Journal of Pharmaceutical and Biomedical Analysis 56 (2011) 1069–1078

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Critical assessment of accelerating trypsination methods

Hanne Kolsrud Hustoft[a], Leon Reubsaet[b], Tyge Greibrokk[a], Elsa Lundanes[a], Helle Malerod[a,*]

[a] Department of Chemistry, University of Oslo, PO Box 1033, Blindern, NO-0315 Oslo, Norway
[b] School of Pharmacy, University of Oslo, PO Box 1068, Blindern, NO-0316 Oslo, Norway

ARTICLE INFO

Article history:
Received 28 June 2011
Received in revised form 3 August 2011
Accepted 4 August 2011
Available online 11 August 2011

Keywords:
Tryptic digestion
Accelerating digestion methods
Efficiency of protein identification
LC–MS based proteomics

ABSTRACT

In LC–MS based proteomics, several accelerating trypsination methods have been introduced in order to speed up the protein digestion, which is often considered a bottleneck. Traditionally and most commonly, due to sample heterogeneity, overnight digestion at 37 °C is performed in order to digest both easily and more resistant proteins. High efficiency protein identification is important in proteomics, hours with LC–MS/MS analysis is needless if the majority of the proteins are not digested. Based on preliminary experiments utilizing some of the suggested accelerating methods, the question of whether accelerating digestion methods really provide the same protein identification efficiency as the overnight digestion was asked. In the present study we have evaluated four different accelerating trypsination methods (infrared (IR) and microwave assisted, solvent aided and immobilized trypsination). The methods were compared with conventional digestion at 37 °C in the same time range using a four protein mixture. Sequence coverage and peak area of intact proteins were used for the comparison. The accelerating methods were able to digest the proteins, but none of the methods appeared to be more efficient than the conventional digestion method at 37 °C. The conventional method at 37 °C is easy to perform using commercially available instrumentation and appears to be the digestion method to use. The digestion time in targeted proteomics can be optimized for each protein, while in comprehensive proteomics the digestion time should be extended due to sample heterogeneity and influence of other proteins present. Recommendations regarding optimizing and evaluating the tryptic digestion for both targeted and comprehensive proteomics are given, and a digestion method suitable as the first method for newcomers in comprehensive proteomics is suggested.

1. Introduction

Proteomics is defined as the large-scale study of proteins in particular for their structures and functions [1], and their investigation has become very important since they are the main components of the physiological metabolic pathways of eukaryotic cells. Proteomics is increasingly important in areas like cancer prevention, biomarker discovery, food safety, toxicology, protein interaction studies, medical diagnostics and treatment [2]. Most proteomics studies today involve the use of a MS for identification and possibly quantification of proteins. Peptides are more easily ionized in the MS and easier to keep in solution than proteins, hence, the proteins are usually enzymatically digested to get peptides. Due to sample complexity, separation is necessary, and in proteomics traditionally two different approaches are followed, gel-based or the in-solution workflow. When using the gel-based approach, the proteins are separated in one or two dimensions (1D/2D) dependent on the sample complexity and the digestion is then performed in-gel [3]. In the in-solution based workflow, proteins can be separated with LC prior to the digestion to reduce the sample complexity. Additionally, the proteolytic peptides are subsequently separated with multidimensional (MD) LC–MS methods [2,4,5]. The in-solution based approach tends to be the simplest in terms of sample handling and speed, but on the other hand it requires sophisticated LC–MS equipment that requires constant maintenance. The focus in this study will be on the in-solution approach.

Trypsin is most commonly used for proteolytic cleavage of proteins. It has a well defined specificity; it hydrolyzes only the peptide bonds in which the carbonyl group is contributed by either an arginine (Arg) or a lysine (Lys) residue, except when they are bound to proline (Pro). Other enzymes like chymotrypsin, pepsin, Lys-C, Asp-N and Glu-C can also be applied for protein digestion. However, since trypsin is most commonly used, only trypsin is included in this study.

Prior to the in-solution digestion most protein samples need to be denaturated, reduced and alkylated in order for the enzyme to be able to cleave the proteins to peptides. The conventional tryptic digestion method is rather time consuming and is normally carried out overnight (12–16 h) at 37 °C, but digestion times up to 24 h have been reported due to sample heterogeneity and the presence

* Corresponding author. Tel.: +47 228 55573; fax: +47 228 55441.
E-mail address: hellema@kjemi.uio.no (H. Malerod).

0731-7085/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2011.08.013

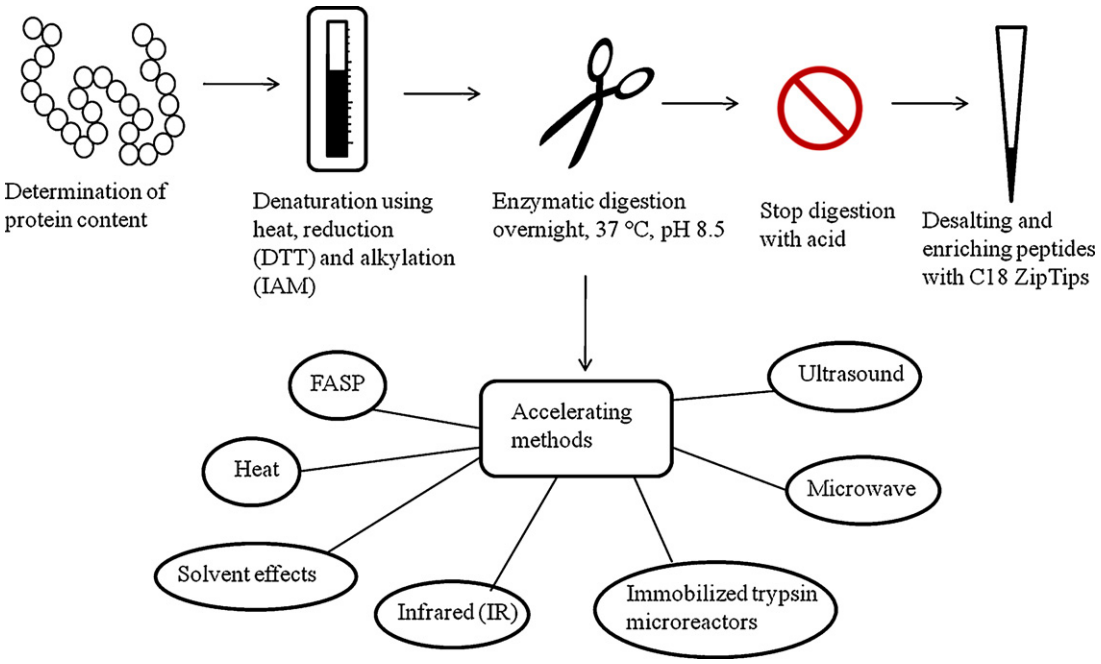

Fig. 1. The procedure of a classical workflow with various accelerating methods for in-solution based tryptic digestion.

of both easily and more resistant digestible proteins [3,6,7]. Protein digestion is therefore often considered as the bottleneck in proteomics. Accelerating methods like for instance heating [8], solvent effects [9–11], ultrasonic energy [12], infrared (IR) assisted [13–15], microwave assisted [16–20], and microreactors where the trypsin is immobilized on a solid support [21–24] have been reported. A classical sample preparation workflow in in-solution digestion with different accelerating methods is presented in Fig. 1.

The accelerating methods can be grouped into temperature related digestion (heating, IR and microwave assisted digestion), immobilized digestion (filter-aided sample preparation (FASP) and immobilized trypsin microreactors) and other ways (ultrasound and solvent effects) to accelerate the protein digestion. Based on preliminary in-house experiments we considered it necessary to ask the following questions; *can* these methods digest proteins completely within minutes as reported? And *can* these methods replace the conventional digestion method and reduce the time for sample preparation, even in comprehensive proteomics?

The aim of this study was to evaluate the efficiency of protein identification of the accelerated potential of IR and microwave assisted digestion, as well as solvent aided and FASP digestion. The methods were compared to the conventional digestion method at 37 °C in the same time range. The evaluation was based on comparing sequence coverage (SQ) and the area of intact protein peak. There is no straightforward procedure to describe the efficiency of protein digestion and act of trypsin, and hence an evaluation of the most used parameters is included. Based on these studies, we present a recommended procedure for how to optimize the digestion as well as a digestion method suited as a first approach for newcomers in comprehensive proteomics.

2. Materials and methods

2.1. Materials and reagents

HPLC grade acetonitrile (ACN, VWR, Radnor, PA, USA), type 1 water from an ultrapure water purification system (Millipore Corporation, Billercia, MA, USA), formic acid (FA, 50%, Fluka, by Sigma–Aldrich, St. Louis, MO, USA) and trifluoroacetic acid (TFA, Sigma–Aldrich) were used to prepare the mobile phases. 1.0 M triethylammonium bicarbonate (tABC) (pH 8.5, Sigma–Aldrich), $CaCl_2$ (dehydrated, Fluka Chemicals by Sigma–Aldrich) and 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris, Life Technologies Inc., Gaithersburg, MD, USA) were used to adjust the pH prior to protein digestion. A *4-protein mixture* was used to evaluate the different methods for protein digestion. The mixture consisted of bovine serum albumin (BSA), lysozyme (chicken egg white), cytochrome-C (cyto-C, bovine heart) and α-lactalbumin (α-lact, bovine), all obtained from Sigma–Aldrich. To prepare the stock solution of the protein mix, 5 mg of each of the solid protein was dissolved in 20 mL of water, resulting in a concentration of 0.25 mg/mL of each protein. When preparing the standard protein solutions, protein precipitation was experienced when the proteins were dissolved in 1.0 M tABC (pH 8.5) buffer prior to denaturation and reduction. The proteins were on the other hand easily dissolved in water prior to buffer addition, which is required for tryptic digestion. Health risks for the chemicals used are given in the Supplementary data Table S-1.

2.2. Protein denaturation, reduction and alkylation

DL-dithiothreitol (DTT) (Sigma–Aldrich) was added in a $1_{DTT} + 50_{protein}$ (w/w) ratio and the protein solution was incubated at 95 °C for 15 min, followed by cooling to room temperature and subsequent alkylation using iodoacetamide (IAM, Sigma–Aldrich) in the dark for 15 min at room temperature. IAM was added in a $1_{IAM} + 10_{protein}$ (w/w) ratio. The protein solution (1.0 mg/mL, 0.25 mg/mL of each protein) was subsequently divided into 100 μL (100 μg protein in total) aliquots in low protein bind eppendorf tubes and frozen at −20 °C until use.

2.3. Tryptic digestion

Prior to tryptic digestion the samples (100 μL) were made alkaline by adding tABC, to a final concentration of 50 mM (pH 8.5), and the proteins were digested using five different methods; the conventional method at 37 °C, IR, microwave energy, solvent aided and with the FASP method. For all methods, TPCK treated trypsin

(Sigma–Aldrich) was applied to cleave the proteins and 5 μL 50% FA was used to stop the protein digestion.

2.3.1. The conventional tryptic digestion method

Trypsin was added to the reduced and alkylated proteins at four different enzyme to protein (E + P) ratios; 1 + 5, 1 + 10, 1 + 20 and 1 + 40 (w/w). The proteins were allowed to digest at 37 °C at different times 5, 10, 60, 120, 180, 240, 300 min and overnight using a Thermoshaker® from Grant Instruments Ltd. (Cambridge, UK).

2.3.2. IR assisted protein digestion

Trypsin was added in a 1 + 20 E + P ratio (w/w) and placed in an in-house made IR oven, according to Wang et al. [13]. The incubation time for protein digestion ranged from 5 to 300 min (5, 6, 7, 8, 9, 10, 60, 120, 180, 240 and 300 min). The temperature in the solution was monitored during the IR assisted protein digestion using a thermometer.

2.3.3. Microwave assisted protein digestion

For the microwave assisted protein digestion the experimental set-up was based on the paper by Pramanik et al. [16], with some experimental and instrumental changes. Trypsin was added to 350 μL of the 4-protein mixture (1 mg/mL; 0.25 mg/mL of each protein) and cyto-C (250 ng/mL) in a 1 + 25 E + P (w/w) ratio and the eppendorf tubes were placed in a closed vessel microwave digestion oven (Ethos 1600, Milestone Inc. Shelton, CT, USA) for 10 min. The power of the microwave oven was set to 144 W and the temperature was held at either 37 or 45 °C. 17.5 μL 50% FA was added to stop the protein digestion. Digestion on the Thermoshaker using the same time (10 min) and temperature range (37, 45 and 60 °C) served as control.

2.3.4. Solvent effects

An amount of 40 or 80 μL ACN was added to the protein mixture (totally 100 μg protein in 100 μL) prior to reduction, alkylation and protein digestion. The pH was adjusted using either 50 mM tABC, 50 mM tABC in combination with 50 mM Tris and 10 mM $CaCl_2$, or 50 mM Tris with 10 mM $CaCl_2$. Trypsin was added in a 1 + 20 E + P (w/w) ratio and the proteins were digested for 1 h at 37 °C using the Thermoshaker. 4-protein mixture samples pH-adjusted with the same buffers, but without ACN, were used as control. After terminating the digestion with 5 μL 50% FA, all protein digests were evaporated to almost dryness using the SC 110 speedvac (Savant by Thermo Scientific) and re-dissolved in 100 μL 0.1% FA.

2.3.5. Protein digestion using the FASP method

4 w/v % SDS (Sigma–Aldrich) was added to the 4-protein mixture, and the resulting solution was treated by the FASP method published by Wisniewski et al. [21]. The protocol, attached in the Supplementary data, was consistently followed using a 10 K filter (Millipore Cooperation), and the SDS was substituted with 8 M urea (Sigma–Aldrich). Tryptic digestion was performed while the proteins were attached to the filter using a 1 + 100 E + P (w/w) ratio, and the peptides were eluted using centrifugal power (13 krpm, 20 min) and the filter was finally rinsed by 0.5 M NaCl (Sigma–Aldrich). The samples were desalted using C18 ZipTips prior to the LC–MS analysis using a method described in Section 2.4.

2.4. Tryptic digest clean-up

For tryptic digest clean-up, ZipTips packed with 0.6 μL C18 resin (Millipore Corporation) were used. The ZipTips were activated with 30 μL (3× 10 μL) 100% ACN and equilibrated with 30 μL (3× 10 μL) 0.1% FA. The sample (10 × 10 μL, totally 100 μL) was subsequently loaded on the ZipTip by aspirating and dispensing the sample up and down. The ZipTips were cleaned with 30 μL (3× 10 μL) 0.1% TFA, before the peptides were eluted with 10 μL ACN/0.1% FA (75/25, v/v). The eluate was evaporated to almost dryness and diluted with 50 μL 0.1% FA prior to the LC–MS/MS analysis.

2.5. Reversed phase (RP) LC–IT/Orbitrap-MS/MS

For the Ion-trap (IT) MS/MS analysis an ACE-C18 (1 × 150 mm, 3 μm, 100 Å) column (Advanced Chromatography Technologies, Aberdeen, Scotland, UK) was used. The LC system consisted of an 1100 series autosampler (G1313A) and a cap LC pump (G1376A) with a degasser incorporated (G1379A) (Agilent Technologies, Santa Clara, CA, USA). Mobile phase A contained 0.1% FA in type 1 water, and mobile phase B consisted of ACN + 0.1% FA. A 90 min gradient was used to separate the tryptic peptides. The gradient started at 5% B and increased linearly up to 45% B within 90 min. Subsequently, the gradient increased up to 75% B in 10 min and was kept at 75% B for 4.5 min. The gradient returned to the starting conditions in 0.1 min and the column was re-equilibrated for 20 min. The total analysis time was 125 min. The flow rate was set to 48 μL/min and the injection volume was 5 or 10 μL. Chemstation (Agilent technologies) was used to control the LC-system. An Esquire 3000 + IT-MS (Bruker Daltonics, Bremen, Germany) with capillary ESI source in positive ion-mode was used to detect the peptides in a scan range between 250 and 1700 Da. The resolution was set to 13,000 m/z/s, the target value was 20,000 ions and the maximum accumulation time was 200 ms. The capillary voltage and the end plate offset were set to 3000 and 500 V, respectively, while the nebulizer, drying gas and temperature were 30 psi, 5 L/min and 350 °C, respectively. Signal intensity threshold for data-dependent acquisition mode (DDA) was set to 2000. The three most abundant ions were chosen at the time for MS/MS to ensure complete compound information. EsquireControl Version 5.1 (Bruker Daltonics) was used to operate the IT-MS.

For the RP LC - Orbitrap MS/MS analysis, 10 μL of the sample was loaded on a Thermo Biosphere C8 (1 × 50 mm, 5 μm, 300 Å) column (Thermo Fisher Scientific Inc., Waltham, MA, USA) using the Ultimate 3000 pump and RS autosampler (Dionex by Thermo Fisher Scientific). The same mobile phases were used as listed above, and a 60 min gradient was used. The gradient started at 5% B and increased linearly to 45% B in 60 min, prior to an increase to 100% in 10 min and kept at 100% B for 4.5 min, before returning to the starting conditions during 0.1 min. The mobile phase composition was kept at 5% B for 20 min for column re-equilibration. The total analysis time was 94.5 min. The flow rate was 50 μL/min. The LTQ Discovery Orbitrap MS (Thermo Fischer Scientific Inc.) with capillary ESI source was operated in positive ion-mode in a scan range from 100 to 2000 Da. MS/MS data were acquired from 0 to 94.5 min and the signal-to-noise (S/N) threshold was set to 3. Collision induced dissociation (CID) was used for MS/MS. Chromeleon Xpress (Dionex Coperation) and Xcalibur 2.0.7 (Thermo Fisher Scientific Inc.) were used to control the LC system and the LTQ Discovery Orbtitrap MS, respectively.

2.6. Data analysis

All IT-MS data were analyzed and processed by Data Analysis and BioTools 2.2 from Bruker Daltonics. Mascot search engine (Matrix Science, http://www.matrixscience.com) was used for protein identification by comparing the MS/MS spectra with the SwissProt database. Carbamidomethylation of cysteine and oxidation of methionine were used as fixed and variable modifications, respectively. The peptide and MS/MS tolerance was set to 1.0 and 0.2 Da. Trypsin was chosen as the enzyme and number of missed cleavages was between 0 and 2. The charge state was specified to 2+ and 3+.

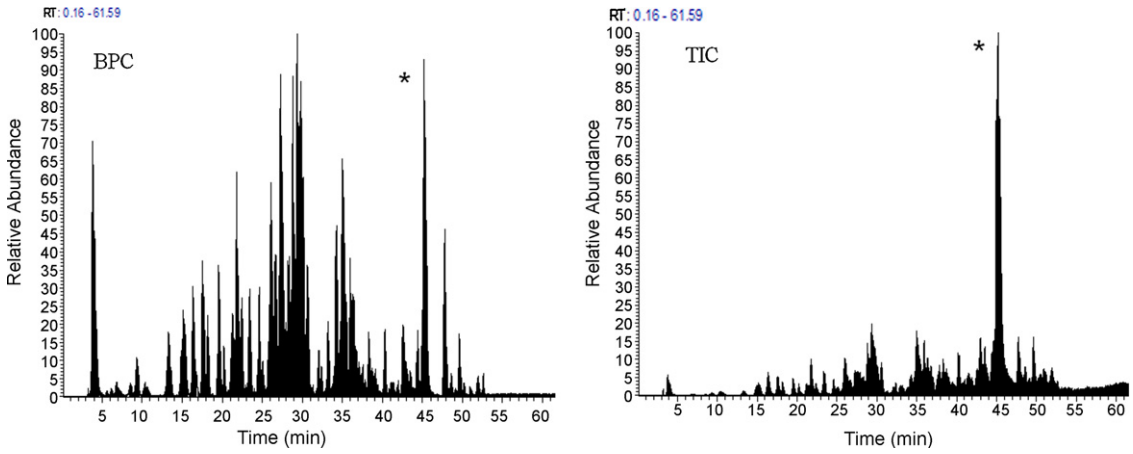

Fig. 2. BPC and TIC of the 4-protein mixture (totally 100 μg) when digested 5 min at 37 °C using a E + P ratio of 1 + 40. The Thermo Biosphere C8 (1 × 50 mm, 5 μm, 300 Å) column and the 60 min gradient were used to separate the peptides. The flow rate was set to 50 μL/min and the injection volume was 10 μL. The Orbitrap-MS was used for detection. The intact protein peak is marked with *.

The Orbitrap-MS data were processed by Proteome Discoverer 1.0 (Thermo Fisher Scientific Inc.) using the Sequest search engine and the SwissProt database. Trypsin was chosen as enzyme, and the maximum missed cleavage was between 0 and 2 when submitting the MS/MS spectra. Precursor mass criterion and tolerance were set to the *same measured mass-to-charge* and 10 ppm, respectively. Fragment mass tolerance was 0.8 Da and methionine oxidation and cysteine carbamidomethylation were set as dynamic side chain modifications. Furthermore, against decoy database search and target FDR *strict* and *relaxed* parameters were set to 0.01 and 0.05, respectively.

3. Results and discussion

Since tryptic digestion, which traditionally has been carried out overnight, is considered as a bottleneck in proteomics, several alternative accelerating methods have been suggested. The efficiency and robustness of these methods however need to be evaluated before being used in proteomics which requires efficient, repeatable and reliable digestion of all proteins present in a sample. When comparing accelerating methods with the conventional method it is of importance to use the exact same conditions, excluding only the accelerating factor in the control sample in order to have an accurate evaluation. Erroneous comparisons are done when comparing overnight digestion with accelerated digestion method carried out within e.g. 5 min [8,13,14,17,18]. Many proteins are easily digested and can provide high sequence coverage (SQ) after 5 min similar to the coverage obtained after overnight digestion. Hence, the accelerating method would appear to be more efficient and timesaving than it actually is by using easily digestible proteins for testing. The aim of this study was to evaluate and compare the accelerated digestion methods with the conventional digestion method at 37 °C.

3.1. Optimizing the tryptic digestion

3.1.1. Description of trypsin action

The efficiency of protein identification is often evaluated by the SQ [25,26], which is easily calculated by search engines like Mascot and Proteome Discoverer in a database search. The SQ might however, be a misleading parameter to use since not all proteins will have a high SQ due to being large in size and more difficult to digest. Additionally, SQ depends on the mass spectrometer as well as the parameters used in the database search. SQ is related to the number of missed cleavages. Incomplete digestion results in generation of longer peptides with possible cleavage sites that are not cleaved. These sites are referred to as the number of missed cleavages, and can be set between 0 and 9, but the default is 2. The SQ is normally higher when the number of missed cleavages is high, and the SQ with no missed cleavages represents a more complete digestion than at a higher number of missed cleavages. When using Mascot as search engine for protein identification, the number of missed cleavage is reported together with the identified peptides, and the distribution of 0, 1 or 2 missed cleavages can give an indication of the protein identification efficiency.

The trypsin action can also be evaluated by the mass chromatograms using both the base peak chromatogram (BPC) and total ion chromatogram (TIC), to both check if peptides are generated [8] and the presence of any intact protein [15,20]. As an example, using the 4-protein mixture, the BPC and TIC from a 5 min digestion at 37 °C are shown in Fig. 2. Many peptides were generated, reflected by the number of peaks, but an intact protein peak (highlighted with *) was additionally observed in both mass chromatograms. The intact protein peak was more abundant in the TIC than in the BPC, due to TIC representing the summed intensities of a mass range, while the BPC only presents the most intense ions.

The intact protein peak was identified to be undigested cyto-C. Cyto-C is a rather small protein (~12 kDa) and more easily detected in the mass range of the MS. The other proteins are too large (≥14 kDa) to be detected by MS instruments like the Orbitrap-MS. A MALDI MS would have overcome this problem and detected the other proteins, but was not available for this study.

Due to the accessibility of two different MS instruments some of the samples were analyzed using both the IT-MS and the Orbitrap-MS. The TICs of the undigested cyto-C peak is apparently much more abundant when using the Orbitrap-MS compared to the IT-MS, as seen in Fig. S-1. This is probably because the Orbitrap-MS is more sensitive than the IT-MS.

As seen, evaluation of protein digestion can be difficult since it is depended of the nature of the proteins, parameters in a database search and the MS instrument. Studying the digestion in comprehensive proteomics is more challenging than in targeted proteomics, where the E + P ratio and the digestion time can be optimized for each protein. We suggest using an easily digested protein like cyto-C to be added to the protein sample as a "digestion marker". If intact cyto-C is detected, then the digestion can be considered as insufficient. It is however important that the protein originate from another species than that under investigation

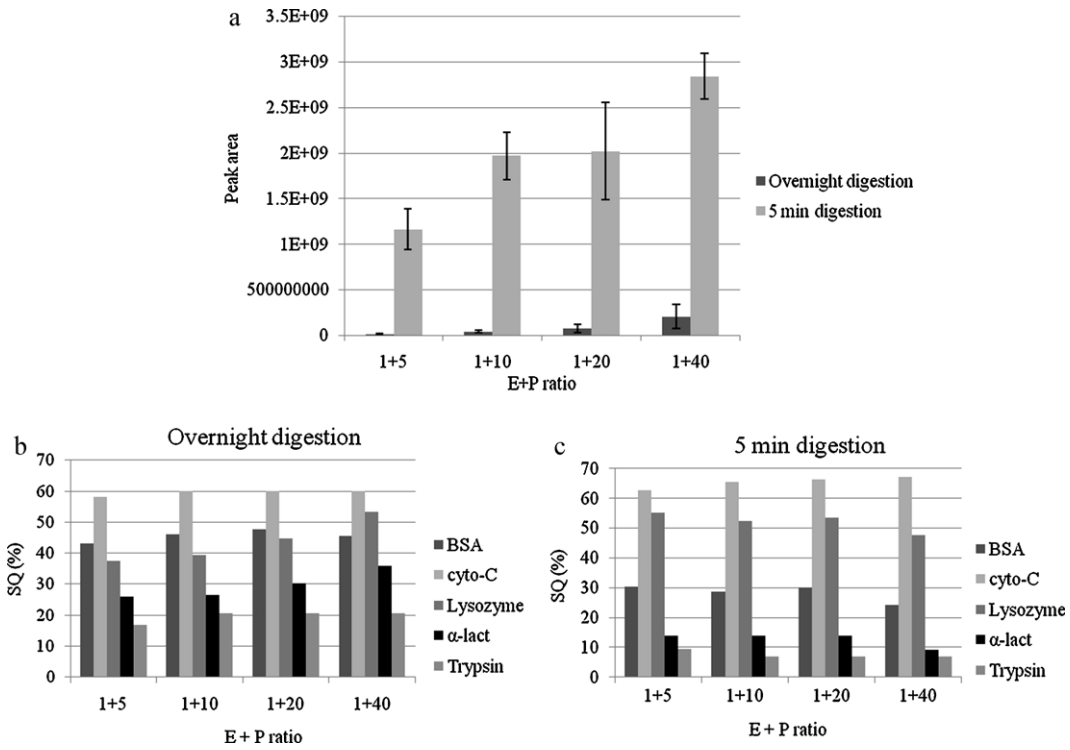

Fig. 3. The area of intact protein peak (a) and SQ (b and c) of the intact cyto-C peak when digesting the 4-protein mixture overnight and 5 min at 37 °C using four different E + P ratios; 1 + 5, 1 + 10, 1 + 20 and 1 + 40. The Biosphere C8 column, a 60 min gradient and the Orbitrap-MS were used to separate and detect the tryptic digested peptides. Other conditions were as in Fig. 2. The mass tolerance was 0.8 Da and the precursor tolerance was 10 ppm. The number of missed cleavages was 2.

to avoid interferences in protein identification and quantification. Alternatively, other non-protein reagents that can be cleaved by enzymes can be used as the "digestion marker." In this study, the efficiency was evaluated by the SQ when the number of missed cleavages altered from 0 to 2, and by the area of the intact cyto-C peak.

3.1.2. Enzyme to protein (E + P) ratio

An optimal enzyme to protein ratio is crucial for an efficient protein digestion. The relative concentration of enzyme needs to be high, but not so high that autolysis occurs [27]. The 4-protein mixture was digested using an enzyme to protein (E + P) ratio between 1 + 5 and 1 + 40 (w/w) and the digestion time was set to 5 min and 12–16 h (overnight) at 37 °C on the Thermoshaker. The rather fast digestion of 5 min was included in order to have the optimal E + P conditions for the accelerated digestion methods as well, which are often carried out in about 5 min range. The area of the intact cyto-C peak, for the 5 min and overnight digestion, is shown in Fig. 3a.

As seen from Fig. 3a, a higher undigested cyto-C peak was detected when the 4-protein mixture was digested for only 5 min compared to the overnight digestion. The E + P ratio influenced the digestion for both digestion times in the same manner. More protein was digested using the 1 + 5 ratio, while the largest undigested peak was observed with the 1 + 40 ratio.

As previously mentioned, autolysis may occur when using a high enzyme concentration. The SQ of the proteins and trypsin is therefore another important parameter when determining the optimal enzyme to protein ratio. The SQ of the model proteins using the four different E + P ratios are shown in Fig. 3b and c. A higher SQ was found for BSA and α-lactalbumin when the digestion time was 12–16 h (overnight), indicating that the protein identification efficiency is strongly dependent of the digestion time. Cyto-C and lysozyme on the other hand, were identified with almost similar SQ at both digestion times.

When determining the most optimal E + P ratio, both SQ and the intact protein peak were taken into consideration. The 1 + 40 ratio was not selected due to the presence of a larger intact protein peak compared to the other ratios. The 1 + 5 ratio was also not selected due to the potential increased risk for autolysis. The SQ of trypsin did not vary much when altering the E + P ratio, but the highest SQ for the 5 min digestion was observed for the 1 + 5 ratio. Based on this, the 1 + 20 ratio was chosen as the optimal ratio since BSA, lysozyme and α-lactalbumin were all detected with a higher SQ compared to the 1 + 10 ratio. Furthermore, the difference in area of the intact protein peak was not significant for the 1 + 10 and 1 + 20 ratio for both digestion times. Hence, the E + P ratio used in this study was 1 + 20 if not otherwise is mentioned.

3.1.3. Digestion time

As Capelo et al. among others claim, the digestion step is the main bottleneck regarding speed of the traditional workflow of protein identification, and for complex samples it may take as much as 24 h due to heterogeneity [3,6,7].

The digestion time was varied from 5 min to 5 h in order to explore if digestion could be carried out faster than 12–16 h (overnight). The intact protein peak of cyto-C, which decreased with time, indicated sufficient digestion after 2–4 h (Fig. 5 in Section 3.3.1). The SQ did not vary significantly when using longer digestion times. BSA and cyto-C were identified with the highest SQ at 5 h and 10 min, respectively, and α-lactalbumin and lysozyme at 1 and 2 h (control samples in Fig. S-2 in Supplementary data).

There was little difference in peak area of the intact protein and SQ when the digestion time was 2 and 5 h, and these four model proteins can therefore be digested within 2 h. Hence, many proteins can be digested within a shorter time period than overnight (12–16 h). Proc et al. performed a study indicating a digestion time of 9 h for a complex protein sample of 45 proteins of which some were easy and some difficult to digest [28]. For targeted proteomics, the

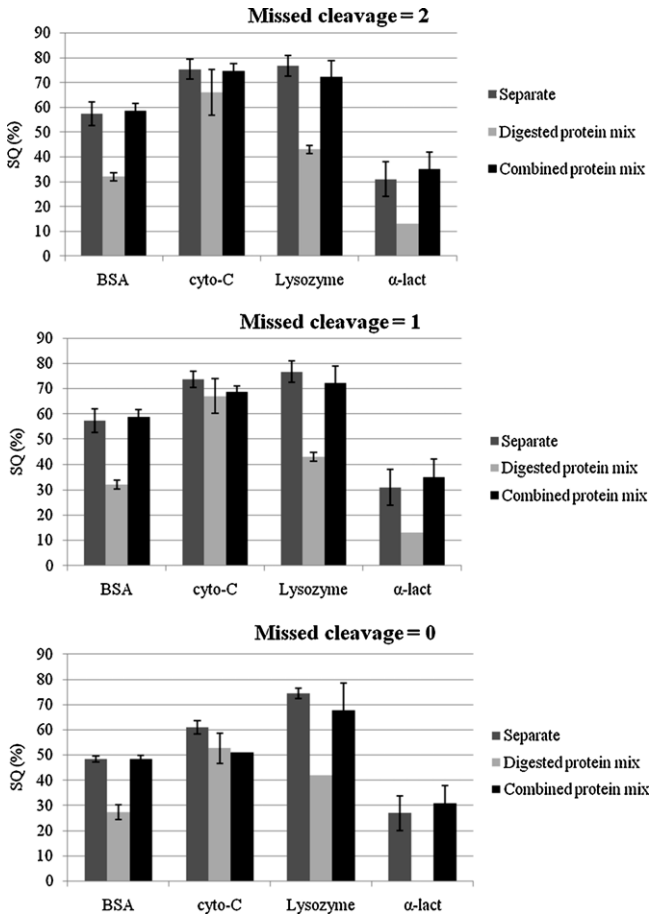

Fig. 4. The SQ of the four model proteins (25 μg of each protein) digested overnight at 37 °C separately, in a mixture and separately digested in a combined protein mixture ($n = 3$). The E + P ratio was 1 + 20. The peptides were separated on the ACE-C18 column using the 90 min gradient and the IT-MS/MS was used for identification. The injection volume was 10 μL. Number of missed cleavages was altered from 0 to 2 and peptide and MS/MS tolerance were set to 1.0 and 0.2 Da, respectively. α-Lactalbumin was only identified in two out of three sample replicates in the protein mixture, and a standard deviation was therefore not calculated.

needed digestion time should be determined during method development. In comprehensive proteomics on the other hand where the protein heterogeneity may be large, the digestion time should be sufficiently long to digest the majority of all proteins.

3.2. Influence of other proteins

Proteins may act differently in various environments such as in a complex matrix, or protein mixture compared to that of a single protein in solution. The influence of other proteins was evaluated using the four model proteins, digested both as single proteins and in a mixture. An additional mixture was prepared by combining the tryptic digested peptides from all four proteins, digested separately. The average SQ with the corresponding standard deviation of three sample replicates of the four proteins is shown in Fig. 4. The number of missed cleavage was altered from 0 to 2.

As seen from Fig. 4, the SQ was lower when the proteins were digested in a mixture compared to when they were separately digested, for all number of missed cleavages. The lower SQ could be due to lower peptide signal because of ion suppression in the MS, or that the proteins have less access to trypsin in a mixture resulting in less cleavage. However, since similar SQ was obtained for the combined protein mixture and the separately digested proteins, ion suppression is not the limitation, and hence the reduced SQ for the

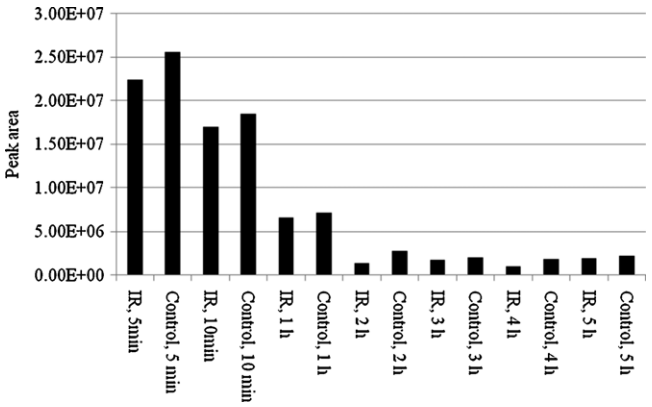

Fig. 5. The area of the intact protein peak after digesting totally 100 μg (25 μg of each protein) of the 4-protein mixture for 5 and 10 min and 1–5 h in the IR oven and on the Thermoshaker (control) at 37 °C. The peptides were separated on the ACE-C18 column using the 90 min gradient and the IT-MS/MS was used for detection. Other conditions were the same as in Fig. 4.

protein mixture is due to less access to trypsin. This proves that the presence of other proteins has a large influence of the digestion and a long digestion time may be necessary in order to digest the majority of the proteins as in comprehensive proteomics.

3.3. Accelerated protein digestion methods

3.3.1. IR assisted digestion

3.3.1.1. Digestion time. The in-house made IR oven based on the designs by Wang et al. [13–15] was used for IR assisted protein digestion of the 4-protein mixture. Conventional digestion on the Thermoshaker at 37 °C served as control. The temperature in the solution was monitored during the IR assisted protein digestion (Fig. S-3 in Supplementary data) and the temperature for three sample replicates was found to stabilize at 37 °C after approximately 2 min, and remained stable at 37 °C for at least 10 min. The protein mixture was digested for 5 min, the same time as used by Bao et al. [15]. The digestion time was further increased to 10 min, since the oven used 2 min to stabilize the temperature at 37 °C. Hence, the solution was at 37 °C for 3 and 8 min, respectively. The SQ was overall the same for all samples examined, whether they were incubated for 5 or 10 min. The digestion time was further increased, up to 5 h (1–5 h), in order to investigate the IR assisted digestion at an extended incubation time. The area of the intact protein peak decreased when increasing the digestion time using both the Thermoshaker and IR oven as seen in Fig. 5 and the SQ in Fig. S-2 in Supplementary data.

Fig. 5 shows that a digestion was accomplished between 2 and 4 h. Only a small decrease in both intact protein peak and SQ was revealed when comparing IR assisted digestion to digestion in the Thermoshaker. The overall difference between the IR assisted digestion and traditional digestion at 37 °C was very small both when considering the SQ (Fig. S-2 in Supplementary data) and the intact protein peak, using fast (5–10 min) as well as long (1–5 h) digestion time.

3.3.1.2. Concentration levels. The concentration of the protein mixture was decreased in order to explore if the IR assisted digestion was more effective at a lower protein concentration level than 100 μg/100 μL. The protein mixture was diluted with type 1 water to 20–60 μg/100 μL (0.20–0.60 μg/μL), and digested for 2 h in the IR oven and on the Thermoshaker at 37 °C with a 1 + 20 E + P ratio. The tryptic peptides were subsequently separated by RP LC-IT-MS/MS. The peak area relative to protein concent ration remained similar as shown in Fig. S-4a, with the exception of the 0.20 μg/μL 4-protein

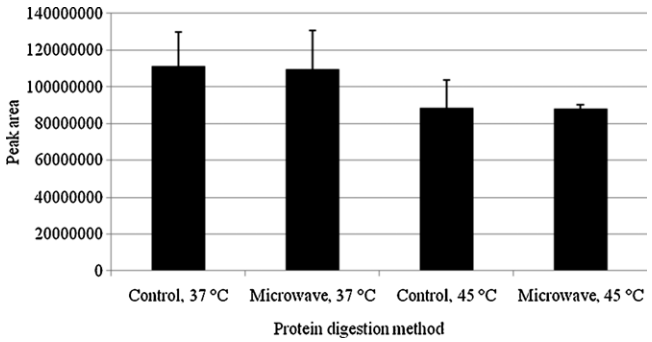

Fig. 6. The average peak area ($n = 3$) of the intact protein peak represented by cyto-C with the corresponding standard deviation when digesting the 4-protein mixture in the microwave and on the Thermoshaker at 37 and 45 °C. The ACE-C18 column and the 90 min gradient were used to separate the peptides, while the IT-MS/MS was used for detection. Other conditions were the same as in Fig. 4.

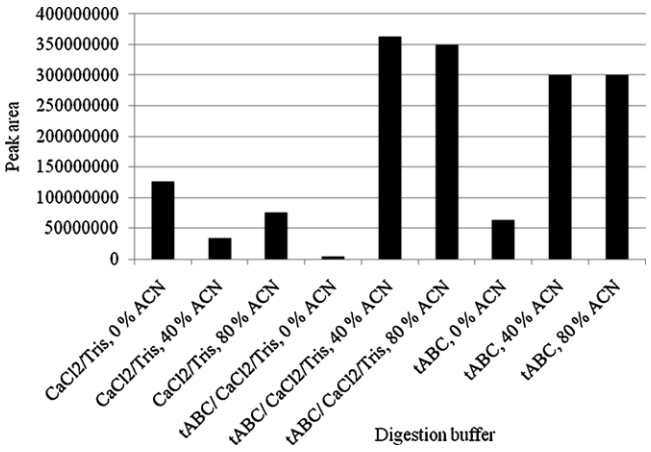

Fig. 7. The area of the intact protein peak (cyto-C) when digesting the four model proteins with either $CaCl_2$/Tris, tABC/$CaCl_2$/Tris or tABC in combination with 0, 40 or 80% ACN. The digestion time was 1 h at 37 °C. A 10 μL sample (10 μg digested protein) was loaded on the ACE C18 column and the peptides were separated using a 90 min gradient using the IT-MS/MS for detection. Other conditions were similar to Fig. 4.

mixture. The SQ decreased slightly with decreasing protein concentration, as seen from Fig. S-4b-d in Supplementary data, as the reaction rate of digestion is proportional to the protein concentration [10] and since some of the peptides will have an intensity lower than the limit of detection.

When comparing the SQ of the IR assisted digestion procedure with the traditional digestion at 37 °C, no significant differences were found, even at concentration levels down to 0.20 μg/μL. Thus, proteins can be digested in a IR oven, but compared to the conventional digestion procedure, there are no indications that the IR method provide an improved protein identification efficiency for 0.20 to 1.00 μg/μL protein (which are commonly employed) at digestion times from 5 min up to 5 h.

3.3.2. Microwave assisted digestion

Microwave assisted protein digestion was evaluated using a temperature controlled microwave oven. The 4-protein mixture and cyto-C were used as test samples. The power of the microwave was set to 144 W and the temperature was held at 37 and 45 °C, for 10 min as described in Section 2.3.3 in Experimental. Digestion at 37 and 45 °C on the Thermoshaker was used for control. The average SQ with the corresponding standard deviation from three sample replicates of the protein mixture is shown in Fig. S-5 in Supplementary data. No differences between microwave and temperature assisted protein digestion were observed for the four model proteins. The area of the cyto-C peak decreased with increasing temperature as seen in Fig. 6, but no significant difference in the area of the intact protein peak was found between the temperature and microwave initiated digestion.

The actual temperatures measured in the sample solution immediately after the microwave irradiation were approximately 45 and 60 °C, and hence deviated from the set temperatures of the instrument which were 37 and 45 °C, respectively. Digestion of cyto-C on the Thermoshaker at 60 °C was therefore also carried out and the average area of the intact protein peak, from two sample replicates is shown in Fig. S-6. The difference in peak area between the two sample replicates using the different digestion methods was between 10 (Thermoshaker, 45 °C) and 20% (microwave, 45 °C). The intact peak of cyto-C decreased slightly with increasing temperature, but no significant differences in SQ using the different methods or temperatures were found (data not shown).

Hence, microwave irradiation seems to have no additional effect on the tryptic digestion of the four model proteins when compared to the conventional digestion method. On the other hand, the increased digestion temperature apparently enhanced the action of trypsin giving a decrease in the intact peak of cyto-C. However, the BPC of four chosen peptides of cyto-C also showed a decrease in the peak area when increasing the digestion temperature, indicating that the action of trypsin decreased at higher temperatures using TPCK treated trypsin (data not shown). Nevertheless, a modified (like methylated) trypsin which is commercially available, that have an optimum digestion temperature at 60 °C, could be more efficient in combination with microwave assisted digestion [8]. In this study, however the TPCK treated trypsin was used since the aim was to repeat the experimental setup performed by Pramanik et al. [16].

3.3.3. Solvent aided digestion

Tris with $CaCl_2$ is the buffer that is commonly applied when the concentration of ACN is high during protein digestion, while tABC is mostly used at lower ACN concentration level. It should be kept in mind that protein precipitation may occur at high concentration levels of ACN. The natural presence of Ca^{2+} in most samples, binds at the Ca^{2+}-binding loop of trypsin and prevents autolysis [29]. A concentration of 1 mM $CaCl_2$ is therefore recommended, but not always necessary if the contribution of Ca^{2+} from natural sources is high [30]. The area of the intact cyto-C peak after digestion at the various ACN concentrations in combination with different buffers are shown in Fig. 7, while the SQ of each protein is presented in Fig. S-7 in Supplementary data.

When the protein digestion was performed without ACN, the tABC/$CaCl_2$/Tris gave the most efficient digestion, with a very low undigested protein peak. However, when increasing the concentration of ACN to 40 and 80%, the $CaCl_2$/Tris buffer gave the best digestion. A larger undigested protein peak was observed when tABC was used in the solution. Only cyto-C was identified when using tABC in combination with 40 and 80% ACN as seen in Fig. S-7. There was a large difference in SQ when the number of missed cleavage varied between 0 and 2, indicating insufficient digestion. The difference in SQ when using the other buffer–ACN combinations were smaller, and hence the digestion appeared to be more efficient. The reason for the low trypsin protein identification efficiency at higher ACN concentration may be due to the lack of Ca^{2+}-ions. In contrast to Tris buffers, where $CaCl_2$ always is added, $CaCl_2$ contribution was low when the ratio of ACN was as high as 80%. Others speculate that poor digestion at high organic content in the sample solution may be due to increased protein precipitation because of interactions with ACN [31]. Protein precipitation was observed with tABC in combination with 40 and 80% ACN, and could have affected the digestion in the present study. For the four

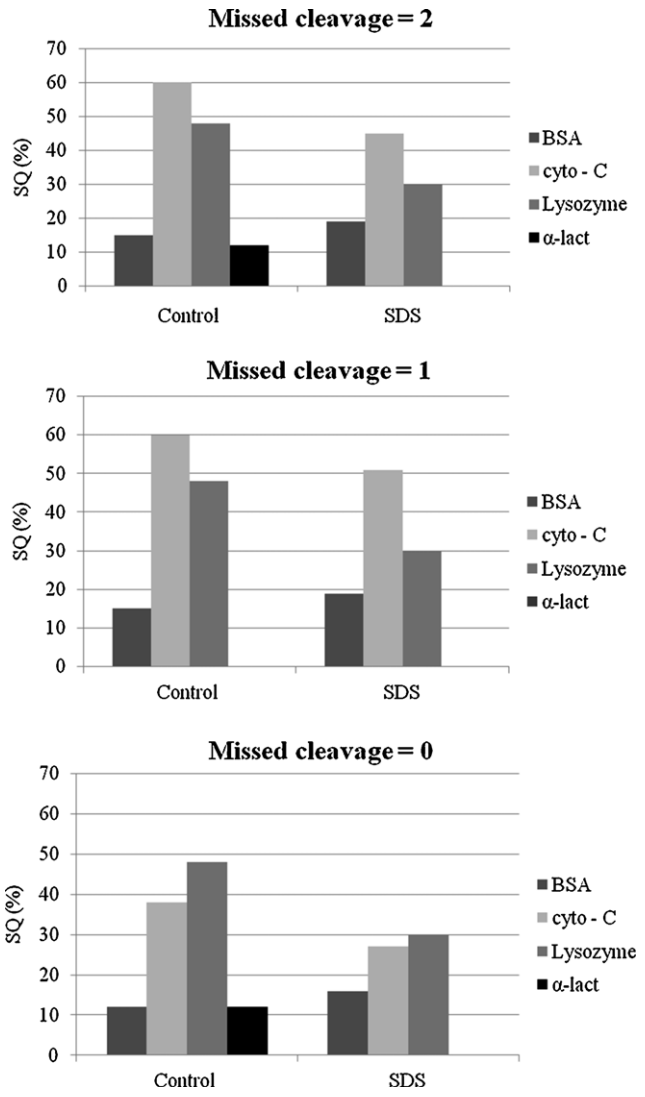

Fig. 8. 30 μg of the 4-protein mixture dissolved in an aqueous buffer (control) and a SDS containing buffer. The protein mixtures were digested on a 10 kDa spinfilter using the FASP method. The LC and MS conditions were as in Fig. 4.

model proteins addition of higher amount of organic solvent to the sample solution did not improve the digestion, as earlier found by Russell et al. [32]. However, adding $CaCl_2$/Tris to the tABC buffer decreased the intact protein peak considerably compared to using tABC solely. It is therefore recommended to include $CaCl_2$/Tris in the digestion buffer in order to have a more sufficient digestion. Furthermore, a $CaCl_2$/Tris buffer should be used when ACN is included in order to reduce the risk of protein precipitation and improve the efficiency of protein identification.

3.3.4. Filter-aided sample preparation (FASP) assisted digestion

According to the FASP procedure, total solubilisation of the proteome can be achieved using 4% SDS, which can subsequently be exchanged by urea on a 10 K filter device [21]. Digestion of the proteins on the filter is followed by elution of the generated peptides through centrifugal force. In the present study we wanted to verify the FASP procedure, both for its ability to deplete SDS and to digest the proteins. The 4-protein mixture was dissolved in an aqueous (control) and a SDS containing buffer, prior to SDS depletion and protein digestion on the 10 K filter. The average SQ with the corresponding standard deviation is shown in Fig. 8. α-Lactalbumin was not identified in the SDS containing samples. As seen from Fig. 8, lower SQ was obtained when the protein mixture was dissolved in the SDS containing buffer. The same trend was observed when number of missed cleavages was altered from 0 to 2.

The spinfilter was not able to deplete all SDS using the reported protocol, this can lead to problems with subsequent the LC–MS analysis because of a large SDS peak. Due to high abundance of SDS, peptides with lower abundance coeluting with SDS may be difficult to identify. Actually, Millipore's product information state that the 10 K Amicon filter devices used will not deplete more than 40% of SDS in 5% SDS solution, and this might be the reason for the lower SQ for the SDS dissolved samples.

The FASP procedure was found rather time consuming, but recently the FASP method was made commercially available through a FASP™ Protein Digestion Kit, from Protein Discover, and in this protocol some of the centrifugation steps are decreased in time, but still it takes more than 2 h to complete the protocol, prior to 4–18 h of trypsin digestion.

3.4. Tryptic digest clean-up

Purifying the tryptic peptides with C18 Ziptip is very useful in order to remove salts and other reagents that may interfere with the subsequent LC–MS/MS analysis. However, the capacity of the ZipTips needs to be considered prior to the sample loading. In this study, 0.6 μL C18 ZipTips were used and the capacity of these tips were determined to be 10 μg. A digestion of 10 μg protein of the four model proteins were ZipTipped prior to the LC–MS/MS analysis and compared with a non ZipTipped sample.

The more hydrophilic peptides, eluting early, are not identified in the ZipTipped sample, this may be due to washing the ZipTip with an aqueous mobile phase. Additionally, the overall intensity is also lower compared to the non ZipTipped sample, and this can be due to the limited capacity. Even though, more than 60% of all peptides were identified for all proteins. Hence, when needing to purify the tryptic digest prior to LC–MS analysis, ZipTips are still recommended since they are easy to handle and commercially available at a reasonable price.

4. Conclusions and recommended sample preparation procedure for LC–MS based proteomics

The protein identification efficiency of four potential accelerated digestion methods using IR and microwave assisted, solvent aided and the FASP protocol was evaluated. These methods were compared to the conventional digestion method at 37 °C using the same conditions, excluding the accelerating factor. The four model proteins were digested using all methods, but none appeared to be more efficient than the conventional digestion method at 37 °C in the same time range as the accelerating methods. This is in contrast with statements made in the original literature. These statements were mostly based on erroneous comparisons, by comparing fast digestion in the minutes range with overnight digestion, and the use of easily digestible proteins. Since the accelerated methods were not able to digest the four model proteins more efficiently, they would certainly not be more suitable for more complex protein samples. We therefore recommend using the conventional digestion method at 37 °C, also due to commercially available instrumentation and ease of handling.

Both sequence coverage and intact protein peak were used to evaluate the digestion. Sequence coverage is often the easiest parameter to use, but it is dependent of the nature of proteins, the MS instrument and the parameters used for the database search and it can therefore be difficult to use for comparing different digestion methods. There are limitations regarding the use of the intact protein peak as well, but small proteins like cyto-C or other non-protein reagents which may be cleaved by trypsin, can be used as a digestion marker in order to evaluate the action of trypsin.

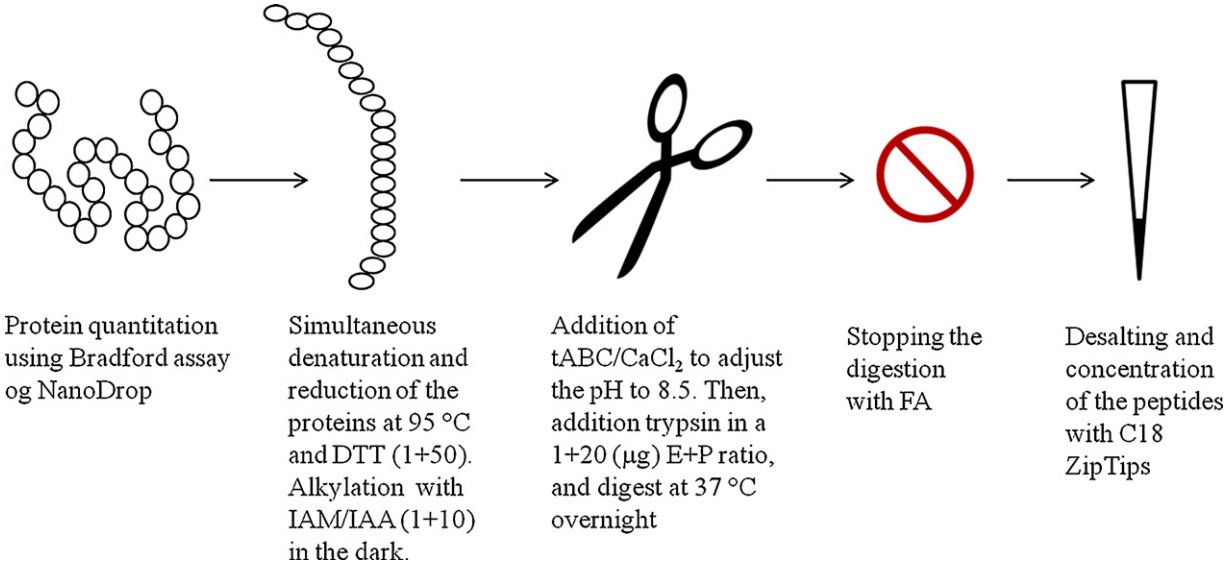

Fig. 9. A recommended procedure for in-solution comprehensive proteomics.

For targeted proteomics, the protein digestion can be optimized for each protein and the digestion can be carried out faster than 12–16 h (overnight) at 37 °C. In this case the optimal digestion method can be found varying both digestion times as well as enzyme to protein ratios. In comprehensive proteomics, on the other hand, it is more intricate to optimize the digestion due to sample heterogeneity and the large variety of proteins. Often a long digestion time is needed in order to cleave the majority of proteins to peptides due to the nature of proteins and influence of other proteins present. We suggest therefore especially for newcomers in the field of comprehensive proteomics to use the conventional overnight digestion method at 37 °C as shown in Fig. 9 in order to have a sufficient digestion for both easily and more resistant digestible proteins.

Acknowledgements

The Norwegian Research council is acknowledged for financial support (Project 197431/F20). The authors would also like to thank Steven R. Wilson for valuable discussions, Inge Mikalsen for making the IR oven and helping with different technological challenges, and Siri Olsen for contributing to the laboratory work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jpba.2011.08.013.

References

[1] N. Anderson, N. Anderson, Proteome and proteomics: new technologies, new concepts, and new words, Electrophoresis 19 (1998) 1853–1861.
[2] J.L. Capelo, R. Carreira, M. Diniz, L. Fernandes, M. Galesio, C. Lodeiro, H.M. Santos, G. Vale, Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques, Anal. Chim. Acta 650 (2009) 151–159.
[3] D. López-Ferrer, B. Cañas, J. Vázquez, C. Lodeiro, R. Rial-Otero, I. Moura, J.L. Capelo, Sample treatment for protein identification by mass spectrometry-based techniques, TrAC Trends Anal. Chem. 25 (2006) 996–1005.
[4] H. Loftheim, T.D. Nguyen, H. Malerød, E. Lundanes, A. Åsberg, L. Reubsaet, 2-D hydrophilic interaction liquid chromatography-RP separation in urinary proteomics – minimizing variability through improved downstream workflow compatibility, J. Sep. Sci. 33 (2010) 864–872.
[5] H. Malerod, E. Lundanes, T. Greibrokk, Recent advances in on-line multidimensional liquid chromatography, Anal. Methods 2 (2010) 110–122.
[6] J. Capelo, R. Carreira, L. Fernandes, C. Lodeiro, H. Santos, J. Simal-Gandara, Latest developments in sample treatment for 18O-isotopic labeling for proteomics mass spectrometry-based approaches: a critical review, Talanta 80 (2010) 1476–1486.
[7] J.M. Armenta, M. Perez, X. Yang, D. Shapiro, D. Reed, L. Tuli, C.V. Finkielstein, I.M. Lazar, Fast proteomic protocol for biomarker fingerprinting in cancerous cells, J. Chromatogr. A 1217 (2010) 2862–2870.
[8] J. Havliš, H. Thomas, M. Šebela, A. Shevchenko, Fast-response proteomics by accelerated in-gel digestion of proteins, Anal. Chem. 75 (2003) 1300–1306.
[9] M.N. Gupta, I. Roy, Enzymes in organic media, Eur. J. Biochem. 271 (2004) 2575–2583.
[10] M.B. Strader, D.L. Tabb, W.J. Hervey, C. Pan, G.B. Hurst, Efficient and specific trypsin digestion of microgram to nanogram quantities of proteins in organic–aqueous solvent systems, Anal. Chem. 78 (2005) 125–134.
[11] W. Hervey IV, M. Strader, G. Hurst, Comparison of digestion protocols for microgram quantities of enriched protein samples, J. Proteome Res. 6 (2007) 3054–3061.
[12] Y.-P. Li, Y.-Y. Li, X.-W. Zhou, C.-H. Chen, S.-L. Zhang, P.-Y. Yang, Ultrasound-assisted urea-free chemical denaturants combined with thermal denaturation to accelerate enzymatic proteolysis, Fenxi Huaxue 38 (2010) 663–667.
[13] S. Wang, H. Bao, L. Zhang, P. Yang, G. Chen, Infrared-assisted on-plate proteolysis for MALDI-TOF-MS peptide mapping, Anal. Chem. 80 (2008) 5640–5647.
[14] H. Bao, T. Liu, X. Chen, G. Chen, Efficient in-gel proteolysis accelerated by infrared radiation for protein identification, J. Proteome Res. 7 (2008) 5339–5344.
[15] H. Bao, Infrared-assisted proteolysis using trypsin-immobilized silica microspheres for peptide mapping, Proteomics 9 (2009) 1114–1117.
[16] B.N. Pramanik, U.A. Mirza, Y.H. Ing, Y.H. Liu, P.L. Bartner, P.C. Weber, A.K. Bose, Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: a new approach to protein digestion in minutes, Protein Sci. 11 (2002) 2676–2687.
[17] S. Lin, Z. Lin, G. Yao, C. Deng, P. Yang, X. Zhang, Development of microwave-assisted protein digestion based on trypsin-immobilized magnetic microspheres for highly efficient proteolysis followed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis, Rapid Commun. Mass Spectrom. 21 (2007) 3910–3918.
[18] S. Lin, G. Yao, D. Qi, Y. Li, C. Deng, P. Yang, X. Zhang, Fast and efficient proteolysis by microwave-assisted protein digestion using trypsin-immobilized magnetic silica microspheres, Anal. Chem. 80 (2008) 3655–3665.
[19] J.R. Lill, E.S. Ingle, P.S. Liu, V. Pham, W.N. Sandoval, Microwave-assisted proteomics, Mass Spectrom. Rev. 26 (2007) 657–671.
[20] P.M. Reddy, W.-Y. Hsu, J.-F. Hu, Y.-P. Ho, Digestion completeness of microwave-assisted and conventional trypsin-catalyzed reactions, J. Am. Soc. Mass Spectrom. 21 (2010) 421–424.
[21] J. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nat. Meth. 6 (2009) 359–362.
[22] J. Krenkova, N.A. Lacher, F. Svec, Highly efficient enzyme reactors containing trypsin and endoproteinase lysc immobilized on porous polymer monolith coupled to MS suitable for analysis of antibodies, Anal. Chem. 81 (2009) 2004–2012.
[23] H. Yamaguchi, M. Miyazaki, T. Honda, M.P. Briones-Nagata, K. Arima, H. Maeda, Rapid and efficient proteolysis for proteomic analysis by protease-immobilized microreactor, Electrophoresis 30 (2009) 3257–3264.

[24] H. Yuan, Y. Zhou, L. Zhang, Z. Liang, Y. Zhang, Integrated protein analysis platform based on column switch recycling size exclusion chromatography, microenzymatic reactor and [mu]RPLC-ESI-MS/MS, J. Chromatogr. A 1216 (2009) 7478–7482.
[25] F. Xu, W.-H. Wang, Y.-J. Tan, M.L. Bruening, Facile trypsin immobilization in polymeric membranes for rapid, efficient protein digestion, Anal. Chem. (2010).
[26] A. Shevchenko, Evaluation of the efficiency of in-gel digestion of proteins by peptide isotopic labeling and MALDI mass spectrometry, Anal. Bioanal. Chem. 296 (2001) 279–283.
[27] M. Ye, S. Hu, R.M. Schoenherr, N.J. Dovichi, On-line protein digestion and peptide mapping by capillary electrophoresis with post-column labeling for laser-induced fluorescence detection, Electrophoresis 25 (2004) 1319–1326.
[28] J.L. Proc, M.A. Kuzyk, D.B. Hardie, J. Yang, D.S. Smith, A.M. Jackson, C.E. Parker, C.H. Borchers, A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin, J. Proteome Res. 9 (2010) 5422–5437.
[29] F.F. Nord, M. Bier, L. Terminiello, On the mechanism of enzyme action. LXI. The self digestion of trypsin, calcium-trypsin and acetyltrypsin, Arch. Biochem. Biophys. 65 (1956) 120–131.
[30] University of Minnesota, Center for Mass Spectrometry and Proteomics, <http://www.cbs.umn.edu/msp/protocols.html> (Accessed date: 27 June 2011).
[31] H.M. Santos, C. Mota, C. Lodeiro, I. Moura, I. Isaac, J.L. Capelo, An improved clean sonoreactor-based method for protein identification by mass spectrometry-based techniques, Talanta 77 (2008) 870–875.
[32] W.K. Russell, Z.-Y. Park, D.H. Russell, Proteolysis in mixed organic–aqueous solvent systems: applications for peptide mass mapping using mass spectrometry, Anal. Chem. 73 (2001) 2682–2685.