VetCore - Facility for Research

(Head: Ao. Univ.-Prof. Dr.med.vet. Dieter Klein)

Optimising and Establishing Single-Pot, Solid-Phase-Enhanced Sample Preparation (SP3) as a Standard Digestion Protocol

Bachelor Thesis

University of Veterinary Medicine, Vienna

Adam Iebed

Vienna, July 2021

Supervisor: Ao. Univ.-Prof. Dipl.-Ing. Dr. Ebrahim Razzazi-Fazeli

Reviewer: Dr in rer.nat. Corina Itze-Mayrhofer

Acknowledgements

I would like to express my appreciation to Ao. Univ.-Prof. Dipl.-Ing. Dr. Ebrahim Razzazi-Fazeli for giving me the opportunity to work on my research project and his support and trust.

A very special thanks goes to Dr. Karin Hummel and Dr. Katharina Nöbauer, who not only were always welcoming and patient but also gave me all the support I needed throughout the working process.

I also want to thank the whole VetCore-Facility of the University of Veterinary Medicine Vienna for their motivating work environment.

Further thanks go to Sophie Loidolt, BSc, who shared her recent experiences with me and always found the right words when I needed them.

Last but not least, I am grateful for the patience and the emotional support from my family and friends along my way.

Table of Contents

1. Introduction

1.1. Proteomics

The proteome can be defined as the whole protein repertoire of any organism that can be or is expressed. The research field of proteomics deals with the proteome, which includes the structure, expression, any kind of function, modification and interaction of proteins at any time and stage (Aslam *et al.*, 2017). Because proteins play an essential role in any given mechanism in all organisms, the goal is to understand the pathways and processes by determining the different properties (Patterson and Aebersold, 2003). The human genome is estimated to have around 40,000 genes, with alternative RNA splicing and post-translational modification (PTM) potentially encoding up to 2,000,000 proteins (Cho, 2007). Therefore, genomics does provide lots of information about sequences and activity, but it can't predict protein-protein interactions and PTMs, which most proteins undergo (Cho, 2007). Especially after the completion of the "Human Genome Project", focus shifted to the proteome, not with the intention to identify proteins, but to understand and determine their physiological and pathological functions (Cho, 2007; Hughes *et al.*, 2014) .

Because proteins vary so much from cell to cell, also depending on the current state and external influences, there are different types of proteomic strategies (Aslam *et al.*, 2017). Those different approaches are top-down, middle-down, bottom-up and shotgun strategy (Zhang *et al.*, 2013). Bottom-up is performed by analysis of peptides from the protein after proteolysis. Whenever peptides of a complex mixture are analysed, it is called shotgun analysis. Top-down uses the strategy of characterising intact proteins, which comes in handy when determining PTMs and protein isoforms. However, there are also downsides and limitations because of difficulties with protein fragmentation when protein ionisation is performed in gas phase (Zhang *et al.*, 2013). To overcome these limitations, a combination of bottom-up and top-down analysis has been introduced and called middle-down. It is working with partially digested proteins, so the fragments are larger than in bottom-up. Bottom-up and shotgun proteomics have been used widely in combination with high resolution LC-MS in the past decade (Zhang *et al.*, 2013).

Other commonly used methods for protein separation are two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE), which exploit the molecular mass and charge to separate thousands of intact proteins on a single gel. In a first step proteins are solubilised using detergents like 3-[(3-

cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) and sodium dodecyl sulfate (SDS) and denatured with chaotropes such as thiourea and urea. Protein identification then either is done using matrix-assisted laser desorption/ionisation-mass spectrometry (MALDI-MS) or LC-MS after in-gel digestion (Issaq and Veenstra, 2008; Zhang *et al.*, 2013).

Besides sampling and protein extraction, sample preparation is a crucial step in proteomic workflows. Depending on the protocol, proteins have to be reduced, alkylated, digested and purified from unwanted cellular compounds and other contaminations. There are several standard digestion protocols, in which samples are prepared before nano LC-MS analysis.

1.2. Sample Preparation for Proteomics

Samples often are very heterogeneous protein mixtures with different properties (Zhang *et al.*, 2013). One of the most important steps in sample preparation is protein digestion. To do that efficiently, the protein has to be unfolded to get better access to the cleavage sites. There are various protocols, that use different strategies to achieve the same goal (Zhang *et al.*, 2013), like filter-aided sample preparation (FASP), in-solution digestion (InSol) and single-pot, solid-phase-enhanced sample preparation (SP3), as can be seen in Figure 1. These methods have been used more frequently over the last couple of years (Sielaff *et al.*, 2017).

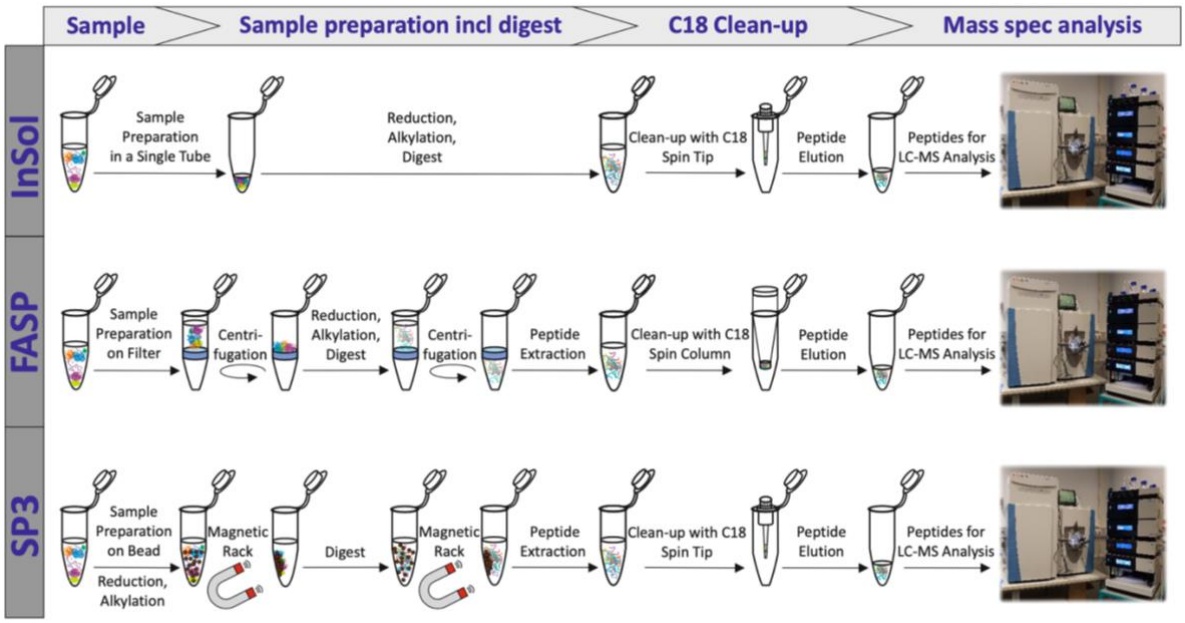

***Figure 1**: Schematic illustration of protein digestion protocols InSol, FASP and SP3*

In-solution digestion became more popular than in-gel digestion after MS began being used as the standard protein analysis tool (Medzihradszky, 2005). It offers more control over the outcome because conditions can be modified more easily and sample recovery is more reliable (Medzihradszky, 2005). In-solution digestion can be used after protein extraction into solution, then protein reduction, alkylation and digestion is performed in the same solution (Medzihradszky, 2005; León *et al.*, 2013). One major disadvantage of in-solution digestion is the difficulty to eliminate SDS or other detergents used to improve solubilisation. This can be achieved more easily with in-gel digestion (Konigsberg and Henderson, 1983; Medzihradszky, 2005).

FASP was first introduced as a method combining the advantages of in-gel digestion and in-solution digestion. The advantages of in-gel digestion include robustness against impurities and in-solution digestion minimises sample handling and can be automated more easily (Wiśniewski *et al.*, 2009). SDS often is the detergent of choice to solubilise cells or tissues, unfortunately detergents in general interfere with enzymatic digestion and later also with HPLC separation and MS analysis. Removal of SDS has been seen as impossible without substantial sample loss until it was discovered, that using an ultrafiltration membrane equilibrated with 8 M urea remove detergents (Nagaraj *et al.*, 2008; Wiśniewski *et al.*, 2009). FASP protocol takes advantage of this and uses urea for dissociating SDS from proteins and getting the micelle concentration below the critical point by reducing the detergent concentration. This strategy enables the removal of SDS completely with a single step buffer exchange. Proteins are then reduced, alkylated and digested directly on the filter. After that, peptides can be extracted by centrifugation (Erde *et al.*, 2017).

As PCR had prompted the production of templates for DNA sequencing, subsequent purification had become an issue because organic extraction, filtration and centrifugation were either time-consuming or expensive (Deangelis *et al.*, 1995). The use of magnetic particles had been introduced with a method called solid-phase reversible immobilisation (SPRI). It deploys carboxyl coated magnetic particles that bind DNA when combined with polyethylenglycol (PEG) and salt (Deangelis, Wang *et al.*, 1995), a process similar to the one used in hydrophilic interaction liquid chromatography, which was introduced in 1990 (Alpert, 1990; Moggridge *et al.*, 2018). However, methods using paramagnetic beads were not common in proteomics. SP3 uses this principle with carboxyl coated paramagnetic beads (Hughes *et al.*, 2014). Proteins are reduced and alkylated before being immobilised on-bead. By adding an organic solvent to the beads in aqueous solution, proteins and peptides are

trapped in a solvation layer (Hughes *et al.*, 2014). In a further step, proteins are purified by rinsing the beads on a magnetic rack using ACN and EtOH, which have proven effective in removing contaminants (Hughes *et al.*, 2014). After digestion, the tube is placed on a magnetic rack and peptides are recovered in the supernatant after being acidified (Hughes *et al.*, 2014).

1.3. Mass Spectrometry

Mass spectrometry is one of the most important tools for identification, characterisation and quantification of proteins in life science research (Domon and Aebersold, 2006). Mass spectrometric measurements use the mass-charge ratio of ions to separate different molecular ions (Aebersold and Mann, 2003). Mass spectrometers consist of an ion source, where the analytes are ionised before being transferred into a mass analyser, which measures the mass-charge ratio of the ions. Finally, they enter a detector to register the number of ions (Aebersold and Mann, 2003). Electrospray ionisation (ESI) ionises the analytes out of solution and is coupled to liquid chromatography (LC). Another method to ionise the proteins is matrix-assisted laser desorption/ionisation (MALDI). MALDI uses laser pulses to ionise samples in dry crystalline matrix (Aebersold and Mann, 2003).

Before MS analyses, sample compounds shall be separated. High performance liquid chromatography (HPLC) is an effective chromatographic separation tool. HPLC uses a solid stationary phase that interacts with a mobile phase sample mixture, single sample components then are separated on the stationary phase (Yandamuri *et al.*, 2013).

Nano HPLC is based on the same principle as standard HPLC, but with advantages of higher efficiency, the ability to analyse small amounts of sample and because of the low flowrate of 40-600 nl/min, it is more compatible with MS, especially for peptide analysis (Zotou, 2012). Those are some of the reasons why nano HPLC is a widely used separation method in shotgun proteomics.

1.4. Why SP3 is a Promising Novel Protein Digestion Protocol

SP3, as stated above, uses paramagnetic beads to immobilise proteins and peptides for clean-up and digestion (Hughes *et al.*, 2019). This offers a wide range of possibilities. It can be used with several binding solvents like EtOH, ACN, isopropanol and acetone and with different binding pH (Moggridge *et al.*, 2018). SP3 has also proven to work under conditions that are not compatible with lots of other protein preparation methods, like with SDS, Triton X100,

Tween 20, NP-40 and deoxycholate (Moggridge *et al.*, 2018). SP3 can also be applied for a wide range of solvent/lysate ratios and even for capturing individual proteins from noncomplex mixtures (Moggridge *et al.*, 2018). It is a robust high throughput method enabling automated processing. Because of the single tube, it offers rapid and simple handling with minimal sample loss, even in sub-microgram amounts of material (Hughes *et al.*, 2014; Moggridge *et al.*, 2018).

The usage variety offers the possibility of modified workflows adjusted to individual experimental conditions and extended compatibility (Moggridge *et al.*, 2018).

The aim of this study was to determine, whether the performance of SP3 is comparable to other established protein digestion protocols. In addition, the suitability of the method for being used in a core facility was investigated. Furthermore, an improvement of the protocol was carried out to achieve a higher protein identification rate.

2. Material and Methods

2.1. Protein Digestion

Standardised yeast extract (Promega, Madison, WI, USA) was used in all experiments after being diluted from 10 µg/µl stock solution to 1 µg/µl with phosphate buffered saline (PBS) (Sigma-Aldrich , St. Louis, MO, USA). The protease used for digestion was a mixture of Trypsin and LysC (Promega, Madison, WI, USA).

2.1.1. Single-Pot, Solid-Phase-Enhanced Sample Preparation (SP3)

For the bead mixture preparation 20 µl of two different types of carboxylate coated hydrophilic SpeedBeads™ (Cytiva, Marlborough, MA, USA) were combined. After being washed with 160 µl of MilliQ-water they were placed on a magnetic rack (PCR Strip Magnetic Separator 12 Strip, Peabody, MA, USA) to remove the water. The beads were then rinsed two more times with 200 µl water and finally 20 µl of water were added. The prepared beads were stored at +4 °C until further use.

Following solutions needed for the protocol were prepared beforehand:

- Sample loading: 100 mM Triethylammonium bicarbonate (TEAB) (Sigma-Aldrich) solved in MilliQ-water
- Reduction: 200 mM Tris-(2-carboxyethyl)-phosphin hydrochloride (TCEP) (Carl Roth, KA, Germany) solved in 100 mM TEAB
- Alkylation: 800 mM Chloroacetamide (CAA) (Sigma-Aldrich) solved in 100 mM TEAB
- Washing solution: 56.5 % LiChrosolv® Acetonitrile (ACN) hypergrade for liquid chromatography–mass spectrometry (LC-MS) (Merck KGaA, Darmstadt, Germany) mixed with 100 mM TEAB (Sigma-Aldrich)
- Ethanol washing solution: 80 % EtOH (Thermo Fisher Scientific, Waltham, MA, USA)
- Acidification solution: 40% trifluoroacetic acid (TFA) (Fisher Scientific)

The sample was prepared for digestion by diluting the yeast extract (1 µg/µl) in a mixture of TCEP, and CAA solved in TEAB. Then the thermomixer (Thermomixer C/comfort, Eppendorf,

Hamburg, Germany) was preheated to 99 °C. After placing the sample on the thermomixer, the temperature was reduced to 70 °C. For reduction and alkylation of disulphide bonds, samples were incubated for 25 minutes at 70 °C. The sample (30 µl) was loaded in a mixture of 113 µl ACN and 55 µl 100 mM TEAB onto 2 µl beads for protein binding. The solution was then sonicated (performed in Sonorex DT 103 H, BANDELIN electronic GmbH & Co. KG, Berlin, Germany) for 10 minutes and incubated at room temperature for another 20 minutes. It was put on a magnetic rack and the liquid phase was removed from the beads. Proteins bound on the beads were rinsed with “washing solution”, then twice with 80 % ethanol and 100 % ACN at the end with short incubation times and resuspension in between every step. Yeast proteins were digested using a combination of Trypsin/LysC. For digestion beads were solved in 70 µl 100 mM TEAB kept for 14 hours at 37 °C before cooling down to 4 °C until the tubes were removed from the thermomixer.

For peptide extraction, the suspension was acidified to a pH below 2 using 40 % TFA. Liquid was then removed using the magnetic rack. The beads were captured on the wall of the tube and the liquid containing the peptides was transferred into a fresh tube.

2.1.2. Modified Single-Pot, Solid-Phase-Enhanced Sample Preparation (SP3)

The modified SP3 protocol for optimisation used fewer rinsing steps. The first one consisting of ACN and 100 mM TEAB was left out. Also the starting bead amount was doubled and the first flow-through after protein binding was loaded on the same amount of fresh beads again. This resulted in a quadruple of the standard bead amount.

2.1.3. in-solution Digestion

in-solution digestion was performed applying a standard protocol according to the Trypsin/LysC user manual (Promega). The in-solution protocol was used as a reference method as well as for digestion of SP3 fractions from various washing steps. Following solutions were prepared before starting:

- Reduction: 50 mM Dithiothreitol (DTT) (Sigma-Aldrich)
- Alkylation: 100 mM Iodoacetamide (IAA) (Sigma-Aldrich)
- 8 M urea (Carl Roth) in 50 mM TRIS (Carl Roth)
- 50 mM TRIS (Carl Roth)

For the standard protocol, first the sample was evaporated and solved in 8 M urea in 50 mM TRIS (pH 8). Then 1.1 µl DTT was added and the sample was incubated for 30 minutes at 37 °C to reduce the disulphide bounds of the proteins. After that, 0.9 µl IAA was added and the sample was incubated for 30 minutes at 25 °C for alkylation. For digestion Trypsin/LysC mix was added. Then the sample was diluted with 50 mM TRIS to reduce the urea concentration below 1 mol/l. Proteins were digested for 8 hours at 37 °C and cooled down to 4 °C until the tubes were removed from the thermomixer. To stop the digestion, 0.4 µl concentrated TFA was added.

The digestion of the washing fractions did not require reduction and alkylation because this was already performed in the SP3 protocol.

2.1.4. Filter-aided sample preparation (FASP)

An ultrafiltration unit with a cut-off of 10 kDa (Pall Cooperation, Washington Port, NY, USA) was used for the FASP protocol. It was prepared by rinsing the filter three times with 500 µl 8 M urea in 50 mM TRIS followed each time by centrifugation for 20 minutes at 10,000 rcf and discarding the flow-through afterwards. The first centrifugation step started with another minute added, to check if the membrane is still intact. Protein samples were filled up to 500 µl with 8 M urea in 50 mM TRIS and were loaded onto the filter. The sample mixture then was centrifuged for 20 minutes at 10,000 rcf and the flow-through was discarded. In the next step proteins were reduced with 200 mM DTT added directly onto the filter and incubated at 37 °C for 30 minutes. The proteins were then alkylated with 500 mM IAA with 30 minutes incubation at 25 °C. Liquid phase was filtered by centrifugation for 15 minutes at 10,000 rcf. Afterwards, the filter was washed two times with 100 µl 50 mM TRIS and the collection tube was also washed with 100 µl 50 mM TRIS to get rid of all urea remains. Then 20 µl 50 mM TRIS was put into the collection tube to prevent the membrane from drying out. After that digestion was carried out using Trypsin/LysC (Promega, Madison, WI, USA) mix for 14 hours at 37 °C. After centrifugation for 15 minutes at 10,000 rcf, the digested peptides were recovered with 3 times 50 µl of 50 mM TRIS and 15 minutes centrifugation at 10,000 rcf after each step and acidified with 1 µl concentrated TFA.

2.2. Peptide Clean-up and Purification

After digestion, the peptides have to be purified to get rid of other organic compounds, unwanted liquid, salts etc. Therefore, depending on the protein amount either Pierce™ C18 Spin Tips or C18 Columns containing reversed-phase sorbent were used for all digestion protocols before LC-MS analysis.

2.2.1. Pierce™ C18 Spin Tips

Spin Tips have a maximum capacity of 10 µg protein. These were used for clean-up of all SP3 and in-solution digested samples. The washing solution consisted of 0.1 % TFA. A mixture of 80 % ACN, 20 % MilliQ-water and 0.1 % TFA was used as wetting/elution solution.

The C18 spin tips were inserted into an adapter and placed into a 1.5 ml receiver tube. During the whole protocol each loading/washing/elution step was followed by centrifugation at 1,000 rcf for 1 minute (Centrifuge 5424, Eppendorf). Before peptide loading, tips were wetted with 20 µl wetting solution and equilibrated with 20 µl washing solution. After transferring tip and adapter to a new 1.5 ml receiver tube, the peptide sample was loaded onto the tip. In order to ensure binding of the peptides to the resin, the flow-through was reloaded onto the tip. Afterwards, two washing steps with 0.1% TFA followed to remove salts and contaminants. After transferring tip and adapter to a new 1.5 ml receiver tube once again, peptides were eluted twice each with 20 µl elution solution.

After evaporation of elution solution using a vacuum concentrator (Concentrator plus, Eppendorf) the peptides were resuspended in 25 µl 0.1 % TFA to reach a protein concentration of 0.1 µg/µl before LC-MS analysis.

2.2.2. Pierce™ C18 Spin Columns

Spin columns have a bigger capacity and can process up to 30 µg digested protein. They were used for clean-up of all FASP samples. The activation solution consisted of 50 % ACN and 50 % MilliQ-water, the equilibration and wash solution of 5 % ACN, 95 % MilliQ-water, and 0.5 % TFA. A mixture of 70 % ACN, 30 % MilliQ-water and 0.1 % TFA was used as elution solution.

The bottom and top cap of the C18 spin columns were removed and the column was placed into a 1.5 ml receiver tube. During the whole protocol each loading/washing/elution step was

followed by centrifugation (Centrifuge 5424, Eppendorf) at 1,500 rcf for 1 minute and discarding of the flow-through. Before peptide loading, column wall and resin were rinsed with 200 µl activation solution and equilibrated with 200 µl equilibration solution. The peptide sample then was loaded onto the column. In order to ensure binding of the peptides to the resin, the flow-through was reloaded. Afterwards, two washing steps with washing solution followed to remove salts and contaminants. After that, peptides were eluted twice each with 20 µl elution solution.

After evaporation of elution solution using a vacuum concentrator the peptides were resuspended in 300 µl 0.1 % TFA to reach a protein concentration of 0.1 µg/µl before LC-MS analysis.

2.3. Peptide Analysis using nano LC-MS/MS

Three technical replicates of a sample were placed in an autosampler and every replicate was analysed by LC-MS as duplicates. Per run, 3 µl peptides were injected into a nano HPLC Ultimate 3000 RSLC system (Thermo Fisher Scientific, Waltham, MA, USA). First they were pre-concentrated and desalted with a 5 mm Acclaim™ PepMap™ µ-Precolumn (300 µm inner diameter, 5 µm particle size, and 100 Å pore size) (Thermo Fisher Scientific). For loading and desalting of the peptides, 2 % ACN in ultra-pure H_2O (Optima® LC/MS W6-212 Water, Thermo Fisher Scientific) with 0.05 % TFA was used as a mobile phase with a flow rate of 5 µl/min.

For separation of the peptides, a 25 cm Acclaim PepMap™ C18 column (75 µm inner diameter, 2 µm particle size, and 100 Å pore size) (Thermo Fisher Scientific) with a flow rate of 300 nl/min was used. The gradient started with 4 % B (80 % ACN with 0.08 % formic acid) for 7 min, increased to 31 % in 60 min and finally to 44 % in 5 min. It was followed by a washing step with 95 % B. Mobile Phase A consisted of ultra-pure H_2O with 0.1 % formic acid. To perform the mass spectrometric analysis, the LC was directly coupled to a high resolution Q Exactive HF Orbitrap mass spectrometer via an electrospray ionisation (ESI) nano source interface (Thermo Fisher Scientific). For the MS full scans ultra-high-field-Orbitrap mass analyser (Thermo Fisher Scientific) was used in the mass range of m/z 350−2,000 with a resolution of 60,000. The maximum injection time (MIT) was 50 ms and the automatic gain control (AGC) was set to $3e^6$. The 10 ions with the highest intensity were subjected to HDC cell for further fragmentation via higher energy collisional dissociation (HCD) activation at a resolution of 15,000 over a mass range between m/z 200 and 2,000 and with the intensity threshold at $4e^3$.

Ions with charge state of +1 and >+7 were excluded. Normalised collision energy (NCE) was set at 28. The AGC was set at $5e^4$ for each scan and the MIT was 50 ms. To suppress repeated peak fragmentation, dynamic exclusion of precursor ion masses over a time window of 30 s was used.

2.4. Protein Identification and Quantification

Protein and peptide identification and quantification was done by comparing the raw data provided from mass spectrometry to a preselected database using the software Thermo Proteome Discoverer (PD) 2.4.0.305™ software (Thermo Fisher Scientific) and by comparing samples with different preparation to each other.

2.4.1. Protein Database

The *Saccharomyces cerevisiae* (taxonomy ID: 4932) database was downloaded from www.uniprot.org containing 27,574 different protein sequences. Contamination of the sample from hair, skin particles etc can occur in any step of the sample preparation despite careful working. To identify these peptides, another database was added, containing common contaminant proteins. This database called “cRAP” was downloaded from www.thegpm.org/crap containing 116 protein sequences.

2.4.2. Protein Identification

Protein identification in PD was done in two steps, first the database search was performed ("Processing Step") and in a second step, the statistical verification of the results took place ("Consensus Step"). The "Processing Step" was designed as shown in Figure 2. The nodes used are for spectra selection ("Spectrum Files"), which are retrieved to be processed ("Spectrum Selector") and finally the database search is performed ("Sequest HT"). Posterior error probabilities as well as *q*-value of the identified peptide spectra matches are then calculated ("Percolator").

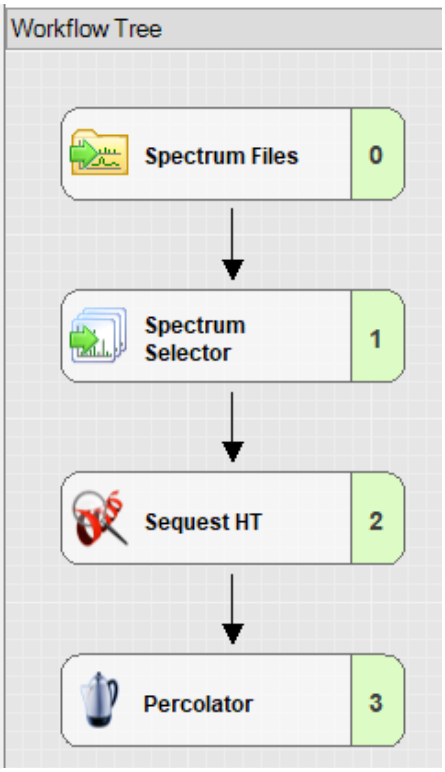

Figure 2: *Workflow "Processing Step" for identification (Source: Thermo Proteome Discoverer 2.4.0.305™ software)*

The "Percolator" is used to tell apart correct and incorrect peptide-spectrum matches (PSM) by false discovery rate (FDR) estimation with a decoy database search strategy. The *q*-value for PSMs was set to 0.05, the strict target false discovery rate (FDR) to 0.01 and the relaxed FDR to 0.05.

All nodes were used with the default parameters except for the "Sequest HT"-node. All changed parameters are shown in Table 1.

Table 1:" *Processing Step": Setting Sequest HT-node*

Max. missed cleavage sites	2
Min. peptide length	6
Max. peptide length	144
Precursor mass tolerance	10 ppm
Fragment mass tolerance	0.02 Da
Dynamic peptide modification	Oxidation: +15.995 Da
Dynamic protein modification	Acetyl: +42.011 Da (N-Terminus)
Dynamic protein modification	Met-loss: -1331.040 Da
Dynamic protein modification	Met-loss + Acetyl: -89.030 Da
Static peptide modification	Carbamidomethyl: +57.021 Da (C-Terminus)

The “Consensus step” consisted of eight nodes as can be seen in Figure 3. In this second step, the search files were selected (“MSF Files”) and the redundantly identified PSMs were put together to peptide groups (“PSM Grouper”). In a further step, the confidences of PSMs and peptides were calculated (“Peptide Validator”), results were filtered by preselected filters (“Peptide and Protein Filter”) and the scores of the filtered proteins were calculated (“Protein Scorer”). Now after the protein confidences were assigned (“Protein FDR Validator”) proteins that share the same set of peptides were grouped together (“Protein Grouping”). All these nodes were used with the default parameters. An additional node (“Protein Marker”) was used to mark all contaminants found in the “cRAP” database to easily filter them later on.

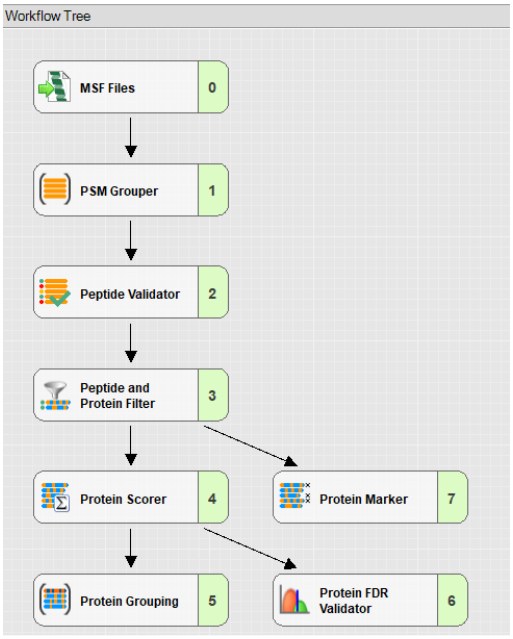

Figure 3: *Workflow „Consensus Step” for identification (Source: Thermo Proteome Discoverer 2.4.0.305™ software)*

2.4.3. Protein Quantification

Label free quantification was done in PD just to compare protein abundances of the analysed MS1 spectra. Therefore, a similar workflow as in the identification step with some modifications was used. The “Processing Step” is replaced (Figure 4) the “Spectrum Files” node with the “Spectrum Files RC” to recalibrate the precursor mass. Additionally, the node “Minora Feature Detector” was used to detect chromatographic peaks and features.

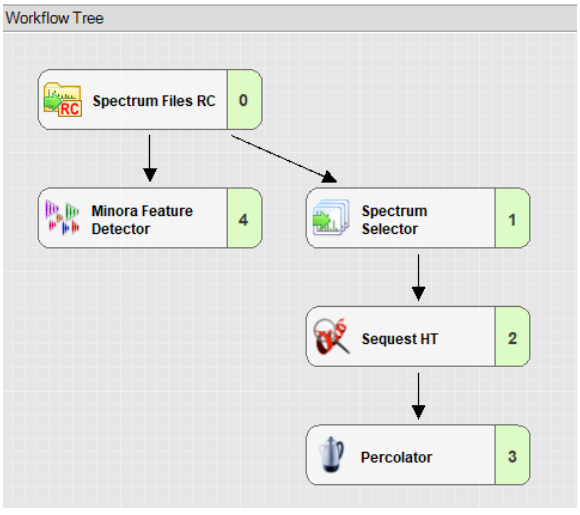

Figure 4: *Workflow “Processing Step” for quantification (Source: Thermo Proteome Discoverer 2.4.0.305™ software)*

The “Consensus Step” for the quantification had four additional nodes: “Feature Mapper” to map the features across multiple files, “Precursor Ions Quantifier” to calculate quan ratios, “Protein Annotation” to add protein annotation available in the used protein database and “Peptide in Protein Annotation” to annotate the flanking residues and to show the position of the peptide in the protein sequence Figure 5. All these nodes were used with the default parameters as only the sample abundance plots were of interest in this thesis.

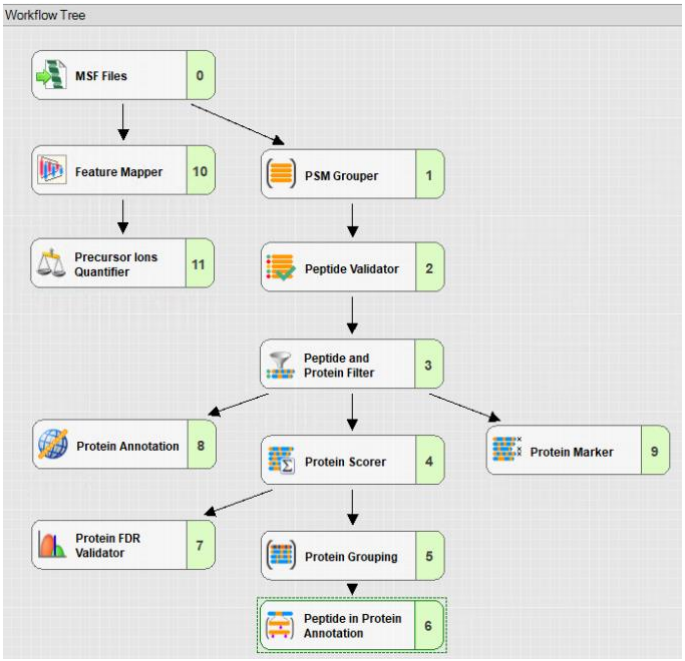

Figure 5: *Workflow “Consensus Step” for quantification (Source: Thermo Proteome Discoverer 2.4.0.305™ software)*

2.5. Statistical Evaluation

To determine whether the differences in the number of protein ID between the various methods and protocol modifications are statistically relevant, data was analysed using IBM SPSS Statistics version 27. Two sample t-test was performed for two independent groups, whereas for three and more one-way ANOVA test was done. Differences between two groups were considered as significant with a p-value < 0.05.

3. Results

3.1. Standard Protocol

The first section deals with the determination of the *status quo* of the protein digestion methods. The SP3 standard protocol without modifications (described in 2.1) was used to compare with other standard protocols and for the evaluation of the bead storage stability.

3.1.1. Comparison of SP3 to Other Standard Digestion Protocols

To determine how effective SP3 digestion is, a comparison to in-solution digestion and FASP was carried out. The results are shown by the number of protein identifications in Figure 6. Digestion was performed with 2.5 µg yeast extract for all protocols and additionally with 30 µg for FASP, the appropriate amount of the standard FASP protocol. As can be seen, FASP delivered comparable identification rates to in-solution with lower protein amount. However, SP3 reached around a third fewer IDs than in-solution digestion and FASP with 30 µg but around the same as FASP with 2.5 µg protein. This was shown to be statistically significant at a significance level of $p < 0.05$.

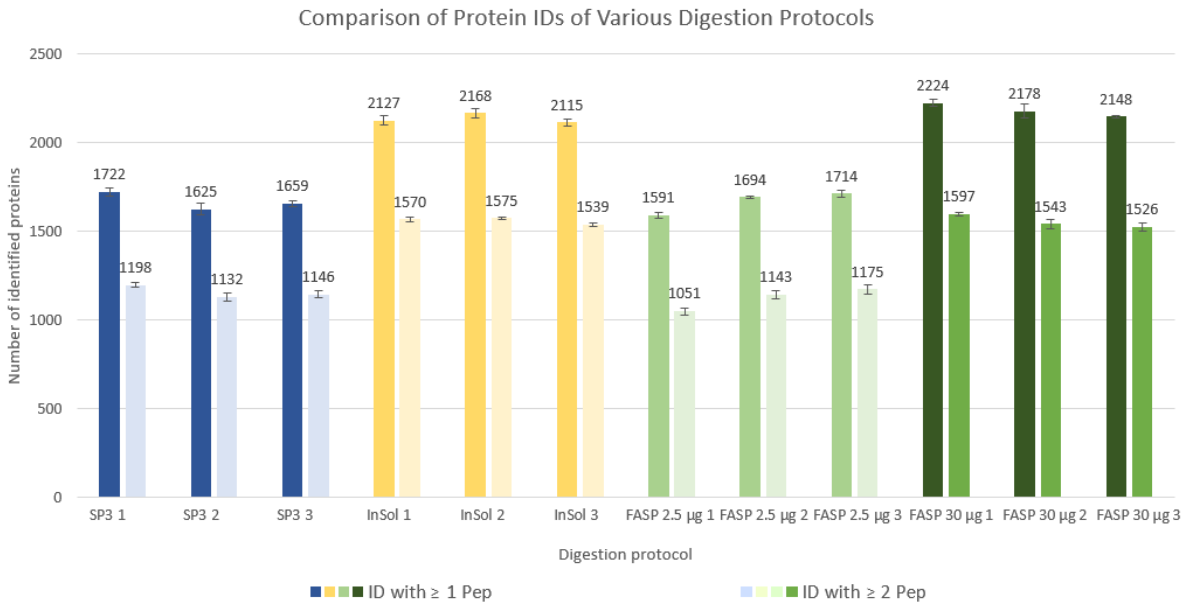

Figure 6: *Protein IDs in bar chart comparing SP3, InSol, FASP loaded with 2.5 µg and 30 µg protein*

The log(10)-transformed protein abundance of the different groups is shown as a scatterplot in Figure 7. The abundance of every protein of a digestion method is compared to another digestion method. The R^2 value of SP3 vs. FASP 2.5 µg and InSol is not as high as FASP 30 µg vs. FASP 2.5 µg and InSol vs. FASP 2.5 µg. That underlines, that InSol and FASP 30 µg differ from SP3 and FASP 2.5 µg in their performance to deliver comparable protein abundance.

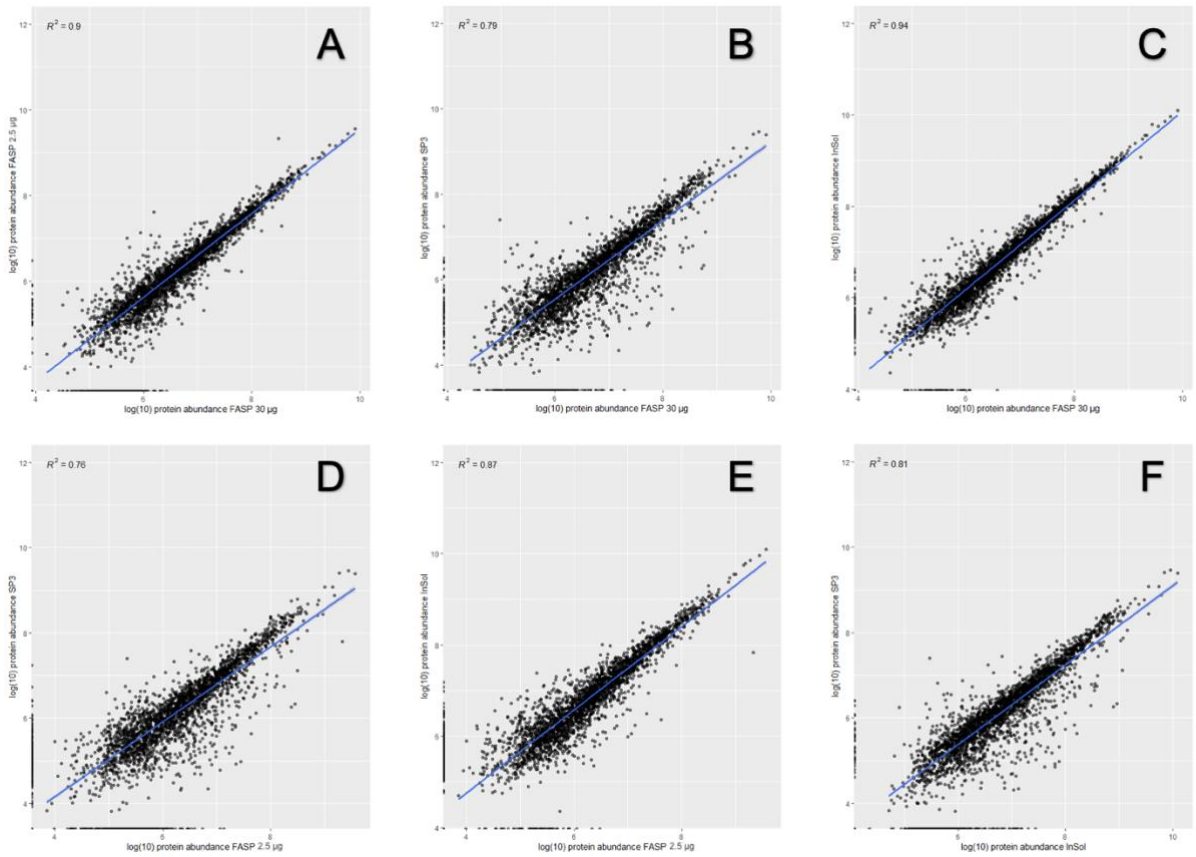

Figure 7: *Scatterplot log(10)-transformed abundances of identified proteins per group of each digestion protocol, FASP 2.5 µg vs. FASP 30 µg (A), SP3 vs. FASP 30 µg (B), InSol vs. FASP 30 µg (C), SP3 vs. FASP 2.5 µg (D), InSol vs. FASP 2.5 µg (E), SP3 vs. InSol (F)*

Figure 8 shows the protein abundance after log(10)-transformation of all technical and LC-MS replicates of each sample as a boxplot. In general, total abundance was roughly in the same range within all digestion methods, nevertheless in-solution samples showed the highest mean abundance and the SP3 samples the lowest. Furthermore, it points out the consistency in abundance within SP3 and in-solution samples compared to FASP. Higher variability could be noticed for the FASP protocol for both loaded sample amounts.

Even though SP3 performed well regarding reproducibility, however there is a need for improvement in terms of the total abundances. Therefore, in order to increase protein binding for SP3, several different strategies were used and described in 3.2 summarised in Table 2 in this chapter.

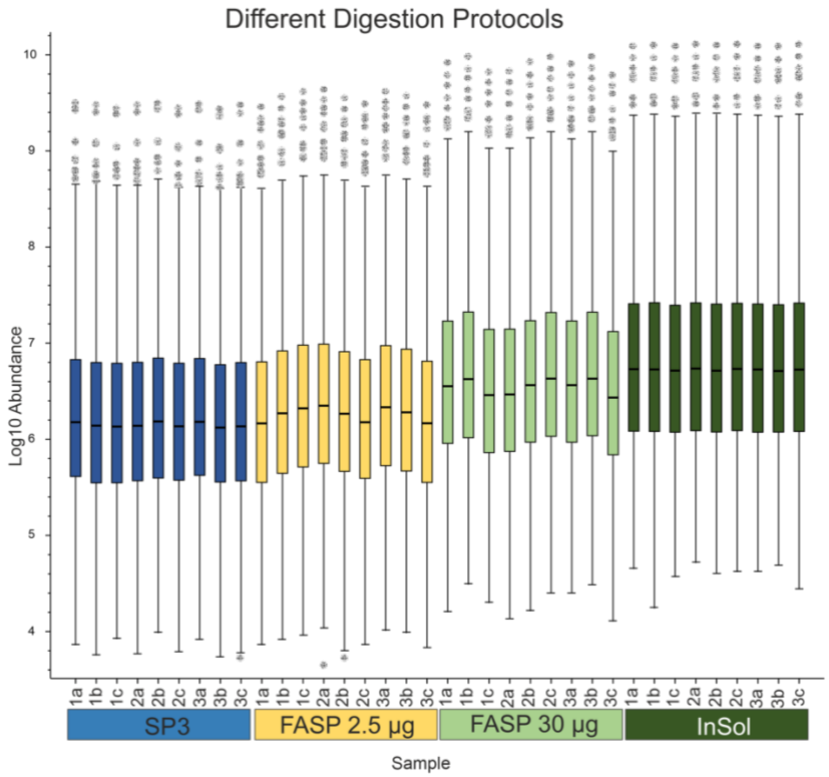

***Figure 8**: Log(10)-transformed protein abundance of technical and LC-MS replicates as boxplot of each digestion method (SP3, FASP, in-solution)*

3.1.2. Influence of Storage Time of Paramagnetic Beads

To determine the storage stability of the carboxylate coated hydrophilic SpeedBeads™ mixture, SOP protocol was performed with the same beads on the first day, after 7 and 14 days. As reference, stored beads were also analysed in comparison to fresh beads in order to monitor day-to-day variability (data not shown). Figure 9 shows the number of identified proteins of SP3 standard protocol with beads after different storage times. The results showed no statistically significant differences ($p < 0.05$) in mean protein IDs related to the beads stored for several days at +4 °C.

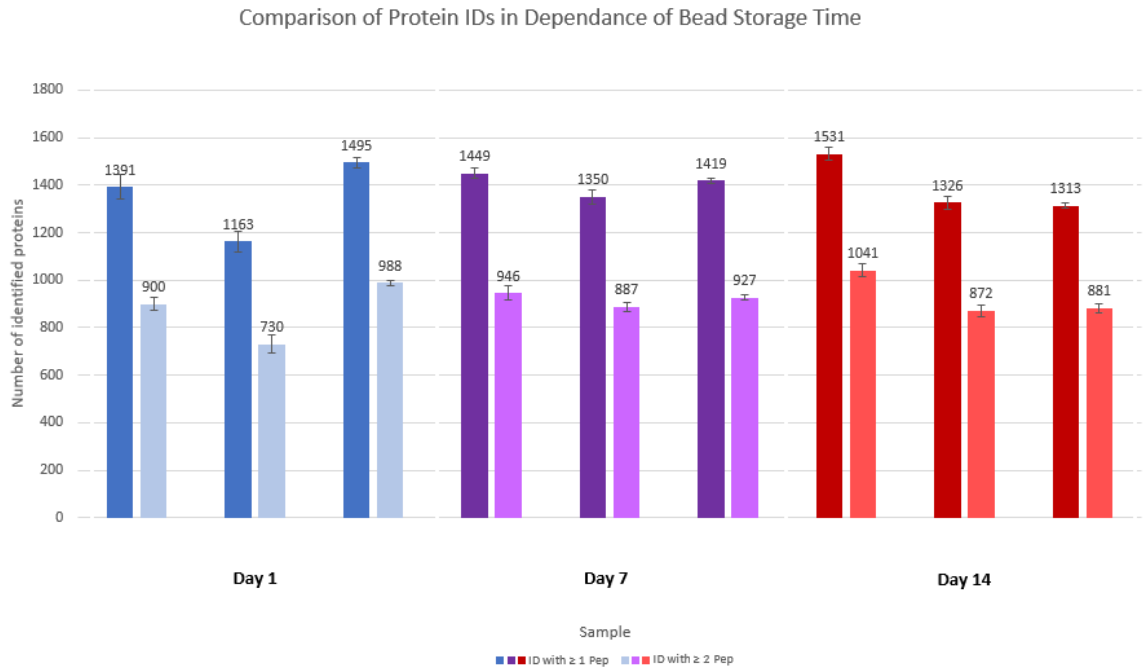

Figure 9: *Bar chart of protein IDs comparing sample preparation with freshly mixed and stored paramagnetic beads*

The log(10)-transformed abundances of each sample with technical and LC-MS replicates using beads with different storage times are displayed in Figure 10. On day 1 proteins had the lowest consistency in abundance when compared to day 7 and 14, but overall the highest abundance of the groups was on a similar level.

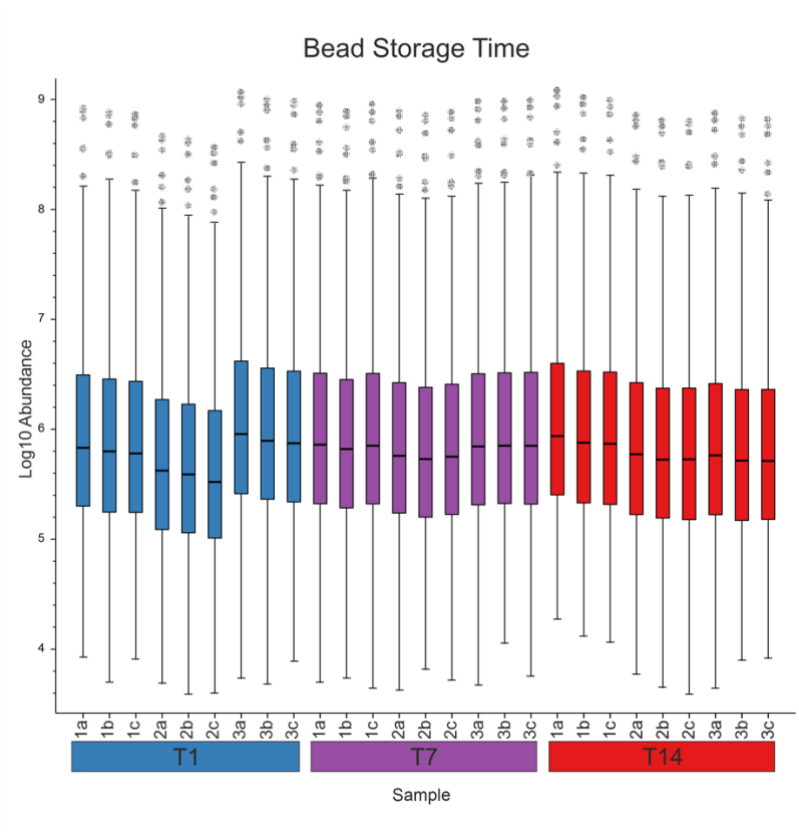

Figure 10: *Log(10)-transformed abundances of technical and LC-MS replicates as boxplot with SP3 protocol using beads on day 1, day 7 and day 14*

3.2. Optimisation of SP3 Sample Preparation

The second section describes the results of the modifications applied to the original standard protocol as seen in Table 2.

Table 2: *Summary of modified protocols*

Modification	Protein amount	Bead solution amount	Modified step	Mean protein IDs
Incubation time	2.5 µg	2 µl	Incubation on magnetic rack 15 minutes instead of 1 minute	1344
Washing step	2.5 µg	2 µl	Leaving out washing step with ACN and 100 mM TEAB mixture	1496
Bead amount	2.5 µg	4 µl	Doubling initial bead amount	1671
Reloading FT	2.5 µg	8 µl	In addition to 4 µl bead solution, first FT was loaded onto 4 µl fresh bead solution and combined with sample	1639

3.2.1. Effect of Incubation Time on Magnetic Rack on Protein IDs

The first approach of optimisation was to decrease the loss of beads and accordingly sample bound on the beads. In order to achieve this, incubation time on the magnetic rack was increased from 1 minute to 15 minutes before removing the liquid phase (further termed as “flow-through”). As seen in Figure 11, the mean number of protein IDs was slightly higher with increased incubation time. Still, there was a noticeable inconsistency in protein IDs in either of the protocols, hence the difference was not shown to be statistically significant ($p < 0.05$).

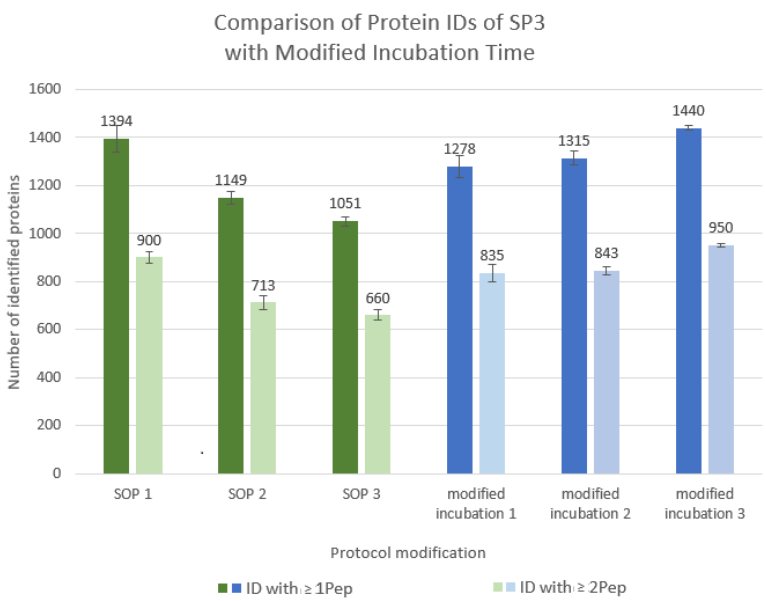

Figure 11: *Bar chart of protein IDs comparing difference between standard protocol (SOP) and longer incubation time on magnetic rack (modified incubation)*

The total log(10) abundances of the samples in the modified protocol was higher compared to the standard protocol. There were slight variabilities in both groups with no recognisable pattern seen in Figure 12.

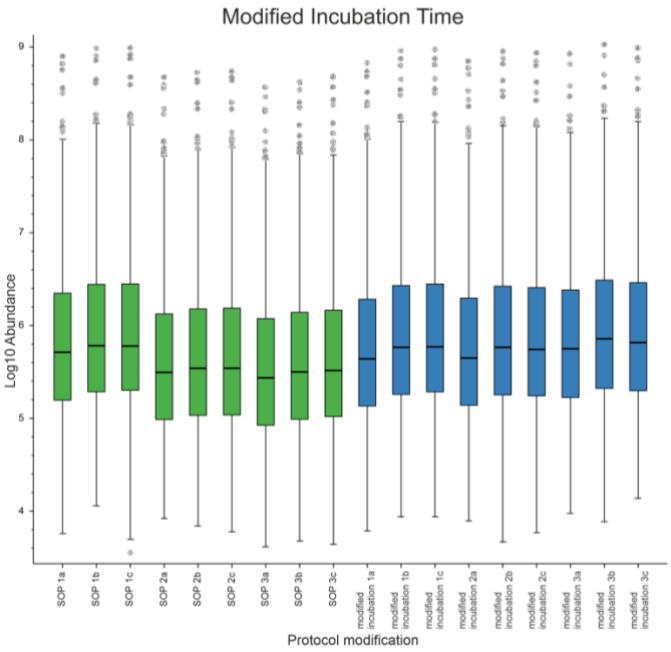

Figure 12: *Log(10)-transformed protein abundances of technical and LC-MS replicates as boxplot with standard SP3 protocol and modified protocol with longer incubation time on magnetic rack*

3.2.2. Effect of Reduction of Washing Steps on Protein IDs

Current protocols suggest, that rinsing only with EtOH and ACN is most effective (Moggridge *et al.*, 2018; Hughes *et al.*, 2019). In this modified method, the washing step using a mixture of ACN and 100 mM TEAB corresponding to the protein loading conditions was skipped in order to make handling easier and decrease bead loss in supernatant. The number of protein IDs showed to be more stable with a higher output in the modified method compared to the standard protocol as shown in Figure 13. The mean number of protein IDs was higher in the modified protocol than in the standard SOP, even though this difference was not shown to be statistically significant ($p < 0.05$).

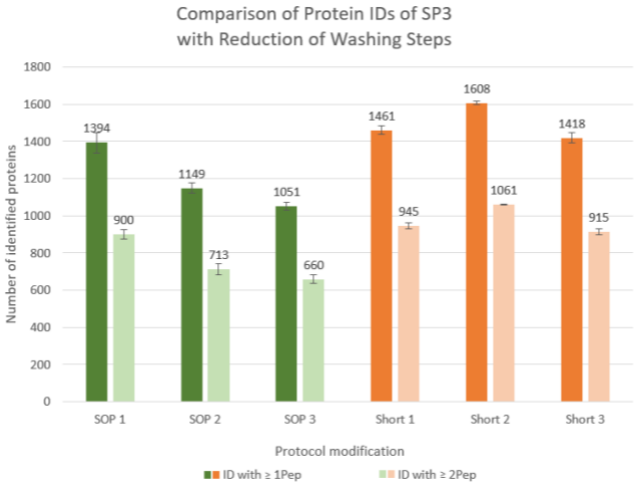

Figure 13: *Bar chart of protein IDs comparing difference between standard protocol (SOP) and protocol with one washing step left out (Short)*

The protein abundance of all samples with technical and LC-MS replicates comparing the yeast protein digestion with the shortened method and standard protocol is shown in Figure 14. Even though there was variation within the technical replicates, the modified method showed better consistency and slightly higher abundance in total.

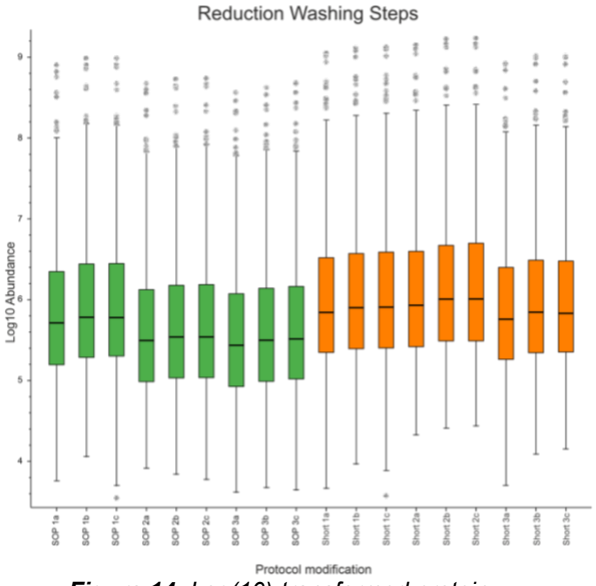

Figure 14: *Log(10)-transformed protein abundances of technical and LC-MS replicates as boxplot with standard SP3 protocol and modified protocol with one washing step left out*

3.2.3. Effect of Bead Amount on Protein IDs

Since a loss of bead material in the flow-through was suspected, the flow-throughs were retained, digested using an in-solution protocol and then analysed by LC-MS. As seen in Figure 15, there were 833 proteins discovered in flow-through after the protein binding incubation step. Also, there was a loss of about 93 proteins in the first washing step using ACN and 100 mM TEAB mixture corresponding to the loading solvent. No protein loss was recorded in EtOH nor the ACN fractions.

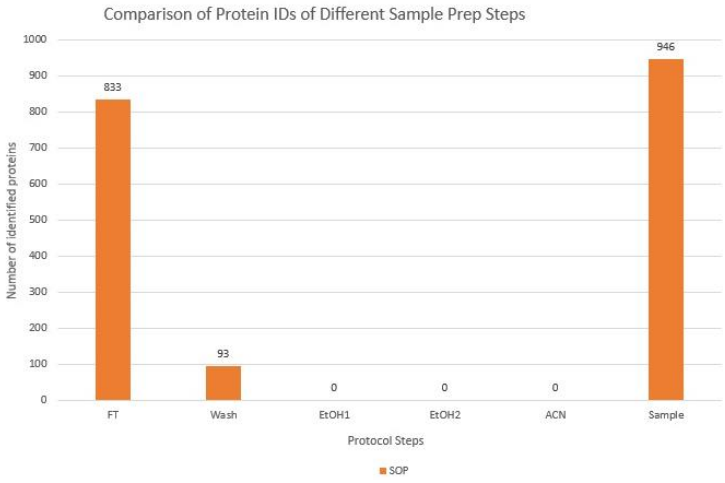

Figure 15: *ID number of proteins with one or more peptides after InSol digestion of the rinsing fractions (flow-through, wash, EtOH 1, EtOH 2, ACN), in comparison to the SP3 digested sample*

To improve protein binding, the bead to sample ratio (wt/wt) was increased from 200:2.5 (2 µl bead solution) to 400:2.5 (4 µl bead solution) by doubling the amount (volume) of the bead solution combined with the reduction of washing steps. Figure 16 showed higher and very consistent numbers of protein IDs when using 4 µl instead of 2 µl beads. The standard method with 2 µl bead solution showed an outlier in one technical replicate. Differences were not found to be statistically significant ($p < 0.05$).

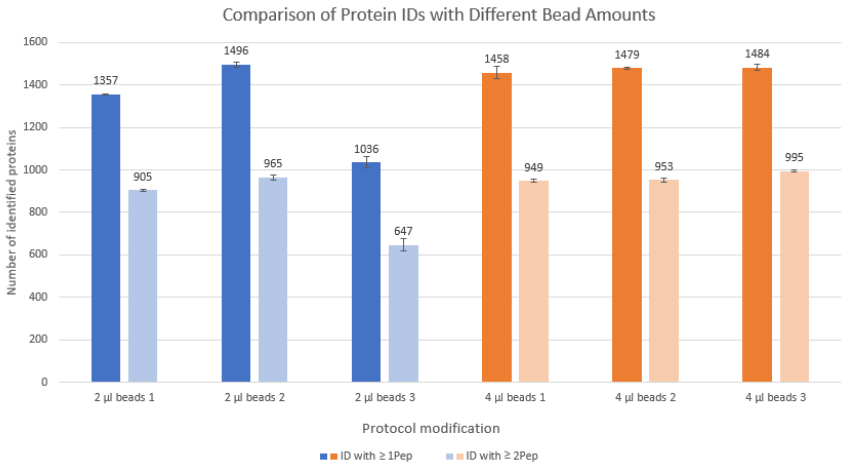

Figure 16: *Bar chart of protein IDs comparing the difference between paramagnetic bead amounts used to bind proteins (2 µl and 4 µl bead solution)*

Figure 17 plots the protein abundances comparing the three technical replicates with two LC-MS replicates for both approaches. Except one technical replicate, the mean abundances of the samples with the standard protocol (2 µl beads) were in the same range as the mean abundances of the sample with the modified method (4 µl beads).

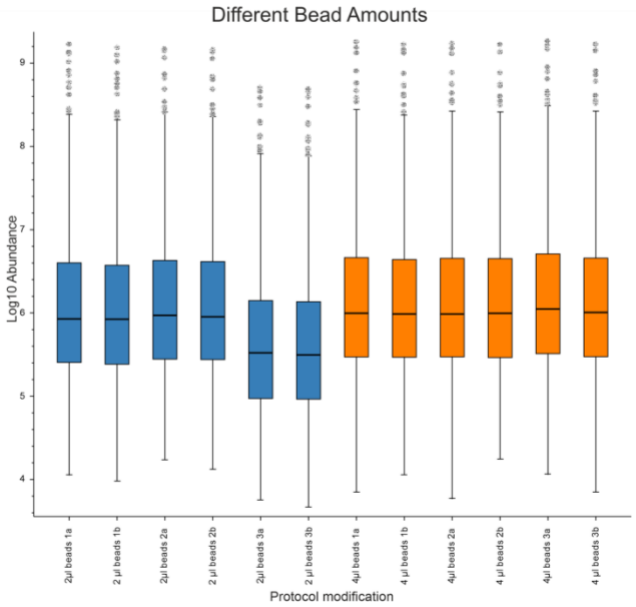

Figure 17: *Log(10)-transformed protein abundances of technical and LC-MS replicates as a boxplot comparing sample preparation with 2 µl and 4 µl bead solution*

The flow-through of the binding step of every sample was digested using an in-solution protocol and analysed with LC-MS. Except for one sample using the standard protocol, the protein IDs of both methods were consistent within the method. The fractions of the modified method (4 µl beads) even showed higher numbers of protein IDs (Figure 18). This trend was statistically significant for the protein IDs with at least one identified peptide per protein ($p < 0.05$), but not for the IDs with more than two peptides.

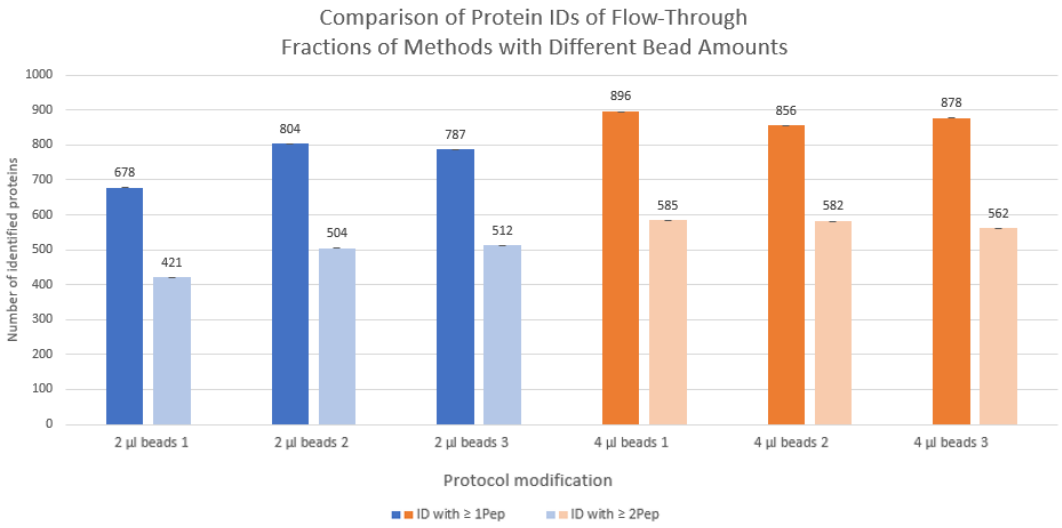

Figure 18: *Bar chart of protein IDs comparing the fractions of the two methods using different paramagnetic bead amounts for protein binding (2 µl and 4 µl bead solution)*

In the protein abundance boxplot in Figure 19, the replicates of each method showed very reproducible mean values. In general, the method with 4 µl bead solution had higher total abundances in the flow-through fraction then the method using 2 µl bead solution.

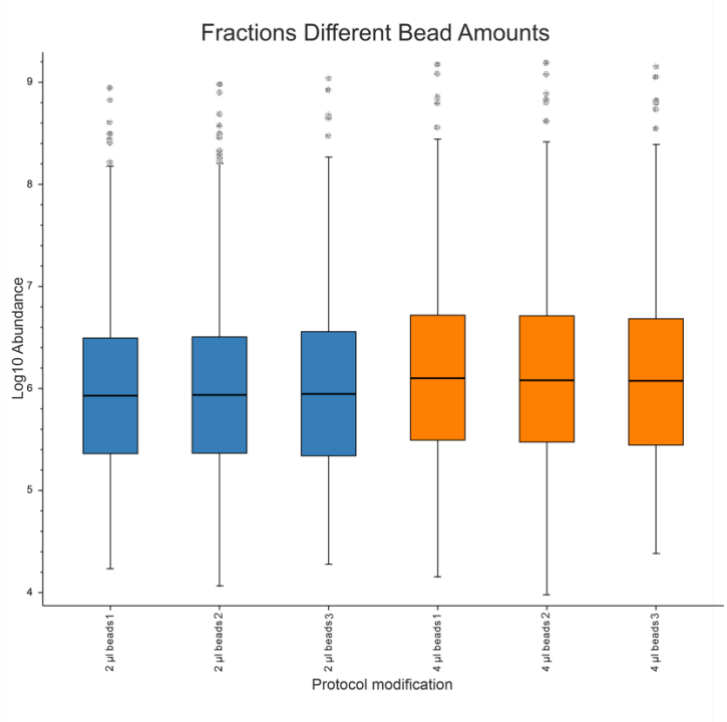

Figure 19: *Log(10)-transformed protein abundances of technical replicates as boxplot comparing fractions of the two methods using 2 µl vs. 4 µl bead solution*

3.2.4. Effect of Reloading Flow-Through on Protein IDs

To further improve the binding of proteins on beads, an additional step was undertaken. The flow-through after protein binding, that was normally discarded, was loaded onto 4 µl fresh beads. The incubation steps were performed as according to standard protocol. Then the protein loaded beads of each approach (original sample and flow-through) were combined in the original tube and the further protocol was carried out regularly. The modified binding conditions did not increase the number of identified proteins compared to only doubling the bead amount of the original protocol, as shown in Figure 20. One sample of the protocol without reloading the flow-through had a very low number of identified proteins. No statistically significant increase in the mean protein ID numbers could be observed ($p < 0.05$).

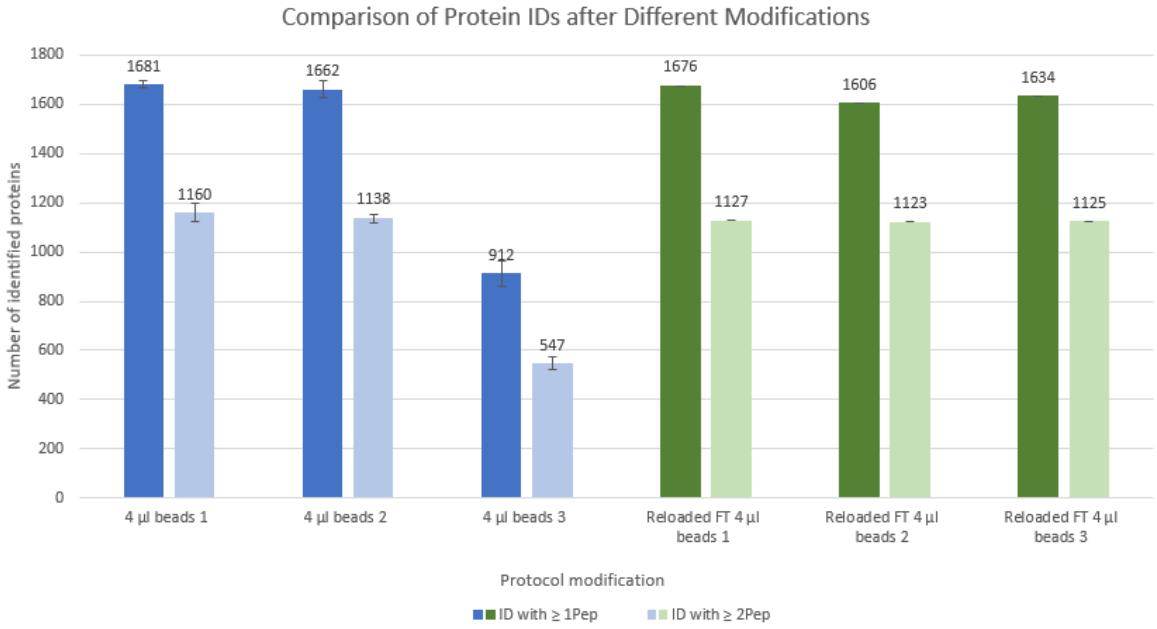

Figure 20: *Bar chart of protein IDs comparing modified methods with different paramagnetic bead amounts used to bind proteins (4 µl and 8 µl bead solution) with the 8 µl split by 4 µl for loading the sample and 4 µl for loading the flow-through*

Figure 21 shows the protein abundances of the method with modified bead amounts and one group with an additional loading step. There was one outlier with a low abundance, despite that, the mean protein abundances within and between the groups were very consistent.

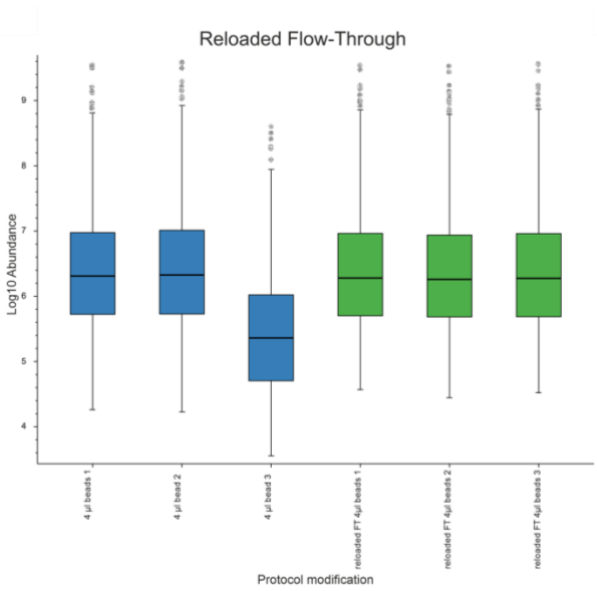

Figure 21: *Log(10)-transformed protein abundances in boxplot of technical replicates comparing the modified methods with different paramagnetic bead amounts used to bind proteins (4 µl and 8 µl bead solution) with the 8 µl split by 4 µl for loading the sample and 4 µl for loading the flow-through*

By performing in-solution digestion of the binding flow-through, the lost proteins in the supernatant were determined. Figure 22 shows the number of protein IDs after MS analysis. The method using initially 4 µl bead volume had lower and less consistent protein IDs than after reloading the flow-through onto 4 µl fresh bead solution. So was observed, that there were still proteins lost in the flow-through after two consecutive loading steps. No statistically significant differences between the two methods could be noticed ($p < 0.05$).

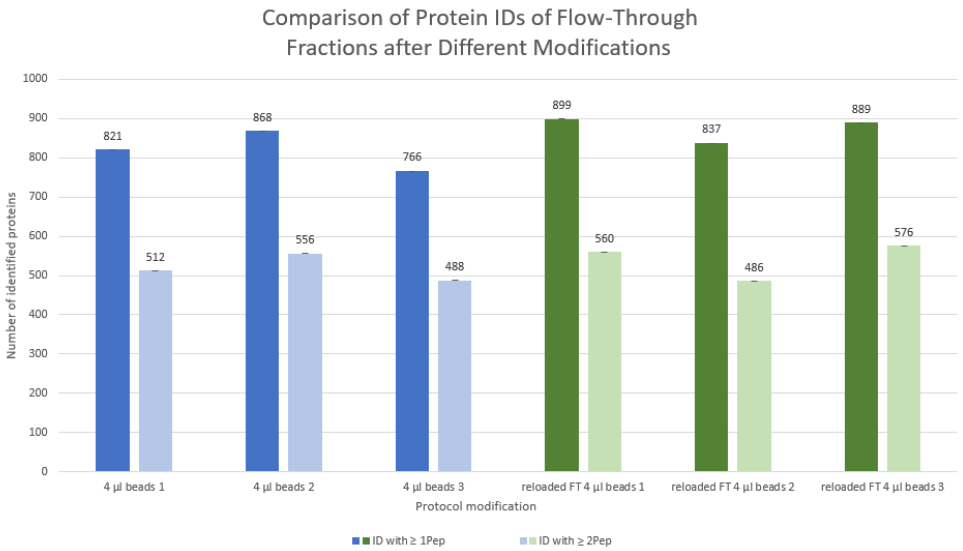

Figure 22: *Bar chart of protein IDs comparing the fractions of modified method using different paramagnetic bead amounts used to bind proteins (4 µl and 8 µl bead solution) with the 8 µl split by loading 4 µl onto the sample and 4 µl onto flow-through*

The abundance of the flow-through samples of the modified methods are plotted in Figure 23. Protein abundances of the replicates were roughly in the same range, both within the same group as well as in comparison to the others.

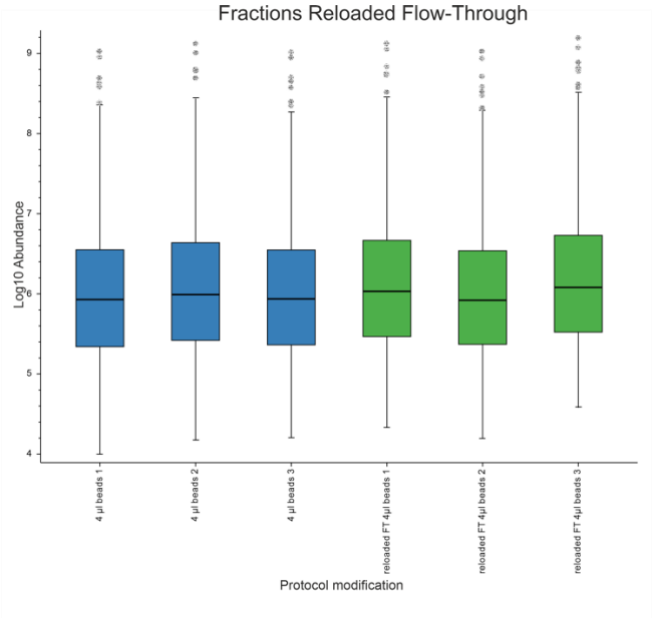

Figure 23: *Log(10)-transformed protein abundances of technical replicates in boxplot comparing fractions of modified method using different paramagnetic bead amounts used to bind proteins (4 µl and 8 µl bead solution) with the 8 µl protocol beads split by 4 µl for loading the sample and 4 µl for loading the flow-through*

4. Discussion

SP3 can be seen as a promising new protein digestion method for studying whole proteome of organisms including low abundant proteins. Most publications suggest, that SP3 is a superior and more reliable method even compared to established strategies like in-solution digestion and FASP, especially when analysing sub-microgram amounts (Moggridge *et al.*, 2018). It also offers a low per assay cost and rapid handling (Hughes *et al.*, 2014, 2019). Those and more properties give a good reason for using a standard processing approach in a core facility.

To establish SP3 Method, a comparison of the standard protein digestion methods in-solution and FASP using 2.5 µg yeast extract and in addition the same FASP protocol with 30 µg was undertaken. Results of this investigation showed, that in-solution and FASP protocol with 30 µg had a higher protein output than SP3, which is in contradiction with most data found (Hughes *et al.*, 2019). To establish the possibilities of SP3, several attempts were made to modify the protocol to improve the protein IDs and the protein abundances, which were used as a parameter for all experiments. Before modification, the standard protocol was used to rule out storage effects on the mixed bead solution containing different hydrophilic paramagnetic beads. Aside from one sample of the first day, most samples showed the same range in the numbers of protein IDs, which ruled out negative effects of storage at 4 °C.

Following modifications of the SP3 standard protocol were assessed:

- longer incubation on the magnetic rack to improve bead capture on the tube wall
- leaving out first washing step with ACN and 100 mM TEAB mixture
- increasing bead amount in first binding step
- loading the binding supernatant to fresh beads again and combining it with the original sample

The results show, that leaving out one washing step and increasing the bead amount turned out to be the most promising modifications. Figure 13 shows the results after reduction of the wash step compared to the standard protocol. It shows more reproducibility and higher protein IDs, which could have been induced by less sample handling. Another possibility would be, that the proteins normally lost when rinsed with the washing solution were kept on the beads.

Figure **16** shows usage of higher bead amounts with a skipped washing step already integrated into the protocol. Even after leaving out the outlier of the standard protocol, the mean number of protein IDs of the modified protocol with higher bead amount showed more identifications. The outlier could be led back to insufficient sample preparation. Longer incubation on the magnetic rack had slightly better protein identifications (Figure 11), however, the difference was not significant and rapidness of the method decreases, when increasing the time on the magnetic rack from 1 minute to 15 minutes. This makes the modification not worth the additional time. An alternative would be increasing contact time between beads and protein during an extended shaking phase.

Reloading flow-through to fresh beads increased sample handling which might have led to the lower protein IDs. Although in theory the proteins lost in the flow-through as seen in Figure 15 should have been recovered. In general, when looking at the flow-through fractions throughout the different protocols it is distinctive (Figure 15, Figure 18 and Figure 22), that all varied in the same range of around 800 protein IDs with at least one identified peptide per protein. This matches with the observations made by Sielaff *et al.* (2017), where the supernatant analysed by SDS-PAGE also showed protein loss until rinsing the beads with EtOH and ACN. Authors suggest, that working in neutral rather than low pH helps minimising protein loss into the supernatant (Sielaff *et al.*, 2017). Those ideal conditions were already integrated in the applied protocol but nevertheless protein loss during the binding step was observed.

Even after modifying the protocol, SP3 deliver lower IDs, up when compared to in-solution with the same protein amount. SP3 offers rapid handling and unlike in-solution digestion sufficient removal of unwanted cellular components or solvents like detergents, which might interfere with digestion or LC-MS analysis. Although FASP also achieves good contaminant removal, SP3 is more efficient when handling low protein amounts. That is why SP3 could still be compatible as a standard protein digestion method. More improvement steps have to be made like changing the binding solution from mixture of ACN and 100 mM TEAB to EtOH and changing the volume of the binding solution (Hughes *et al.*, 2019).

This thesis shows, that SP3 still has the potential to be applied as a suitable alternative to FASP and in-solution when dealing with low protein amounts and detergent containing samples. However, there is adaption of the protocol steps needed to make SP3 reproducible, robust and more sensitive.

5. Summary

Single-pot, solid-phase-enhanced sample preparation (SP3) is a novel protein digestion method based on paramagnetic beads, claiming to be able to outperform other methods. In this thesis SP3 was compared to in-solution digestion and FASP. Yeast proteins were digested using a combination of Trypsin and LysC with the different protocols. Then the peptides were purified with C18 columns before analysis by nano HPLC coupled to an Orbitrap mass spectrometer. To compare the different protocols, protein IDs were chosen as an important parameter, but also other parameters like the protein abundances of the technical and LC-MS replicates were taken into consideration. Although being described as a superior protocol, especially in sub-microgram analyses. In this study however, SP3 did not show the appropriate performance compared to in-solution digestion protocol.

Therefore, different parameters and steps of the sample preparation were investigated in detail. First, the influence of the storage time of bead mixture was tested (fresh, 7 and 14 days old). Storage of beads for 14 days prior to use did not show any strong influence on the number of identified proteins. Furthermore, the loading conditions were modified: higher bead amount; increase of incubation time on the magnetic rack; variation of the number of washing steps as well as reloading the flow-through on fresh magnetic beads in order to minimise the protein loss during the sample loading step.

Omitting the washing step under loading conditions and using higher bead concentration, not only enabled reproducible numbers of protein IDs and enhanced the protein output but also reduced sample handling resulting in faster preparation.

Despite improved protein identifications, protein loss still was higher than with in-solution digestion. Further prospects would be changing the protein binding conditions by using different solvents like ethanol. This thesis led to the conclusion, that SP3 still may have advantages for example when handling samples containing detergents and dealing with low available protein amounts but still the protocol needs further improvement.

Zusammenfassung

SP3 ist eine neue Proteinverdaumethode basierend auf der Wirkung paramagnetischer Kügelchen, diese soll vergleichbare Methoden übertreffen können. In dieser Arbeit wurde Hefeextrakt mit einer Enzymmischung aus Trypsin und LysC verdaut und dabei SP3 mit dem Verdau in Lösung und bzw. Ultrafiltrationseinheiten vergleichen. Als Parameter für die Effizienz der Probenaufarbeitung und des Proteinverdaus wurden die Zahlen der Proteinidentifizierungen und die Proteinabundanz der Proben verwendet. Obwohl SP3 als überlegenes Protokoll, speziell im Sub-Mikrogrammbereich, beschrieben wird, zeigen die Ergebnisse, dass SP3 durch Verdau in Lösung und mit Ultrafiltrationseinheiten übertroffen wurde.

Hierbei wurden unterschiedliche Schritte des Originalprotokolls zur Probenaufarbeitung modifiziert und analysiert. Als erster Schritt wurde die Haltbarkeit der Kügelchen in Lösung überprüft, indem derselbe Versuch mit Kügelchen aus derselben Lösung an drei unterschiedlichen Tagen durchgeführt wurde (frisch, nach 7 und 14 Tagen). Das Lagern der Kügelchen hatte keinen Einfluss auf die Leistung der Kügelchen beim Proteinverdau. Zusätzlich wurden die Beladungsbedingungen der Proteine auf die Kügelchen verändert, indem eine größere Menge an Kügelchen verwendet wurde. Um eine bessere Bindung an den Magneten zu ermöglichen, wurde die Anzahl der Wasch- und Reinigungsschritte modifiziert. Außerdem wurden die Proteine des Überstands nach dem Bindungsschritt auf frische Kügelchen geladen.

Das Auslassen des ersten Waschschritts mit dem Bindungsmedium und die Erhöhung der Menge an Kügelchen haben die zu erwartenden Ergebnisse gebracht. Die Wiederholbarkeit Anzahl der Identifizierungen konnten verbessert werden und die Arbeitsschritte haben sich verringert, was die Praktikabilität und Arbeitszeit des Protokolls verbessert hat.

Obwohl eine Steigerung der Effizienz zu beobachten war, erreicht auch das modifizierte Protokoll nicht die Leistungen des Verdaus in Lösung. Das führt zu dem Schluss, dass SP3 zwar Vorteile wie die Eliminierung von unerwünschten Bestandteilen und der niedrigen benötigten Proteinmenge mit sich bringt, jedoch immer noch Verbesserung nötig sind.

6. List of Abbreviations

SP3	single-pot, solid-phase-enhanced sample preparation
2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
ACN	acetonitrile
AGC	automatic gain control
CAA	chloroacetamide
CHAPS	3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate
DTT	dithiothreitol
ESI	electrospray ionisation
FASP	filter-aided sample preparation
FDR	false discovery rate
HCD	higher-energy collisional dissociation
HPLC	high performance liquid chromatography
IAA	iodoacetamide
InSol	in-solution digestion
LC	liquid chromatography
LC-MS	liquid chromatography–mass spectrometry
MALDI-MS	matrix-assisted laser desorption/ionisation-mass spectrometry
MIT	maximum injection time
MS	mass spectrometry
NCE	normalised collision energy
PBS	phosphate buffered saline
PD	proteome discoverer
PEG	polyethylenglycol

PSM	peptide-spectrum matches
PTM	post-translational modification
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOP	standard operating procedure
SPRI	solid-phase reversible immobilisation
TCEP	Tris-(2-carboxyethyl)-phosphin hydrochloride
TEAB	triethylammonium bicarbonate
TFA	trifluoroacetic acid

7. References

Aebersold, R. and Mann, M. (2003) 'Mass spectrometry-based proteomics', *Nature*. Nature Publishing Group, pp. 198–207. doi: 10.1038/nature01511.

Alpert, A. J. (1990) 'Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds', *Journal of Chromatography A*, 499(C), pp. 177–196. doi: 10.1016/S0021-9673(00)96972-3.

Aslam, B. *et al.* (2017) 'Proteomics: Technologies and their applications', *Journal of Chromatographic Science*. Oxford Academic, pp. 182–196. doi: 10.1093/chromsci/bmw167.

Cho, W. C. S. (2007) 'Proteomics Technologies and Challenges', *Genomics, Proteomics and Bioinformatics*. Elsevier, pp. 77–85. doi: 10.1016/S1672-0229(07)60018-7.

Deangelis, M. M., Wang, D. G. and Hawkins, T. L. (1995) 'Solid-phase reversible immobilization for the isolation of PCR products', *Nucleic Acids Research*, 23(22), pp. 4742–4743. doi: 10.1093/nar/23.22.4742.

Domon, B. and Aebersold, R. (2006) 'Mass spectrometry and protein analysis', *Science*. American Association for the Advancement of Science, pp. 212–217. doi: 10.1126/science.1124619.

Erde, J., Loo, R. R. O. and Loo, J. A. (2017) 'Improving proteome coverage and sample recovery with enhanced FASP (eFASP) for quantitative proteomic experiments', *Methods in Molecular Biology*, 1550, pp. 11–18. doi: 10.1007/978-1-4939-6747-6_2.

Hughes, C. S. *et al.* (2014) 'Ultrasensitive proteome analysis using paramagnetic bead technology', *Molecular Systems Biology*, 10(10), p. 757. doi: 10.15252/msb.20145625.

Hughes, C. S. *et al.* (2019) 'Single-pot, solid-phase-enhanced sample preparation for proteomics experiments', *Nature Protocols*, 14(1), pp. 68–85. doi: 10.1038/s41596-018-0082-x.

Issaq, H. J. and Veenstra, T. D. (2008) 'Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): Advances and perspectives', *BioTechniques*. Biotechniques, pp. 697–700. doi: 10.2144/000112823.

Konigsberg, W. H. and Henderson, L. (1983) 'Removal of Sodium Dodecyl Sulfate from Proteins by Ion-Pair Extraction', *Methods in Enzymology*, 91(C), pp. 254–259. doi: 10.1016/S0076-6879(83)91022-4.

León, I. R. *et al.* (2013) 'Quantitative Assessment of in-solution Digestion Efficiency Identifies Optimal Protocols for Unbiased Protein Analysis', *Molecular & Cellular Proteomics : MCP*, 12(10), p. 2992. doi: 10.1074/MCP.M112.025585.

Medzihradszky, K. F. (2005) 'in-solution digestion of proteins for mass spectrometry', *Methods in Enzymology*. Academic Press Inc., pp. 50–65. doi: 10.1016/S0076-6879(05)05003-2.

Moggridge, S. *et al.* (2018) 'Extending the Compatibility of the SP3 Paramagnetic Bead Processing Approach for Proteomics', *Journal of Proteome Research*, 17(4), pp. 1730–1740. doi: 10.1021/acs.jproteome.7b00913.

Nagaraj, N. *et al.* (2008) 'Detergent-based but gel-free method allows identification of several hundred membrane proteins in single LC-MS runs', *Journal of Proteome Research*, 7(11), pp. 5028–5032. doi: 10.1021/pr800412j.

Patterson, S. D. and Aebersold, R. H. (2003) 'Proteomics: The first decade and beyond', *Nature Genetics*. Nature Publishing Group, pp. 311–323. doi: 10.1038/ng1106.

Sielaff, M. *et al.* (2017) 'Evaluation of FASP, SP3, and iST Protocols for Proteomic Sample Preparation in the Low Microgram Range', *Journal of Proteome Research*, 16(11), pp. 4060–4072. doi: 10.1021/acs.jproteome.7b00433.

Wiśniewski, J. R. *et al.* (2009) 'Universal sample preparation method for proteome analysis', *Nature Methods*, 6(5), pp. 359–362. doi: 10.1038/nmeth.1322.

Yandamuri, N. *et al.* (2013) 'Comparative study of new trends in HPLC: A review', *International Journal of Pharmaceutical Sciences Review and Research*, 23(2), pp. 52–57.

Zhang, Y. *et al.* (2013) ‘Protein analysis by shotgun/bottom-up proteomics’, *Chemical Reviews*. American Chemical Society, pp. 2343–2394. doi: 10.1021/cr3003533.

Zotou, A. (2012) ‘An overview of recent advances in HPLC instrumentation’, *Central European Journal of Chemistry*, 10(3), pp. 554–569. doi: 10.2478/s11532-011-0161-0.

8. List of Figures

9. List of Tables