of Tau Phosphorylation at Threonine 231. *Structure* **2015**, 23 (8), 1448-1458.
87. Seidler, P. M.; Boyer, D. R.; Rodriguez, J. A.; Sawaya, M. R.; Cascio, D.; Murray, K.; Gonen, T.; Eisenberg, D. S., Structure-based inhibitors of tau aggregation. *Nature Chemistry* **2018**, 10 (2), 170-176.

88. Zhang, W.; Tarutani, A.; Newell, K. L.; Murzin, A. G.; Matsubara, T.; Falcon, B.; Vidal, R.; Garringer, H. J.; Shi, Y.; Ikeuchi, T.; Murayama, S.; Ghetti, B.; Hasegawa, M.; Goedert, M.; Scheres, S. H. W., Novel tau filament fold in corticobasal degeneration. *Nature* **2020**, 580 (7802), 283-287.
89. Ait-Bouziad, N.; Lv, G.; Mahul-Mellier, A.-L.; Xiao, S.; Zorludemir, G.; Eliezer, D.; Walz, T.; Lashuel, H. A., Discovery and characterization of stable and toxic Tau/phospholipid oligomeric complexes. *Nature Communications* **2017**, 8 (1), 1678.
90. von Bergen, M.; Friedhoff, P.; Biernat, J.; Heberle, J.; Mandelkow, E.-M.; Mandelkow, E., Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (³⁰⁶VQIVYK³¹¹) forming β structure. *Proceedings of the National Academy of Sciences* **2000**, 97 (10), 5129-5134.
91. Jeganathan, S.; von Bergen, M.; Brtlich, H.; Steinhoff, H.-J.; Mandelkow, E., Global Hairpin Folding of Tau in Solution. *Biochemistry* **2006**, 45 (7), 2283-2293.
92. Wildburger, N. C.; Esparza, T. J.; LeDuc, R. D.; Fellers, R. T.; Thomas, P. M.; Cairns, N. J.; Kelleher, N. L.; Bateman, R. J.; Brody, D. L., Diversity of Amyloid-beta Proteoforms in the Alzheimer's Disease Brain. *Sci Rep* **2017**, 7 (1), 9520-9520.
93. Schauer, S. P.; Mylott, W. R.; Yuan, M.; Jenkins, R. G.; Rodney Mathews, W.; Honigberg, L. A.; Wildsmith, K. R., Preanalytical approaches to improve recovery of amyloid- β peptides from CSF as measured by immunological or mass spectrometry-based assays. *Alzheimer's Research & Therapy* **2018**, 10 (1), 118.