Mol Cell Proteomics Papers in Press. Published on March 4, 2020 as Manuscript RA119.001889

Quantitative profiling of the human substantia nigra proteome from laser-capture microdissected FFPE tissue

Eva Griesser[1,2], Hannah Wyatt[2], Sara Ten Have[1], Birgit Stierstorfer[2], Martin Lenter[2] and Angus I. Lamond[1]

[1]Centre for Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee, DD1 5EH United Kingdom, [2]Drug Discovery Sciences, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach an der Riss, Germany

Corresponding author: Angus I. Lamond, a.i.lamond@dundee.ac.uk

Running title: Laser-capture proteomics of FFPE human substantia nigra

Abbreviations

LCM	laser-capture microdissection
FFPE	formalin-fixed paraffin-embedded
SP3	single-pot solid-particle-enhanced sample preparation
TMT	tandem mass tag
GO	gene ontology
IHC	immunohistochemistry
H&E	haematoxylin & eosin
SDC	sodium deoxycholate
ABC	ammonium bicarbonate
ACN	acetonitrile
AGC	automatic gain control
HCD	higher-energy collisional dissociation
TCA	tricarboxylic acid

Abstract

Laser-capture microdissection (LCM) allows the visualization and isolation of morphologically distinct subpopulations of cells from heterogeneous tissue specimens. In combination with formalin-fixed and paraffin-embedded (FFPE) tissue it provides a powerful tool for retrospective and clinically relevant studies of tissue proteins in a healthy and diseased context. We first optimized the protocol for efficient LCM analysis of FFPE tissue specimens. The use of SDS containing extraction buffer in combination with the single-pot solid-phase-enhanced sample preparation (SP3) digest method gave the best results regarding protein yield and protein/peptide identifications. Microdissected FFPE human substantia nigra tissue samples (~3,000 cells) were then analysed, using tandem mass tag (TMT) labelling and LC-MS/MS, resulting in the quantification of >5,600 protein groups. Nigral proteins were classified and analysed by abundance, showing an enrichment of extracellular exosome and neuron-specific gene ontology (GO) terms among the higher abundance proteins. Comparison of microdissected samples with intact tissue sections, using a label-free shotgun approach, revealed an enrichment of neuronal cell type markers, such as tyrosine hydroxylase and alpha-synuclein, as well as proteins annotated with neuron-specific GO terms. Overall, this study provides a detailed protocol for laser-capture proteomics using FFPE tissue and demonstrates the efficiency of LCM analysis of distinct cell subpopulations for proteomic analysis using low sample amounts.

Introduction

Formalin-fixed and paraffin-embedded (FFPE) tissue is a gold mine for retrospective clinical studies considering the huge tissue repositories available worldwide. The samples are not contagious, can be stored for years at room temperature and preserve the structure and morphology of the tissue (1). FFPE is routinely used in pathology departments worldwide to archive tissue specimens. The fixation with formalin preserves the tissue via cross-linking of formaldehyde with nucleic acids, polysaccharides and protein side chains (2). However, these cross-links and modifications with formaldehyde result in poor extraction of proteins and RNA/DNA. Therefore, for subsequent biochemical analysis, the cross-links need to be reversed, which can be achieved via heating ("antigen retrieval") (2). Variation in the protocols used by different laboratories, regarding fixation times, formalin concentrations, temperature and pH, influence the quality of tissue specimens and influence the efficiency of protein extraction from FFPE tissue (3).

Besides heating of samples, different extraction buffers have been used to achieve efficient protein extraction. These buffers commonly contain Tris and SDS, but also other detergents, such as Rapigest, chaotropes (e.g. guanidine-HCl) and organic solvents (e.g. trifluoroethanol, acetonitrile) have been used to solubilize proteins from FFPE tissue (1,2,4).

Tissue sections are a heterogeneous mix of diverse cell populations present in different cell numbers. Laser-capture microdissection (LCM) is a powerful technique for visualizing and isolating distinct cell subpopulations from heterogeneous tissue in its morphological context. LCM enables enrichment of cells and analysis of protein expression in specific cell subpopulations. For example, using LCM, neuromelanin granules were recently isolated from substantia nigra tissue (5), while Drummond et al. purified neurons and amyloid plaques from Alzheimer's disease brain

tissue (6,7). However, a challenge for combining LCM with proteomics is the small quantity of isolated cells and correspondingly low levels of proteins, which usually prohibits fractionation of samples and their in-depth analysis. However, multiplexing of samples, via the use of isobaric TMT, can increase total peptide levels and facilitate fractionation (8,9).

The substantia nigra is an area in the midbrain containing brown-pigmented granules ("neuromelanin"), which are part of dopaminergic neurons. The loss of these neurons in substantia nigra pars compacta is a hallmark of Parkinson's disease (10). Therefore, several studies have already analysed the nigral proteome and compared samples from healthy donors with diseased ones (11-13). Additionally, studies isolating neuromelanin from substantia nigra tissue have been performed recently (5,14). Overall up to 1,795 proteins per study were identified. However, these previous studies did not involve sample fractionation to improve the depth of analysis. Additionally, they were performed using frozen tissues, whereas FFPE tissue material offers better tissue morphology allowing a more precise identification and finally isolation of target cells.

In this study, we have developed an optimized protocol to facilitate efficient LCM analysis of FFPE tissue specimens. A thorough investigation of each sample preparation step is provided including the optimization of protein extraction from FFPE tissue by testing different extraction buffers and investigating the influence of immunohistochemical (IHC) and haematoxylin & eosin (H&E) staining on proteins. SDS present in the protein extracts was removed with the SP3 digest method, which was modified to improve protein and peptide recoveries. Thus, acidification of the peptide wash solution could significantly increase the peptide recovery. Protein expression of microdissected samples was compared to intact tissue sections from substantia nigra to evaluate the efficiency of LCM for the purification of small cell populations, which resulted in the enrichment of neuron-specific proteins including tyrosine hydroxylase and alpha-synuclein. The

optimized protocol was used to analyse samples containing as few as ~3,000 cells isolated from the substantia nigra, using FFPE tissue. Replicate samples of 15 healthy donors were analysed in five separate TMT10plex batches giving an overview of the abundance and annotated GO terms of nigral proteins. This study provides a detailed workflow for laser-capture proteomics of samples from FFPE tissue using limited sample amounts and represents the most in-depth proteomic dataset for human substantia nigra samples reported to date.

Experimental Procedures

Tissue samples and laser-capture microdissection

All FFPE human brain samples were purchased from the Netherlands Brain Bank (Amsterdam/The Netherlands) under the regulatory conditions of the Boehringer Ingelheim corporate policy regarding the acquisition and use of human biospecimen. Brains were sectioned at 10 µm thickness using a microtome, collected on polyethylene naphthalate membrane glass slides (Carl Zeiss Microscopy GmbH, Jena, Germany) and air dried over night at room temperature. Prior to LCM, sections on glass slides were deparaffinized using two washes of xylene (3 min each) followed by ethanol (100%, 96%, 70%) for 1 minute each. Laser microdissection and pressure catapulting of ~3,000 substantia nigra cells was performed using a 10 × objective under brightfield optics on a PALM MicroBeam system (Carl Zeiss Microscopy GmbH). The area equivalent to ~3,000 cells was estimated by counting and averaging the cell number of three squares from tissue sections stained with haematoxylin. An amount of ~3,000 cells reflected an area of 1.2 - 3x10^6 μm^2 sample depending. Isolated cells were collected in an adhesive cap (Carl Zeiss Microscopy GmbH) and resuspended in extraction buffer followed by the preparation steps described below.

Protein extraction

Unstained FFPE human substantia nigra sections, which were analysed without LCM (e.g. for comparison of buffers), were sectioned using a microtome (4 µm thick), transferred into 1.5 mL tubes and deparaffinized by incubation with 500 µL heptane for 1 h. After the addition of 100 µL methanol samples were vortexed thoroughly and centrifuged for 2 min (15,000 g). The supernatant was removed and tissue samples were air-dried followed by resuspension in 100 µL SDS (2% SDS in 300 mM Tris-HCl pH 8.0), sodium deoxycholate (SDC; 1% SDC in 300 mM Tris-HCl pH 8.5)

or Rapigest (0.2% Rapigest in 50 mM ammonium bicarbonate (ABC)) extraction buffer. Samples were boiled (25 min, 99 °C, 350 rpm) and sonicated for 20 cycles (Bioruptor® Pico bath sonicator, Diagenode, Belgium; 30 s on, 30 s off) followed by heating for 2 h at 80 °C (500 rpm) and another sonication step (20 cycles). After centrifugation (10 min, 21,300 g) the supernatant was transferred into a new tube. Optionally, the extraction steps were repeated with the remaining tissue. Reversibly oxidized cysteines were reduced with 10 mM DTT (45 min, 50 °C, 1,000 rpm) followed by alkylation of free thiols with 20 mM iodoacetamide (45 min, 22 °C, 1,000 rpm, in the dark). Proteins were quantified using the fluorometric EZQ™ assay (Thermo Fisher Scientific, Bremen, Germany).

Laser-capture microdissected tissue samples were resuspended in 100 µL SDS extraction buffer followed by the preparation steps described above. The extraction steps were performed twice and both protein extracts were combined.

Protein digestion using the SP3 method

Protein extracts in SDS buffer were cleaned and digested with the SP3 method as described previously with modifications (15,16). Briefly, 10 µL of a 20 µg/µL SP3 bead stock (Sera-Mag SpeedBead carboxylate-modified magnetic particles; GE Healthcare Life Sciences, Freiburg, Germany) for LCM samples (or a 1:10 protein:bead ratio for sections) and 500 µL acetonitrile (ACN; final concentration of 70%) were added to 200 µL of protein extract and incubated for 10 min (1000 rpm). Tubes were mounted on a magnetic rack, supernatants were removed and beads were washed twice with 70% ethanol and once with ACN (1 mL each). Beads were resuspended in 80 µL 50 mM ABC and digested overnight with trypsin (1:50 trypsin:protein ratio for sections, 0.5 µg for LCM samples; Pierce™ trypsin protease, #90058, Thermo Scientific; 37 °C, 1,000 rpm).

The next day another portion of trypsin (1:50 trypsin:protein ratio for sections, 0.5 µg for LCM samples) was added and further digested for 4 h. Peptides were cleaned by addition of 9 µL 10% formic acid (final concentration of 1%) and 1750 µL ACN (final concentration of 95%) followed by incubation for 10 min. After spinning down (1,000 g) tubes were mounted on a magnetic rack and beads were washed once with 1.5 mL ACN. Peptides were eluted from the beads with 50 µL 2% DMSO and acidified with 2.6 µL 20 % formic acid (final concentration of 1%) followed by centrifugation (15,000 g). Peptide amounts were quantified using the fluorometric CBQCA assay (Thermo Scientific).

In-solution digestion

Prior to digestion protein extracts in Rapigest buffer were diluted 1:2 with 50 mM ABC. Rapigest and SDC extracts were digested overnight (1:50 trypsin:protein ratio; Pierce™ trypsin protease, #90058, Thermo Scientific; 37 ºC). The next day another portion of trypsin was added (4 h; 1:50 trypsin:protein ratio). followed by acidification to a final concentration of 1% formic acid. After centrifugation (10 min, 21,300 g) samples were desalted by solid-phase extraction using Sep-Pak tC18 cartridges (50 mg; Waters, Eschborn, Germany) according to the manufacturer's instructions. Briefly, cartridges were activated with ACN and equilibrated with 0.1% TFA in water (1 mL each). Samples were loaded, washed five times with 1 mL 0.1% TFA in water and peptides were eluted with 70% ACN/0.1% TFA (1 mL) and dried *in vacuo* in a Concentrator plus (Eppendorf, Hamburg).

TMT labelling

The required volume for 1.5 µg peptides per sample were dried *in vacuo* in a Concentrator plus (Eppendorf, Hamburg) and resuspended in 50 µL 100 mM HEPES pH 8.5. TMT10plex tags (Thermo Scientific) were dissolved in anhydrous ACN and added to the peptide sample in a 1:10 peptide:TMT ratio. Additional anhydrous ACN was added to a final volume of 22 µL. Samples were incubated for 2 h (22 °C, 750 rpm). Labelling efficiency was evaluated by desalting aliquots from 20% of the samples with C18 stage tips followed by LC-MS analysis and database search using TMT as variable modification (see supplemental data for details). In case of insufficient labelling more TMT was added. Unreacted TMT was quenched by incubation with 5 µL 5% hydroxylamine for 30 min. Samples belonging to the same batch were combined and dried *in vacuo*.

High pH reversed phase fractionation

TMT labelled samples were fractionated using off-line high pH reversed phase chromatography. Dried samples were resuspended in 5% formic acid and loaded onto a 4.6 x 250 mm XBridge BEH130 C18 column (3.5 µm, 130 Å; Waters). Samples were separated on a Dionex Ultimate 3000 HPLC system with a flow rate of 1 mL/min. Solvents used were water (A), ACN (B) and 100 mM ammonium formate pH 9 (C). While solvent C was kept constant at 10%, solvent B started at 5% for 3 min, increased to 21.5% in 2 min, 48.8% in 11 min and 90% in 1 min, was kept at 90% for further 5 min followed by returning to starting conditions and re-equilibration for 8 min. Peptides were separated into 48 fractions, which were concatenated into 24 fractions and subsequently dried *in vacuo*. Peptides were redissolved in 5% formic acid and analysed by LC-MS.

LC-MS analysis

Label-free peptides were analysed on a Q-Exactive Plus mass spectrometer coupled to a Dionex RSLCnano HPLC (Thermo Scientific). Samples were loaded onto a 100 µm × 2 cm Acclaim PepMap-C18 trap column (5 µm, 100 Å) for 6 min with 2% ACN/0.1% formic acid and a constant flow of 4 µL/min. Peptides were separated on a 75 µm × 50 cm EASY-Spray C18 column (2 µm, 100 Å; Thermo Scientific) at 40 ºC over a linear gradient from 10% to 35% B in 110 min with a flow rate of 200 nL/min. Solvents used were 0.1% formic acid (A) and 80% ACN/0.1% formic acid (B). The spray was initiated by applying 2.5 kV to the EASY-Spray emitter. The ion transfer capillary temperature was set to 250 °C and the radio frequency of the S-lens to 50%. Data were acquired under the control of Xcalibur software in a data-dependent mode using the top 15 most abundant precursor ions. Full scan MS spectra were acquired in profile mode with a resolution of 70,000 covering a mass range of *m/z* 350-1,400. The automatic gain control (AGC) target was set to $1x10^6$ ions with a maximum fill time of 20 ms. Precursor ions were isolated in the quadrupole with a window of *m/z* 1.5 and fragmented in the higher energy collisional dissociation (HCD) cell with a normalized collision energy of 27. Spectra of fragment ions were acquired with a resolution of 17,500, AGC target of $1x10^5$ ions and a maximum fill time of 60 ms. Only peptides with a charge between 2 and 6 were considered for fragmentation. Dynamic exclusion of already fragmented precursor ions was 40 ms and peptide match was set to "preferred".

TMT labelled samples were analysed on an Orbitrap Fusion Tribrid mass spectrometer coupled to a Dionex RSLCnano HPLC (Thermo Scientific). LC and spray parameters used were the same as described above. Peptides were separated over a linear gradient from 10% to 40% B in 210 min with a flow rate of 200 nL/min. The full scan was acquired in the orbitrap covering the mass range

of *m/z* 350 to 1,400 with a mass resolution of 120,000, an AGC target of $2x10^5$ ions and a maximum injection time of 50 ms. Only precursor ions with charges between 2 and 7 were selected, with an isolation window of *m/z* 1.6 for fragmentation using collision-induced dissociation in the ion trap with 35% collision energy. The ion trap scan rate was set to "rapid". The AGC target was set to $4x10^3$ ions with a maximum injection time of 180 ms and a dynamic exclusion of 80 s. During the MS3 analysis, for more accurate TMT quantification, 5 fragment ions were co-isolated using synchronous precursor selection in a window of *m/z* 2 and further fragmented with a HCD collision energy of 55%. The fragments were then analysed in the orbitrap with a resolution of 50,000. The AGC target was set to $5x10^4$ ions and the maximum injection time was 115 ms. The cycle duration was 2 s.

TMT batch 4 was measured with the following changes of parameters. Peptides were separated over a linear gradient from 10% to 40% B in 150 min with a flow rate of 200 nL/min. The full scan was acquired with an AGC target of $2x10^5$ ions. Precursor ions were isolated with an isolation window of *m/z* 1.2. The ion trap scan rate was set to "turbo". The AGC target was set to $1x10^4$ ions with a maximum injection time of 50 ms and a dynamic exclusion of 60 s. During the MS3 analysis the maximum injection time was set to 86 ms. The cycle duration was 3 s.

Data analysis

Acquired tandem mass spectra were searched against the Uniprot human database (downloaded on 14th of June 2018, 20,349 proteins) using MaxQuant (version 1.6.2.3). Trypsin/P was specified as a cleavage enzyme, allowing up to two missed cleavages and a mass tolerance of 4.5 ppm for precursor ions and 0.5 Da (Ion trap, used for SPS-MS3 scans) or 20 ppm (Orbitrap) for MS2 fragment ions. Carbamidomethylation of cysteine was used as a fixed modification, with oxidation

of methionine, deamidation of asparagine and glutamine and acetylation of the protein N-terminus used as variable modifications. For a separate search formylation on lysine and methylation on lysine and arginine were included as variable modifications. For TMT data the type of search was set to reporter ion MS3 using TMT10plex. The fractions of all five TMT batches were searched together. For label-free quantification the "match between runs" option in MaxQuant was only applied to the search for the comparison of microdissected samples with intact tissue sections. All other searches of label-free data, regarding the optimization of protein extraction, were performed without this option. A false discovery rate of 1% was applied to peptide and protein identifications. Contaminants and reverse hits were excluded from the results prior to further analysis. All protein intensities were median-normalized to address differences in sample loadings. For TMT data an additional normalization step was applied to address the variation between batches. The sum of median-normalized and isotope-corrected reporter intensities of all ten TMT channels was calculated per batch and protein. The average of sums higher than zero from five batches was calculated. A normalization factor for each batch and protein group was calculated by division of the sum through the average value. Only protein groups with minimum two peptides per protein and minimum five reporter intensity values in 50 samples were considered for further analysis. Also, as shown by correlation analysis (Figure S9), replicate 7.1 was excluded as an outlier.

Hierarchical clusters, heat maps, correlation matrices, box plots and volcano plots were created using Instant Clue (version 0.5.3) (17) and Venn diagrams were created in DataShop (version 1.2; https://peptracker.com/ds/). The histogram was generated with Perseus (version 1.6.2.3) (18), while other plots were prepared with Excel. Functional enrichment and pathway analysis was performed with the DAVID tool (19,20).

Experimental Design and Statistical Rationale

To test the three different extraction buffers, four technical replicates were analysed per buffer using the same donor from substantia nigra. For the comparison of microdissected substantia nigra with intact tissue sections 11 replicate samples from intact sections and 13 replicates from microdissected samples were analysed, which were prepared from five different donors. An unpaired, two-tailed Student's t-test (equal variance) was performed to identify proteins with significant higher or lower intensities ($p \leq 0.01$ and $p \leq 0.05$). In total, 46 replicate samples of 15 healthy donors from microdissected substantia nigra were analysed in five separate TMT batches. Additionally, four mixes of samples were analysed in batches 3-5. An overview of the distribution of replicate samples in batches is shown in Tables S1 and S2.

Results

Protein extraction from FFPE tissue

Formalin fixation leads to the cross-linking of formaldehyde with several amino acid side chains, including lysine, arginine, histidine and cysteine. This preserves the tissue structure, but also results in poor protein solubilisation from FFPE tissue. Therefore, heating of FFPE tissue samples is commonly used to reverse cross-linking with formaldehyde (2). Here, three different protein extraction buffers, containing either 2% SDS, 1% SDC or 0.2% Rapigest, were compared to optimise the amount of extracted protein and the number of protein/peptide identifications. A higher protein amount (>30%) was extracted from tissue using the SDS buffer, in comparison with the Rapigest and SDC buffers (Figure 1A). Also, significantly ($p \leq 0.001$) more protein groups (>20%) and unique peptides (>30%), were identified using SDS buffer.

A Venn diagram comparing proteins identified with the different extraction buffers tested reveals that 63% of proteins were identical in all samples (Figure 1B). Additionally, ~11% were present in samples from the SDS and SDC buffer extracts. Nearly 17% of all identified proteins were only present in samples from the SDS extraction buffer. The number of missed cleavages in SDS (22%) and Rapigest (20%) samples was significantly ($p \leq 0.001$) higher than in those using SDC buffer (13%) (Figure 1C). The coefficient of variation from four technical replicates, ranging from 16 to 18%, was similar for all three buffers (Figure 1D).

Overall, the SDS containing buffer showed the best results and was consequently chosen for use in the optimised protocol. Additionally, we tested altering the concentration of SDS to investigate if the extracted protein amount could be further increased. However, comparison of buffers with 2% and 5% SDS, respectively, showed little or no difference in either the resulting peptide

recovery, or in the numbers of protein/peptide identifications (Figure S1A). 85% of the identified proteins were identical in samples from both buffers (Figure S1B).

After initial protein extraction, some residual tissue often remains. Thus, an additional extraction step was performed to determine if this could improve total protein recovery. On average, up to an additional ~15% of the protein amount from the first extract could be obtained with a second extraction. With an equal peptide amount injected, 14% less proteins and 20% less peptides were identified from the second extraction (Figure S2A). Furthermore, most of the identified proteins were already detected in the first extract (Figure S2B). However, the majority of protein intensities were lower compared to those in the first extract (Figure S2C). Thus, a second solubilisation step typically does not extract a significant number of additional new proteins from the tissue, but rather increases the total levels of the previously extracted proteins.

To identify specific cells of interest in a morphological context, staining of the tissue sample is often required. Therefore, the influence of IHC and H&E staining of proteins in FFPE tissues was investigated. Sections were stained with rabbit anti-tyrosine hydroxylase pAb using diaminobenzidine and consequently the impact of IHC on proteins was investigated. As shown in Figure S3A, there was no significant difference in protein and peptide identifications between stained and unstained samples. Also, the fraction of missed cleavages was similar (Figure S3B). The correlation matrix in Figure S3C shows clustering of samples based on the donor, instead of staining. Overall, the data show that immunohistochemical staining with diaminobenzidine has little or no negative impact on the efficiency of extraction and detection of tissue proteins.

In comparison, H&E staining resulted in a negative impact on the obtained peptide amount, compared to nuclear- and unstained tissue with a loss of up to 80% (Figure S4A). However, protein and peptide identifications remained similar, when an equal peptide amount was injected (Figure

S4B&C). The sample preparation was modified to improve the obtained peptide amount and reduction and alkylation were performed after the SP3 protein wash step, followed by tryptic digestion. Using this modified protocol, the peptide recovery from H&E stained tissue was similar to the one from unstained tissue. The correlation matrices in Figure S4D&E confirm these data. Samples prepared with the usual method (reduction/alkylation before SP3 protein clean up) clustered on the staining patterns. When using the modified method (reduction/alkylation after SP3 protein clean up), samples clustered by donor identity.

SP3 digest

As SDS will interfere with protein digestion and MS measurements, samples need to be cleaned to remove SDS before tryptic digestion. This was done using the SP3 digest method. However, when using the original method, significant sample loss occurred. Thus, protein and peptide washing steps were optimized. In the protein wash step, different pH values of the washing solution (from pH 3 to pH 8) were tested. Additionally, the final concentration of ACN was increased to 70%. Twenty percent more proteins could be recovered, using a solution with pH 8 and 70% ACN, compared to the original conditions with pH 3 and 50% ACN (Figure S5A). Furthermore, we tested whether the peptide wash step after digestion can be removed, considering that the digestion solutions ammonium bicarbonate and triethylammonium bicarbonate can easily be evaporated. However, it was shown in label-free MS measurements from samples after digestion that the protein wash was not sufficient to remove SDS completely. SDS contamination present in the sample interfered with the analysis, as shown in Figure S5C. In comparison, a sample prepared with the peptide wash step after digestion showed a homogenous elution of peptides over the whole retention time range (Figure S5D). Therefore, different washing approaches were tested, e.g. the

use of a sonication bath, shaker, pipetting and rotator. However, they did not improve the outcome, showing that the peptide wash after digestion is required to remove any residual SDS present in the sample. Next, therefore, we optimized this step to overcome low peptide recovery seen using the original conditions, i.e., with 95% ACN in the washing solution. Acidification to a final concentration of 1% formic acid, together with 95% ACN, improved the recovery of peptides by 60% (Figure S5B). Overall, with the above modifications to the original protocol, the recoveries of proteins and peptides were improved significantly for both washing steps.

Enrichment of neuron-specific proteins in microdissected samples compared to intact tissue sections

Laser-capture microdissection enables the purification of cell subpopulations from heterogeneous tissue specimens. Therefore, even minor cell subpopulations can be enriched using this technique. Here, we compared microdissected samples from substantia nigra tissue specimens with intact sections to test the efficiency of LCM (Figure S6). For each of five healthy donors, one intact tissue section and 3,000 cells, isolated with LCM, were analysed using a label-free shotgun MS approach.

The volcano plot in Figure 2A shows a high number of proteins with a lower intensity in microdissected samples, compared to sections. However, 182 proteins showed a minimum two-fold increase in protein intensity. Bioinformatic analysis of these proteins compared to the human proteome revealed an enrichment of neuron-specific GO cellular components, including synapse, neuron projection and dendrite (Figure 2B). Conversely, 736 proteins had a minimum 2-fold lower intensity compared to entire sections. These were enriched for the GO terms ribosome and nucleosome, as well as the terms extracellular exosome, membrane and cytosol (Figure 2C). When

using a smaller and brain-specific protein set as background (identified proteins from intact sections and microdissected substantia nigra) the enriched GO cellular components of proteins with lower intensity didn't change. However, proteins with higher intensity in microdissected tissue samples showed only enrichment of two GO terms, which didn't include any neuron-specific ones.

As shown in Figure 2D, neuron-specific proteins were detected with higher intensity in microdissected samples from most of the donors, compared to the corresponding tissue sections. For example, tyrosine hydroxylase, the marker for dopaminergic neurons, which are present in substantia nigra tissue, was enriched ~8-fold in donor 2. Alpha-synuclein (SNCA), a component of Lewy bodies, which are a pathological hallmark of Parkinson's disease, was 2.5-fold higher in donors 3 and 5. Tyrosine-protein phosphatase non-receptor type 5 (PTPN5), a brain-enriched protein involved in the regulation of synaptic plasticity, was significantly enriched in all donors, with a minimum 8-fold higher intensity in microdissected samples. Protein phosphatase 1 regulatory subunit 1B (PPP1R1B), a dopamine- and cAMP-regulated neuronal phosphoprotein, showed an average increase of ~12-fold. Overall, the data show the efficacy of LCM for the enrichment of not only specific cell subpopulations, but also proteins, which are only expressed in these subpopulations and are likely to be suppressed by higher abundance proteins from the surrounding proteome.

Analysis of the proteome of substantia nigra

The optimized protocol was applied to laser-capture microdissected human substantia nigra tissue. In total, 46 replicate samples from 15 healthy donors were analysed in five TMT batches. The distribution of donors and technical replicates across the batches is shown in Tables S1 and S2. As

part of the ongoing optimization process, there were variations in the method used between the separate TMT batches. Specifically, 16 fractions were analysed in batches 1 & 2, resulting in the identification of >4,300 protein groups (minimum two peptides per protein) per batch. Next, the number of fractions was increased to 24, resulting in ~4,800 protein group identifications in batch 3 (Table 1). Shorter gradient and run times were tested in batch 4, which reduced the time required to analyse 24 fractions but also resulted in significantly lower protein identifications. Thus, for batch 5, the previous 4 h MS method was used. In comparison to label-free shotgun analysis of the same samples it was possible to increase the number of quantified protein groups and unique peptides by minimum 220% using 16 fractions and even by minimum 300% when analysing 24 fractions (Figure S8).

Overall, 5,677 protein groups (with minimum two peptides per protein) and 53,475 unique peptides were identified (Table 1, Table S3). Per replicate sample 3,500-4,800 protein groups and 21,000-35,000 unique peptides were quantified. When including formylation and methylation as variable modifications around 200 protein groups and 1,400 unique peptides per batch were identified additionally (Table S6). 4.7% and 3.3% of identified peptides were formylated or methylated.

When analysing multiple TMT batches, the effect of batch variation, as shown in Figure S9A, must be controlled for. The correlation of protein intensities for samples within batches is high, with typical Pearson coefficients above 0.9, while samples from different batches show a lower correlation, with coefficients of ~0.8. Therefore, a normalization strategy was applied to the data, which included median-normalization, addressing differences in sample loadings, as well as batch normalization, addressing the differences between batches. Figure S9B shows that the correlation of normalized protein intensities of samples from different batches increased to >0.9. Sample 7.1

was deemed an outlier and consequently excluded from further analysis. The intensities from technical replicates per donor were averaged. The correlation of protein intensities between donors was high, with Pearson coefficients of >0.95 (Figure S10A). The technical coefficient of variation from 4-7 technical replicates was between 11 and 16% for donors 1 to 6 (Figure S10B). The biological coefficient of variation from 15 different donors was 22% (Figure S10C).

Functional enrichment of the nigral proteome

Nigral proteins were sorted based on their intensity (average of all donors) and divided into quartiles (Figure 3A). Proteins from each quartile were associated with GO cellular components and biological processes, along with Kegg pathways, using DAVID analysis. Regarding GO biological processes proteins participating in RNA splicing, mitochondrial translation and autophagy were enriched in quartiles 3 and 4, while TCA cycle, substantia nigra development and neurotransmitter secretion were enriched in quartile 1, corresponding to the highest abundance proteins (Figure 3B). Proteins from the extracellular exosome and proteins annotated with several neuron-specific GO cellular components, such as myelin sheath, synapse, neuron projection, dendrite, axon and synaptic vesicle, were also enriched in quartile 1 (Figure 3C&D). Lower abundance proteins, from the lysosome and trans-Golgi network, were enriched in quartiles 3 and 4. Annotation of Kegg pathways revealed the enrichment of proteins from the spliceosome and phosphatidylinositol signalling system among the lower abundance proteins, while synaptic vesicle cycle, dopaminergic synapse, TCA cycle and oxidative phosphorylation, were among the enriched pathways in quartile 1 (Figure S11).

Discussion

To better understand disease phenotypes it is necessary to analyse specific cell subpopulations, rather than heterogeneous mixtures of cells, which may be present in tissues in varying numbers. The laser-capture microdissection (LCM) technique can be applied to purify either specific cell populations, or tissue regions, for subsequent proteomic analysis. In this study, we optimized a protocol for the proteomic analysis of laser-capture microdissected FFPE tissue, using limited sample amounts. We tested different buffers to achieve efficient protein extraction from FFPE tissue samples, obtaining the best results with SDS containing buffer. SDS was removed from protein extracts, prior to digestion, using a modified SP3 method that improved protein and peptide recoveries. Also, we compared microdissected samples from human substantia nigra with intact tissue sections to investigate the efficiency of LCM for the isolation of distinct cell subpopulations in a morphological context. This resulted in the enrichment of neuron-specific proteins in microdissected samples. Finally, the optimized protocol was applied to analyse a set of FFPE human substantia nigra samples microdissected from healthy donors. Based on the intensity of nigral proteins identified, this demonstrated the enrichment of specific GO terms and Kegg pathways for different protein abundance classes.

Protein extraction of FFPE tissue is challenging, mainly due to the cross-links and modifications induced by formaldehyde treatment. However, heating of FFPE samples is known to reverse these cross-links, resulting in improved protein extraction. Also, different buffer components, including detergents and organic solvents, can enhance protein solubilisation from FFPE samples. Thus, SDS in combination with Tris (9,21) and Rapigest with ammonium bicarbonate (22-24), are widely used in extraction buffers. Therefore, we tested these two compositions in comparison with a buffer containing SDC and Tris used in-house, which revealed that the SDS containing buffer

provided the best results regarding extracted protein amount and numbers of protein/peptide identifications. This is in agreement with previous studies, which used SDS for efficient extraction from FFPE tissue (8,9,21). SDS concentrations of 2% in extraction buffers for FFPE samples are often used (25-27). However, recently proteins were extracted with a buffer containing 5% SDS (28). Thus, we compared both SDS concentrations, which didn't show any significant difference. The SDC buffer yielded also higher protein/peptide identification numbers compared to the Rapigest buffer, despite lower extracted protein amounts. We note that the concentrations of the detergents in the tested buffers were different as it was our aim to compare buffers from published studies with an already in-house used one. However, a conclusion regarding the extraction efficiency of these detergents cannot be drawn from the here described results. Higher concentrations of Rapigest and SDC might increase extracted protein amounts and identification numbers. In addition, two digestion methods were applied, which might differ in sample loss and digestion efficiency resulting in varying numbers of missed cleavages and identifications. However, similar to the detergents a conclusion regarding a better digestion method cannot be drawn from the presented results.

Tissue proteins were extracted in a buffer containing 300 mM Tris, to support the reversal of cross-links. Usually, concentrations up to 100 mM Tris are used, but as Kawashima et al. recently described, a higher concentration of Tris increases the amount of extracted proteins (29). Tris, which contains a primary amine, might act either as a scavenger of released formaldehyde, or as a transamination catalyst. After protein extraction residual tissue was usually still present. It has been shown previously that additional extraction steps can improve the overall yield (30,31). Lai et al. reported protein amounts of over 100% of the first extract in a second and up to 70% in a third extraction (30). However, they homogenised samples only in the second extraction step.

Additionally, proteins were extracted in ammonium bicarbonate alone resulting in insufficient protein solubilisation. Here, a second extraction increased the protein amount by an average of ~15% of the first extract and was implemented into the protocol for the analysis of microdissected substantia nigra samples. However, considering the amount of time the extraction procedure requires, we suggest this step might best be included only when insufficient quantities of proteins are recovered after the first extraction.

For the reliable isolation of morphologically distinct cell populations, staining of proteins of interest (IHC), or histological staining (H&E), is often necessary. Here, we tested the impact of staining techniques on protein recovery and yield. In agreement with previous studies, we demonstrated that IHC didn't alter the number of protein identifications, compared to unstained tissue samples (7). Similar to recent reports, we showed that H&E staining reduces peptide identifications up to 20%, while protein numbers were similar between samples with different stainings (22). However, we observed ~ 80% loss in the peptides obtained after tryptic digestion. Additionally, the correlation matrix showed clustering of samples based on their staining patterns, rather than donor identity. Therefore, to optimize the protocol, the order of reduction/alkylation and protein clean-up was switched, which increased the peptide quantities recovered after digestion and resulted in samples clustered in the correlation matrix, independent of their staining. This improvement is probably related to the absence of eosin, an acidic dye, which binds to basic amino acid side chains of cytosolic proteins. When present in the buffer, eosin might react with iodoacetamide, considering that eosin-5-iodoacetamide, a thiol-reactive fluorophore, was previously used to derivatize proteins (32,33). This reaction would result in bulky protein modifications, which can interfere with digestion and increase the number of missed cleavages.

SDS is known to inhibit digestion enzymes and interfere with downstream LC-MS analysis and therefore needs to be removed prior to these steps. Therefore, samples prepared in SDS buffer were processed using the SP3 method (15), which includes a protein clean-up step before and a peptide clean-up step after digestion. However, in our hands, using the original protocol, protein and peptide recoveries were lower than reported. We therefore tested modifications to improve the SP3 protocol. Following previous studies, the protein wash solution was changed to conditions with pH 8 and 70% ACN (34,35). To avoid sample losses through unnecessary preparation steps and considering that digestion and TMT labelling can be performed in the same buffer system, we tried to remove the peptide wash step. However, chromatograms of samples prepared without clean-up showed interference, probably caused by the presence of SDS suppressing peptide ionization. Conversely, samples prepared using the peptide wash step didn't show any interference with contaminants. Therefore, the peptide wash step was retained and supplemented by acidification of the washing solution to improve peptide recovery. We believe this improves the efficiency of both the original protocol and subsequent studies using the SP3 method.

The comparison of analysing microdissected samples and intact sections showed a higher number of protein identifications in the intact tissue sections, with the majority having a higher intensity. This contrasts with a recent study, where more proteins were identified in microdissected samples, compared to intact sections from breast cancer tissues (36). Neuron-specific proteins were enriched in microdissected samples, whereas ribosomal and nucleosomal proteins demonstrated lower intensities. The substantia nigra consists of different cell populations, including different neurons (dopaminergic, GABAergic, glutamatergic) and glial cells (astrocytes, oligodendroglia and microglia) (37,38). Reports suggest that glial cells are present in a higher number compared to dopaminergic neurons (39). Here, we microdissected 3,000 cells from areas containing dense

dopaminergic neurons, which are visible by brown-pigmented neuromelanin. Thus, the results demonstrate that, by using LCM, we could enrich the neuronal cells compared to the surrounding tissue, which contains more glial cells.

Enriched GO terms from proteins with lower intensity in microdissected samples didn't differ when using different protein sets (human and brain-specific proteome) as background for the functional enrichment analysis. However, neuron-specific GO terms from proteins with higher intensity were only enriched, when the human proteome was used as background showing that there is no difference between both areas analysed regarding the identity of neuronal proteins. Nevertheless, several neuron-specific proteins were detected with a significantly higher intensity in microdissected samples including tyrosine hydroxylase, a marker for dopaminergic neurons (40). Alpha-synuclein (SNCA) is involved in Parkinson's disease via the accumulation and formation of Lewy bodies, which are a pathological hallmark of Parkinson's disease (41). In the substantia nigra SNCA is expressed in neuronal cell bodies and synapses as well as TH-positive neurons (42)

This study analysed a total of five TMT batches of microdissected samples. However, they were not all measured using the exact same procedure, with differences in batch design regarding the number of fractions, number of technical replicates and LC-MS methods, as described in the results. With this limitation, we decided to combine the results of all five batches, despite their differences in analysis and design, to provide a "proof of principle" of laser-capture proteomics, using small sample amounts from a set of healthy donors. When more than one TMT batch is analysed, the problem of variation between batches occurs, as shown here in the correlation matrix and reported in previous studies (43). Usually, a reference sample is included in all batches, which is then used for normalization (43,44). Here, we applied an alternative normalization strategy that

involved summing up intensities of all ten channels per protein within a batch and using these sums from all batches to calculate a reference value for normalization. Using this strategy, the variation between batches could be reduced. This provides a good alternative for batch normalisation in cases where no common reference samples are available.

In this study we used a peptide:TMT ratio of 1:10. The standard ratio suggested from the manufacturer is 1:8. However, the number of overlabelled peptides (average of 13%, data not shown), which have a TMT modification on Ser, Thr or His, didn't change compared to the commonly used ratio (45). Although the majority of formaldehyde cross-links can be reversed by heating, there are still irreversible modifications with formaldehyde present. It has been shown recently that formylation and methylation on lysine are common modifications in FFPE tissue samples (46-48). However, these modifications are rarely included in the database search in studies using FFPE material. Here, we performed an additional search including formylation and methylation as variable modifications. Consequently, the number of identifications increased by approx. 5%.

Recently, a growing number of proteomic studies have been performed using LCM to analyse cellular subpopulations in a pathological context. For instance, Clair et al. purified lung alveoli from cryo sections with a total area of 4 mm^2 and identified more than 3,446 proteins (minimum 2 peptides per protein) in a label-free approach (49). Shapiro et al. isolated pancreatic acinar cells from FFPE tissue with a total area of 1 mm^2 (~10,000 cells corresponding to 2-3 µg protein) and identified in total 2,995 proteins (24) in a murine model of caerulein-induced pancreatitis, with a label-free shotgun approach. Buczak et al. isolated different tumour sectors and compared tumour with non-tumour tissue of hepatocellular carcinoma from FFPE tissue. 40 mm^2 per region (5-10 µg peptide) were dissected and 5,838 and 5,659 protein groups, were identified per batch in a

TMT-based approach using fractions (9). In the current study we isolated as little as 3,000 cells, corresponding to areas of 1-3 mm^2 and 1-4 µg peptide and identified in total over 5,600 protein groups with a range of 3,500 to 4,800 protein groups per batch. Considering that protein amount and content vary between different tissues and cells, we can conclude that protein numbers from this study are in agreement with, if not higher than, protein detection reported in most of the previous studies. Interestingly, Davis et al. reported recently the analysis of Betz and Purkinje cells isolated from frozen human brain tissue in a label-free shotgun approach using only 150 cells resulting in the identification of 2718 and 2025 protein groups with minimum 2 peptides per protein. In this study microdissected tissue was directly collected in RIPA buffer followed by digestion with the SP3 method (50). Additionally, it was shown that digestion directly in the cap could significantly improve identification numbers for a low amount of cells (100-400) revealing that the transfer of microdissected tissue from the cap into tubes is a critical point, which could lead to significant sample losses when working with small sample amounts. Therefore, collection of microdissected tissue directly in buffer might further improve results.

Several proteomic studies have analysed the substantia nigra, motivated by its importance in the pathology of Parkinson's disease and other neurodegenerative diseases (5,11-14). However, in previous studies, the region has often been macrodissected with a scalpel, which is not as precise as LCM. Also, these studies were often performed with cryo sections, rather than with the FFPE tissue used here. For example, in 2014 Licker et al. analysed the substantia nigra from healthy donors and donors with Parkinson's disease and identified 1,795 proteins (13), while Plum et al. analysed the neuromelanin granules, which were isolated by LCM and identified 1,072 proteins (5). 97% of these proteins were also detected in the present study (Figure 4).

We present an optimized protocol for laser-capture proteomics using small sample amounts of FFPE tissue, which was applied to human substantia nigra samples. We demonstrate a detailed investigation of the protein extraction step by comparing different buffers and testing the influence of IHC and H&E staining and show the efficiency of LCM by comparing microdissected samples with intact sections. The substantia nigra, with all its different cell populations, was used for this method development study, rather than a specific cell population, such as tyrosine hydroxylase positive neurons. In conclusion, this study is, to the best of our knowledge, the most detailed proteomic dataset of human substantia nigra reported so far. We note that in Parkinson's disease the loss of dopaminergic neurons occurs in the substantia nigra, while those in the ventral tegmental area are more resistant to degeneration. It will therefore be important in future to distinguish neurons from both brain regions and analyse their proteomes in the course of Parkinson's disease. The laser-capture proteomic method established here can now be applied to analyse the protein expression of dopaminergic neurons in a pathological context.

Acknowledgments

We would like to thank members of the Lamond Laboratory for insightful discussions.

Data Availability

The raw files used for this analysis have been deposited to the ProteomeXchange Consortium (51) via the PRIDE partner repository (52) with the dataset identifier PXD016563, along with the full MaxQuant (53) output.

Reviewer account details:

Username: reviewer47858@ebi.ac.uk

Password: 2X7OrqPt

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article. EG, HW, BS and ML are full employees of Boehringer Ingelheim Pharma GmbH & Co. KG.

References

1. Magdeldin, S., and Yamamoto, T. (2012) Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. *Proteomics* **12**, 1045-1058
2. Gustafsson, O. J., Arentz, G., and Hoffmann, P. (2015) Proteomic developments in the analysis of formalin-fixed tissue. *Biochim Biophys Acta* **1854**, 559-580
3. Longuespee, R., Fleron, M., Pottier, C., Quesada-Calvo, F., Meuwis, M. A., Baiwir, D., Smargiasso, N., Mazzucchelli, G., De Pauw-Gillet, M. C., Delvenne, P., and De Pauw, E. (2014) Tissue proteomics for the next decade? Towards a molecular dimension in histology. *OMICS* **18**, 539-552
4. Maes, E., Broeckx, V., Mertens, I., Sagaert, X., Prenen, H., Landuyt, B., and Schoofs, L. (2013) Analysis of the formalin-fixed paraffin-embedded tissue proteome: pitfalls, challenges, and future prospectives. *Amino Acids* **45**, 205-218
5. Plum, S., Steinbach, S., Attems, J., Keers, S., Riederer, P., Gerlach, M., May, C., and Marcus, K. (2016) Proteomic characterization of neuromelanin granules isolated from human substantia nigra by laser-microdissection. *Sci Rep* **6**, 37139
6. Drummond, E., Nayak, S., Faustin, A., Pires, G., Hickman, R. A., Askenazi, M., Cohen, M., Haldiman, T., Kim, C., Han, X., Shao, Y., Safar, J. G., Ueberheide, B., and Wisniewski, T. (2017) Proteomic differences in amyloid plaques in rapidly progressive and sporadic Alzheimer's disease. *Acta Neuropathol* **133**, 933-954
7. Drummond, E. S., Nayak, S., Ueberheide, B., and Wisniewski, T. (2015) Proteomic analysis of neurons microdissected from formalin-fixed, paraffin-embedded Alzheimer's disease brain tissue. *Sci Rep* **5**, 15456
8. Hughes, C. S., McConechy, M. K., Cochrane, D. R., Nazeran, T., Karnezis, A. N., Huntsman, D. G., and Morin, G. B. (2016) Quantitative Profiling of Single Formalin Fixed Tumour Sections: proteomics for translational research. *Sci Rep* **6**, 34949
9. Buczak, K., Ori, A., Kirkpatrick, J. M., Holzer, K., Dauch, D., Roessler, S., Endris, V., Lasitschka, F., Parca, L., Schmidt, A., Zender, L., Schirmacher, P., Krijgsveld, J., Singer, S., and Beck, M. (2018) Spatial Tissue Proteomics Quantifies Inter- and Intratumor Heterogeneity in Hepatocellular Carcinoma (HCC). *Mol Cell Proteomics* **17**, 810-825
10. Michel, P. P., Hirsch, E. C., and Hunot, S. (2016) Understanding Dopaminergic Cell Death Pathways in Parkinson Disease. *Neuron* **90**, 675-691
11. Chen, S., Lu, F. F., Seeman, P., and Liu, F. (2012) Quantitative proteomic analysis of human substantia nigra in Alzheimer's disease, Huntington's disease and Multiple sclerosis. *Neurochem Res* **37**, 2805-2813
12. Kitsou, E., Pan, S., Zhang, J., Shi, M., Zabeti, A., Dickson, D. W., Albin, R., Gearing, M., Kashima, D. T., Wang, Y., Beyer, R. P., Zhou, Y., Pan, C., Caudle, W. M., and Zhang, J. (2008) Identification of proteins in human substantia nigra. *Proteomics Clin Appl* **2**, 776-782
13. Licker, V., Turck, N., Kovari, E., Burkhardt, K., Cote, M., Surini-Demiri, M., Lobrinus, J. A., Sanchez, J. C., and Burkhard, P. R. (2014) Proteomic analysis of human substantia nigra identifies novel candidates involved in Parkinson's disease pathogenesis. *Proteomics* **14**, 784-794
14. Tribl, F., Gerlach, M., Marcus, K., Asan, E., Tatschner, T., Arzberger, T., Meyer, H. E., Bringmann, G., and Riederer, P. (2005) "Subcellular proteomics" of neuromelanin granules isolated from the human brain. *Mol Cell Proteomics* **4**, 945-957
15. Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., and Krijgsveld, J. (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* **10**, 757
16. Hughes, C. S., Moggridge, S., Muller, T., Sorensen, P. H., Morin, G. B., and Krijgsveld, J. (2019) Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68-85
17. Nolte, H., MacVicar, T. D., Tellkamp, F., and Kruger, M. (2018) Instant Clue: A Software Suite for Interactive Data Visualization and Analysis. *Sci Rep* **8**, 12648

18. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740
19. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57
20. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13
21. Wisniewski, J. R., Dus, K., and Mann, M. (2013) Proteomic workflow for analysis of archival formalin-fixed and paraffin-embedded clinical samples to a depth of 10 000 proteins. *Proteomics Clin Appl* **7**, 225-233
22. Foll, M. C., Fahrner, M., Oria, V. O., Kuhs, M., Biniossek, M. L., Werner, M., Bronsert, P., and Schilling, O. (2018) Reproducible proteomics sample preparation for single FFPE tissue slices using acid-labile surfactant and direct trypsinization. *Clin Proteomics* **15**, 11
23. Kennedy, J. J., Whiteaker, J. R., Schoenherr, R. M., Yan, P., Allison, K., Shipley, M., Lerch, M., Hoofnagle, A. N., Baird, G. S., and Paulovich, A. G. (2016) Optimized Protocol for Quantitative Multiple Reaction Monitoring-Based Proteomic Analysis of Formalin-Fixed, Paraffin-Embedded Tissues. *J Proteome Res* **15**, 2717-2728
24. Shapiro, J. P., Komar, H. M., Hancioglu, B., Yu, L., Jin, M., Ogata, Y., Hart, P. A., Cruz-Monserrate, Z., Lesinski, G. B., and Conwell, D. L. (2017) Laser Capture Microdissection of Pancreatic Acinar Cells to Identify Proteomic Alterations in a Murine Model of Caerulein-Induced Pancreatitis. *Clin Transl Gastroenterol* **8**, e89
25. Tanca, A., Abbondio, M., Pisanu, S., Pagnozzi, D., Uzzau, S., and Addis, M. F. (2014) Critical comparison of sample preparation strategies for shotgun proteomic analysis of formalin-fixed, paraffin-embedded samples: insights from liver tissue. *Clin Proteomics* **11**, 28
26. Tanca, A., Pagnozzi, D., Falchi, G., Biosa, G., Rocca, S., Foddai, G., Uzzau, S., and Addis, M. F. (2011) Impact of fixation time on GeLC-MS/MS proteomic profiling of formalin-fixed, paraffin-embedded tissues. *J Proteomics* **74**, 1015-1021
27. Jiang, X., Jiang, X., Feng, S., Tian, R., Ye, M., and Zou, H. (2007) Development of efficient protein extraction methods for shotgun proteome analysis of formalin-fixed tissues. *J Proteome Res* **6**, 1038-1047
28. Marchione, D. M., Ilieva, I., Garcia, B. A., Pappin, D. J., Wilson, J. P., and Wojcik, J. B. (2019) HYPER-sol: flash-frozen results from archival FFPE tissue for clinical proteomics. *bioRxiv*, 632315
29. Kawashima, Y., Kodera, Y., Singh, A., Matsumoto, M., and Matsumoto, H. (2014) Efficient extraction of proteins from formalin-fixed paraffin-embedded tissues requires higher concentration of tris(hydroxymethyl)aminomethane. *Clin Proteomics* **11**, 4
30. Lai, X., and Schneider, B. P. (2014) Integrated and convenient procedure for protein extraction from formalin-fixed, paraffin-embedded tissues for LC-MS/MS analysis. *Proteomics* **14**, 2623-2627
31. Shi, S. R., Taylor, C. R., Fowler, C. B., and Mason, J. T. (2013) Complete solubilization of formalin-fixed, paraffin-embedded tissue may improve proteomic studies. *Proteomics Clin Appl* **7**, 264-272
32. Haase, H., and Maret, W. (2008) Partial oxidation and oxidative polymerization of metallothionein. *Electrophoresis* **29**, 4169-4176
33. Brown, L. J., Klonis, N., Sawyer, W. H., Fajer, P. G., and Hambly, B. D. (2001) Independent movement of the regulatory and catalytic domains of myosin heads revealed by phosphorescence anisotropy. *Biochemistry* **40**, 8283-8291
34. Sielaff, M., Kuharev, J., Bohn, T., Hahlbrock, J., Bopp, T., Tenzer, S., and Distler, U. (2017) Evaluation of FASP, SP3, and iST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *J Proteome Res* **16**, 4060-4072

35. Donadio, E., Giusti, L., Cetani, F., Da Valle, Y., Ciregia, F., Giannaccini, G., Pardi, E., Saponaro, F., Torregrossa, L., Basolo, F., Marcocci, C., and Lucacchini, A. (2011) Evaluation of formalin-fixed paraffin-embedded tissues in the proteomic analysis of parathyroid glands. *Proteome Sci* **9**, 29
36. De Marchi, T., Braakman, R. B., Stingl, C., van Duijn, M. M., Smid, M., Foekens, J. A., Luider, T. M., Martens, J. W., and Umar, A. (2016) The advantage of laser-capture microdissection over whole tissue analysis in proteomic profiling studies. *Proteomics* **16**, 1474-1485
37. Abe, M., Kimoto, H., Eto, R., Sasaki, T., Kato, H., Kasahara, J., and Araki, T. (2010) Postnatal development of neurons, interneurons and glial cells in the substantia nigra of mice. *Cell Mol Neurobiol* **30**, 917-928
38. Yamaguchi, T., Wang, H. L., and Morales, M. (2013) Glutamate neurons in the substantia nigra compacta and retrorubral field. *Eur J Neurosci* **38**, 3602-3610
39. Garcia-Amado, M., and Prensa, L. (2012) Stereological analysis of neuron, glial and endothelial cell numbers in the human amygdaloid complex. *PLoS One* **7**, e38692
40. Weihe, E., Depboylu, C., Schutz, B., Schafer, M. K., and Eiden, L. E. (2006) Three types of tyrosine hydroxylase-positive CNS neurons distinguished by dopa decarboxylase and VMAT2 co-expression. *Cell Mol Neurobiol* **26**, 659-678
41. Venda, L. L., Cragg, S. J., Buchman, V. L., and Wade-Martins, R. (2010) alpha-Synuclein and dopamine at the crossroads of Parkinson's disease. *Trends Neurosci* **33**, 559-568
42. Taguchi, K., Watanabe, Y., Tsujimura, A., and Tanaka, M. (2016) Brain region-dependent differential expression of alpha-synuclein. *J Comp Neurol* **524**, 1236-1258
43. Brenes, A., Hukelmann, J. L., Bensaddek, D., and Lamond, A. I. (2019) Multi-batch TMT reveals false positives, batch effects and missing values. *Mol Cell Proteomics*
44. Lapek, J. D., Jr., Greninger, P., Morris, R., Amzallag, A., Pruteanu-Malinici, I., Benes, C. H., and Haas, W. (2017) Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. *Nat Biotechnol* **35**, 983-989
45. Zecha, J., Satpathy, S., Kanashova, T., Avanessian, S. C., Kane, M. H., Clauser, K. R., Mertins, P., Carr, S. A., and Kuster, B. (2019) TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution Labeling Approach. *Mol Cell Proteomics* **18**, 1468-1478
46. Zhang, Y., Muller, M., Xu, B., Yoshida, Y., Horlacher, O., Nikitin, F., Garessus, S., Magdeldin, S., Kinoshita, N., Fujinaka, H., Yaoita, E., Hasegawa, M., Lisacek, F., and Yamamoto, T. (2015) Unrestricted modification search reveals lysine methylation as major modification induced by tissue formalin fixation and paraffin embedding. *Proteomics* **15**, 2568-2579
47. Noberini, R., Uggetti, A., Pruneri, G., Minucci, S., and Bonaldi, T. (2016) Pathology Tissue-quantitative Mass Spectrometry Analysis to Profile Histone Post-translational Modification Patterns in Patient Samples. *Mol Cell Proteomics* **15**, 866-877
48. Bennike, T. B., Kastaniegaard, K., Padurariu, S., Gaihede, M., Birkelund, S., Andersen, V., and Stensballe, A. (2016) Comparing the proteome of snap frozen, RNAlater preserved, and formalin-fixed paraffin-embedded human tissue samples. *EuPA Open Proteom* **10**, 9-18
49. Clair, G., Piehowski, P. D., Nicola, T., Kitzmiller, J. A., Huang, E. L., Zink, E. M., Sontag, R. L., Orton, D. J., Moore, R. J., Carson, J. P., Smith, R. D., Whitsett, J. A., Corley, R. A., Ambalavanan, N., and Ansong, C. (2016) Spatially-Resolved Proteomics: Rapid Quantitative Analysis of Laser Capture Microdissected Alveolar Tissue Samples. *Sci Rep* **6**, 39223
50. Davis, S., Scott, C., Ansorge, O., and Fischer, R. (2019) Development of a Sensitive, Scalable Method for Spatial, Cell-Type-Resolved Proteomics of the Human Brain. *J Proteome Res* **18**, 1787-1795
51. Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H. J., Albar, J. P., Martinez-Bartolome, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* **32**, 223-226

52. Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q. W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **44**, D447-456
53. Tyanova, S., Temu, T., and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319

Tables

Table 1: Overview of protein and peptide identifications in all analysed TMT batches. *Minimum 2 peptides per protein.

Batch	Fractions	# Protein groups (2*)	# Unique peptides
1	16	4,343	29,887
2	16	4,424	31,941
3	24	4,812	35,725
4	24	3,514	21,260
5	24	3,606	21,279
Total	104	5,677	53,475

Figures

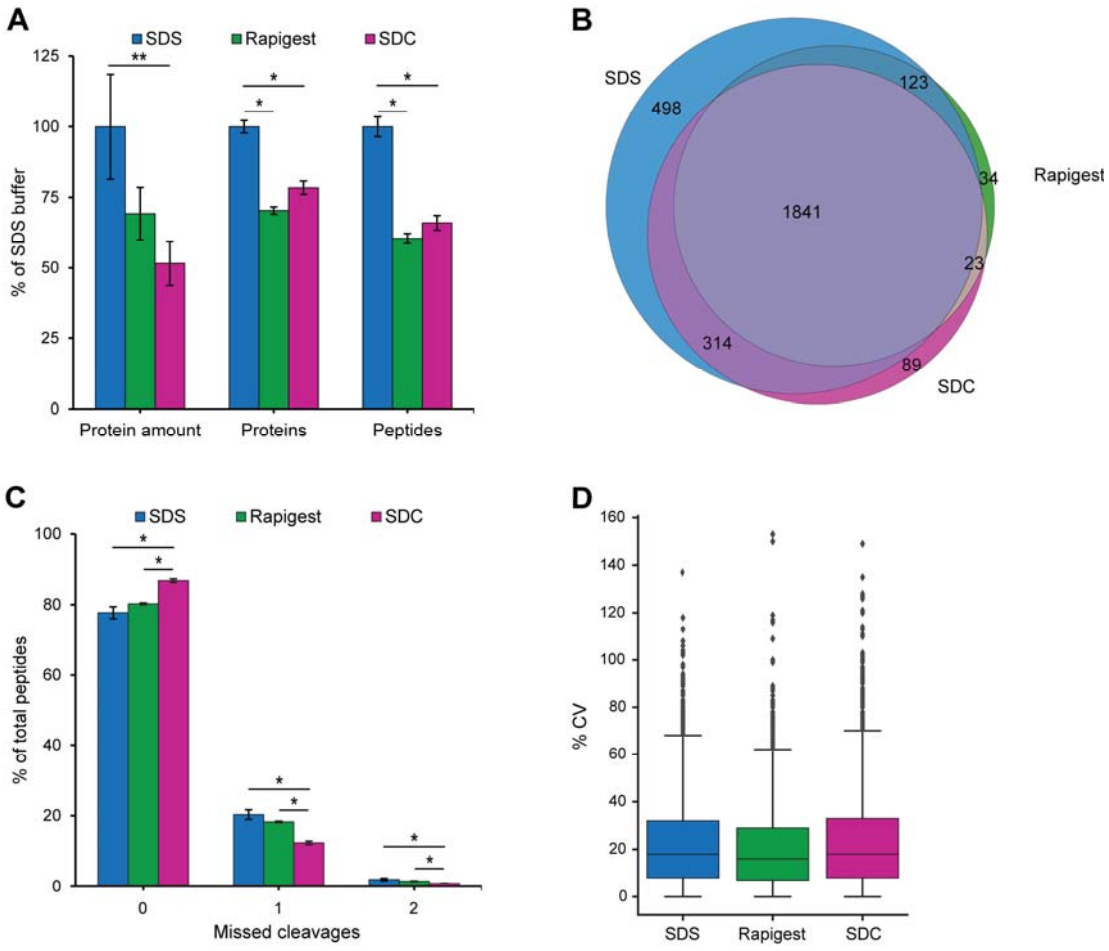

Figure 1: Optimization of protein extraction from FFPE tissue samples. (A) Protein amount extracted from FFPE human substantia nigra (1 section, 4 µm) and number of protein groups (minimum 2 peptides per protein) and unique peptides identified using different protein extraction buffers. Data is shown as means ± SD (n=4, * p≤0.001, ** p≤0.02) and as % of SDS buffer (100% = 45 µg protein, 2,370 protein groups, 18,560 peptides). (B) Venn diagram representing the number of identified protein groups, which are unique and common in samples from different extraction buffers. (C) Fraction of peptides with zero, one and two missed cleavages. Data is shown as means ± SD (n=4, * p≤0.001). (D) Box Plot showing the coefficient of variation (CV) of median-normalized protein intensities from technical replicates (n=4).

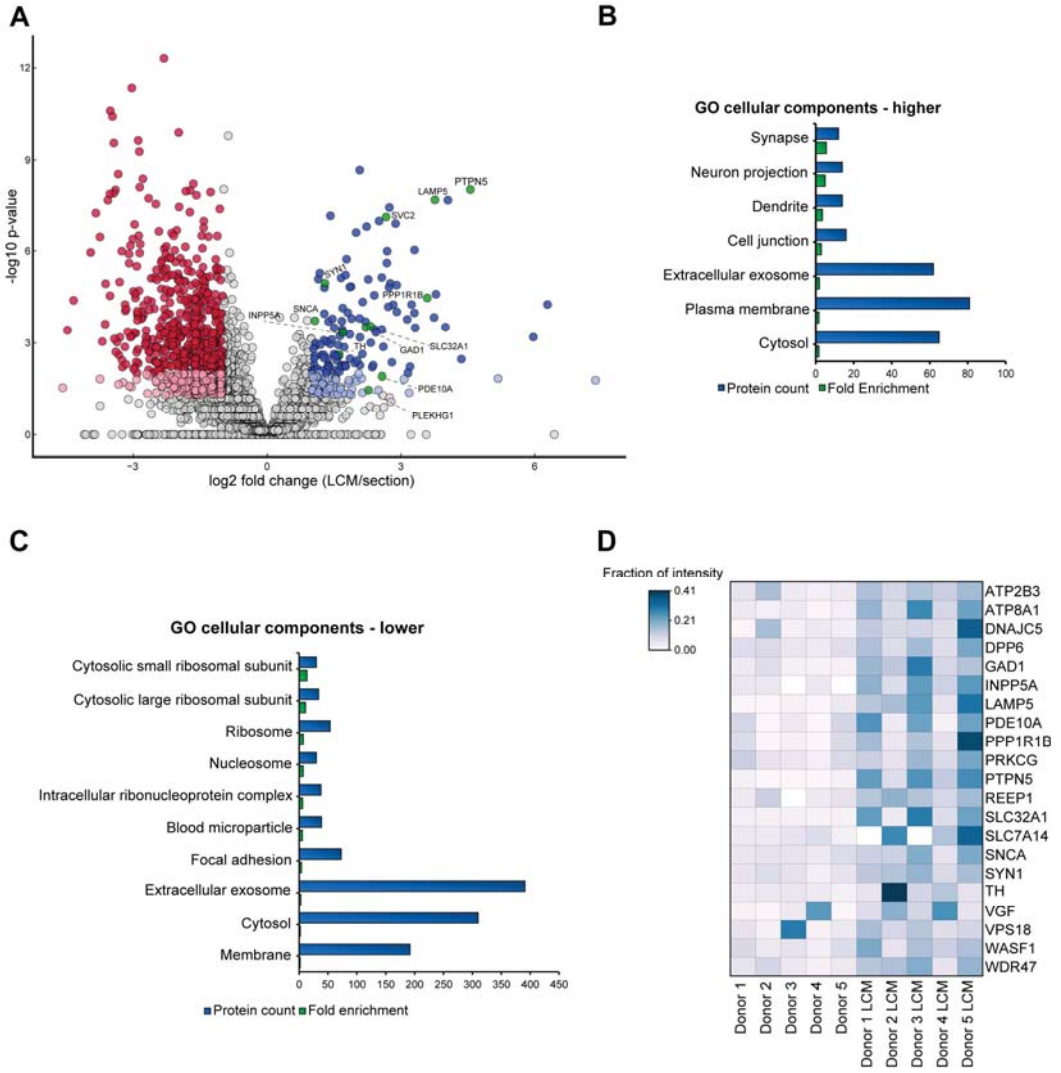

Figure 2: Comparison of microdissected tissue samples with intact sections. (A) Volcano Plot representing the log2 fold changes in protein intensities vs. the negative log10-transformed p-values (two-tailed, equal variance t-test) of microdissected substantia nigra (~3,000 cells) vs. intact tissue sections (10 µm). Data is based on 11 replicate samples from intact sections and 13 replicates from microdissected samples, which were prepared from total 5 biological replicates. Proteins with fold changes ≥2 (blue) or ≤2 (red) and p-values ≤0.01 (dark) or ≤0.05 (light) are displayed in colour. Neuron-specific proteins are presented in green. (B) GO cellular components enriched in microdissected substantia nigra compared to intact sections (Benjamini-corrected p-value <0.01, protein count ≥10). The human proteome was used as

background. (C) GO cellular components with lower intensity in microdissected substantia nigra compared to intact sections (Benjamini- corrected p-value <0.01, protein count ≥10). The human proteome was used as background. (D) Heat map representing the fraction of intensity of microdissected samples and intact sections for neuron-specific proteins.

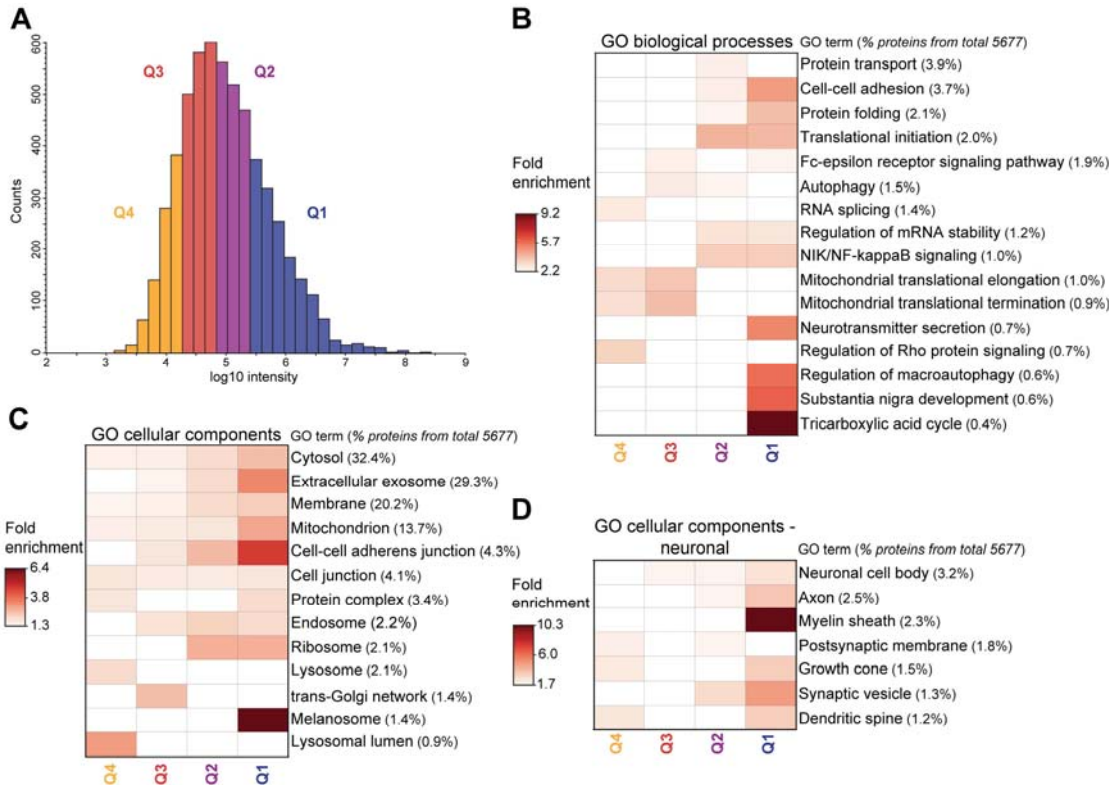

Figure 3: Functional enrichment of the nigral proteome (A) Histogram representing the intensity range of the nigral proteins identified. Depending on their intensity proteins were divided into four quartiles. (B) Heat map showing GO biological processes, which were enriched in each of the quartiles (Benjamini-corrected p-value <0.05, protein count ≥1%). (C) Heat map showing GO cellular components enriched in each of the quartiles (Benjamini-corrected p-value <0.05, protein count ≥1%). (D) Heat map showing neuronal GO cellular components enriched in each of the quartiles (Benjamini-corrected p-value <0.05, protein count ≥1%).

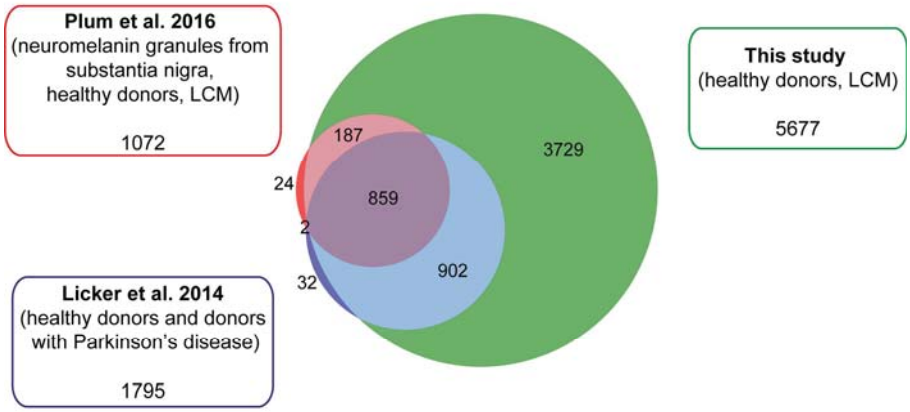

Figure 4: Comparison of protein identifications between different studies. Venn diagram showing the number of identified proteins from published studies compared to the present one.