Enzymatic Digestion of Proteins in Biological Samples for Quantification with LC-MS/MS

An Overview and Evaluation

Suzanne Willems (S3615154)
6-4-2021
Analytical Biochemistry
B. Sleumer & N. van de Merbel

Abstract

The enzymatic digestion of proteins in biological samples is an important step for the quantification with LC-MS/MS and is presented here in an overview. The enzymes used for this are classified as proteases and work by the hydrolysis of the peptide bonds. The most commonly used enzymes are trypsin, chymotrypsin, Glu-C, Lys-C, Arg-C, Asp-N and pepsin, each with their own cleavage sites. The digestion step can be optimized by adjusting several factors: the pH, the temperature, the enzyme:protein ratio and the way the enzyme is used. An evaluation of these enzymes was performed by theoretically digesting hGH 1 (22 kDa) and showed that trypsin resulted in peptides with an optimal length (7-20 amino acids) and charge (at least two positive charges) and had a relatively low price with a wide availability. In general, all the enzymes produced peptides with SSRC (hydrophobicity) scores that lie around the useful range of 10 to 45. Furthermore, a demonstration of signature peptide selection was performed after trypsin digestion on two isoforms of hGH with a mass of 22 kDa by taking several criteria into account: uniqueness and the absence of unstable amino acids. This resulted in two unique peptides for hGH 1 and three for hGH 2. Lastly, the importance of the use of internal standards is discussed. Stable-isotope labeled (SIL) forms of the protein cover for all the steps in the digestion, but may be hard to obtain; SIL forms of the signature peptide often are a good alternative. Next to enzymatic digestion, additional sample cleanup may be necessary to obtain sufficient sensitivity: extraction of proteins from a biological sample can be performed with immunocapture, SPE and LLE; extraction of peptides can be performed with immunocapture and SPE.

Contents

1. Introduction

Quantification of proteins in biological samples by using LC-MS/MS has gained a lot of attention in recent years; due to its sensitivity, selectivity and the possibility to use it in both drug development and in a clinical setting [1–4]. Proteins have many functions, like hormones or antibodies, making them a great target for developing new drugs, which are called therapeutic proteins or biopharmaceuticals [5]. More biopharmaceutical proteins have reached the market and several methods for quantification have been developed, from 'top-down' quantification of intact proteins, 'middle-up' quantification of antibody fragments to 'bottom-up' quantification of enzymatically digested proteins [6,7].

'Bottom-up' quantification of proteins can be divided in seven factors [5] that the quantification relies on: the internal standard, extraction of the protein analyte from the sample, the enzymatic digestion, selection of the signature peptide(s), extraction of the signature peptide from the sample, and separation and quantification with liquid chromatography (LC) and tandem mass spectrometry (MS/MS). Usually, reversed-phase LC is performed, which separates the peptides into fractions based on their interaction with the nonpolar, hydrophobic stationary phase in the column. The more hydrophobic a peptide is, the more interactions are formed between the stationary phase and the peptide and the longer the retention time will be [8]. Next, the fractions obtained with LC are detected by tandem mass spectrometry in a triple quadrupole mass spectrometer, for example with selective reaction monitoring (SRM), based on their mass-to-charge (m/z) ratio. The fractions are first ionized with electrospray ionization (ESI) [3,8]. The ionized peptides then enter the first quadruple of the mass spectrometer, where a peptide with a specific m/z, called the precursor ion, can move to the second quadrupole. In the second quadrupole, the precursor ion is fragmented by using a gas, like N_2, into product ions. The product ions enter the third quadrupole, which again selects ions with specific m/z ratios to move to the mass detector [8]. The digestion step makes easy quantification of proteins by LC-MS/MS possible, considering compounds with a mass higher than 5000 Da cannot usually be directly quantified by LC-MS/MS [2]; it is therefore a crucial step in protein quantification. In this work, the focus will be on the enzymatic digestion of proteins and the steps before the actual quantification in the form of an overview and evaluation.

In the chapter 'Enzymes' several aspects will be discussed, starting with the classification of enzymes and the actual digestion of proteins. Furthermore, the most commonly used enzymes are presented and an evaluation is made, based on their performance of the digestion of the human growth hormone (hGH) isoform 1 as an example. In the following chapter 'Selection of Peptides', the results of the evaluation are used to illustrate how the process of selecting signature peptide(s) works, while taking several criteria into account. Lastly, the chapter 'Sample Preparation' will discuss the use of internal standards and extraction of proteins or peptides from biological samples.

2. Enzymes

2.1 Classification & Types

Enzymes are important tools in the analytical biochemistry field: they catalyze a wide variation of chemical reactions. By considering what kind of chemical reactions enzymes catalyze, the International Union of Biochemistry and Molecular Biology (IUBMB) created the IUBMB Enzyme Nomenclature List. Here, enzymes are classified in seven classes: oxidoreductases, transferases, hydrolases, lyases, isomerases, ligases and translocases [9].

The enzymes that are used for the digestion of proteins fall in the class of hydrolases and can be subclassified as proteases. The enzymes that will be discussed are endoproteases [10], which hydrolyze peptide bonds between amino acids at specific sites, thereby breaking down the protein. In other words, the polypeptide chain, the unfolded form of the protein, is degraded into peptides [11]. Several enzymes are frequently used for the digestion of proteins in biological samples, based on where they cut the protein. Examples of the most commonly used enzymes are trypsin, chymotrypsin, Lys-C, Glu-C, Arg-C, Asp-N and pepsin [3,5].

2.2 Protein Digestion

2.2.1 Working Mechanism

The digestion of proteins is based on the binding of a protein to an enzyme at its binding site. The binding site is mostly shaped as a cavity, in which amino acid side chain residues are exposed. The binding between the protein (ligand) and the enzyme is caused by the interactions between the amino acid residues of the enzyme and the exposed amino acid residues of the protein. These interactions are established by non-covalent bonds, such as electrostatic attractions, Van der Waals attractions, hydrogen bonds and hydrophobic forces. Considering non-covalent bonds are weak bonds, multiple bonds need to be formed between the enzyme and the protein to create binding that is sufficient for the digestion of the protein. This also indicates the importance of the protein fitting in the cavity of the enzyme for the formation of the bonds and the specificity of the enzyme for the protein. Logically, the fewer the bonds that can be formed between the enzyme and the protein, the less the protein fits in the cavity [11].

The reactivity of the surface of an enzyme can be enhanced by interactions between amino acid side chains that are located next to each other. One of these interactions is caused by the accumulation of closely located polar amino acid side chains. This is based on the repulsion that the polar, negatively charged, amino acid side chains have for each other, to strongly attract positively charged groups. Also, the formation of hydrogen bonds between amino acid side chains located close to each other, can lead to the activation of groups that would otherwise be unreactive [11]. For example, in serine proteases, due to hydrogen bond formations, the active site consists of the triad of aspartic acid, histidine and serine. Closely located to the histidine, the negatively charged aspartic acid side chain attracts a proton from the histidine side chain to form a hydrogen bond, which leads to histidine taking up the proton from serine with which it had formed a hydrogen bond. Serine will remain with a negative charge, making it reactive and enabling it to break or create covalent bonds with ligands [11,12].

As stated before, the enzymes used for the digestion of proteins are classified as hydrolases and proteases, meaning that these enzymes hydrolyze peptide bonds [11]. The proteases can also be classified again into four groups/families. Figure 1 shows the four different proteases and their active groups, which are serine proteases (1a), cysteine proteases (1b), aspartic proteases (1c) and metalloproteases (1d).
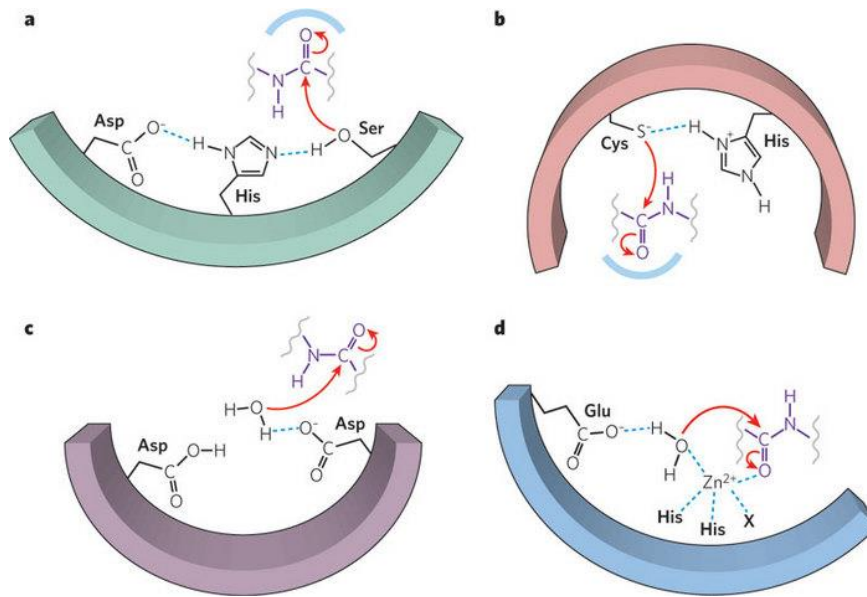
Figure 1: The four different proteases: serine proteases (1a), cysteine proteases (1b), aspartic proteases (1c) and metalloproteases (1d). The red arrow pointing from the active group to the peptide bond indicates the nucleophilic attack. The blue curved lines, shown by 1a and 1b, indicate the oxyanion hole [13].

Serine proteases and cysteine proteases contain groups in the active site with proton-withdrawing properties, making the amino acid residues of serine and cysteine reactive. This results in a nucleophilic attack of the reactive groups on the peptide bond to be hydrolyzed [13,14]. They both have a similar mechanism, with the only difference being the reactive amino acid side chains. Figure 2 shows the general mechanism of cleavage of peptide bonds by serine proteases. The other proteases use the same mechanism, only with different amino acids in the active sites that facilitate the nucleophilic attack. The first step in the mechanism is the enzyme coming in contact with a peptide and finding an amino acid with a specific side chain (table 1) that fits in the binding site (S1 pocket). Every enzyme has its own specific binding site, meaning that every enzyme will cleave after different amino acids (table 2). Trypsin, for example, has a binding site that is narrowly shaped, with the negatively charged amino acid residue of aspartic acid and the residue of serine located deepest inside the binding site. This two characteristics indicate binding of amino acids that have a long side chain with a positive charge, which are arginine and lysine [12,14,15].

As said before, serine is activated due to hydrogen bond formation between aspartic acid (Asp) and histidine (His) [11]. Binding of an amino acid in the binding site results in a small change of conformation in the enzyme, which is the negative side chain of asparagine getting closer to the imidazole ring of histidine. In the second step, the imidazole ring is being protonated by the hydroxyl-group of serine, which makes histidine acting as a basic catalyst. Due to the deprotonation of serine, the oxygen on the side chain of serine remains with a negative charge and facilitates a nucleophilic attack on the carbonyl carbon of the peptide bond. Because of the attack, a tetrahedral intermediate is formed and the carbonyl carbon gets a negative charge, making it unstable. Since oxygen is more electronegative than carbon, oxygen is more stable with a negative charge, thus oxygen takes up an electron pair from the double bond. Now with a negative charge, the oxygen is called an oxyanion. The oxyanion moves to a space within the active site, which is called the oxyanion hole. In the oxyanion hole, hydrogen bonds form between the oxyanion and the amide groups of serine and a closely located glycine (Gly), for even more stabilization of the tetrahedral intermediate [12,14–16].

The third step will cleave the amide group from the peptide bond. Because of the protonation of the imidazole ring in the first step, the ring has now become an acidic catalyst. One of the lone pairs on the oxyanion transfers back to form a double bond between the carbon and oxygen. Once again, the

5

carbon will have a negative charge and is unstable. This makes the peptide bond susceptible for protonation by the imidazole ring and the peptide bond is cleaved. In the fourth step, the amide group is departed from the enzyme. The side chain of the serine is now acylated with the remaining peptide and the intermediate is now called the acyl-enzyme intermediate. A transfer of electron pairs within the imidazole ring removes the positive charge of the nitrogen [12,14,15].

In the fifth step, water enters the active site of the enzyme and is used as a nucleophile for the hydrolysis of the peptide. Again, the histidine will act as a basic catalyst and the imidazole is protonated by the water molecule, thereby facilitating a nucleophilic hydroxyl group. The hydroxyl group attacks the carbon of the acyl-enzyme intermediate, resulting in the oxyanion. In step six, a nucleophilic attack and addition of the hydroxyl group forms a tetrahedral intermediate that is stabilized by the oxyanion in the oxyanion hole where it forms hydrogen bonds [12,14,15].

In step seven, one of the lone pairs on the oxyanion creating the negative charge, transfers back to form a double bond between oxygen and carbon. The transfer results in a negative charge on the carbon atom in the tetrahedral intermediate, making it unstable. This leads to breaking the bond between the peptide and serine through protonation by the imidazole ring of histidine. Within the imidazole ring, the electron lone pairs are transferred to remove the positive charge of the nitrogen. In step eight, the peptide is released and the protonation and electron transfer makes the active state ready for another peptide bond cleavage [12,14,15].
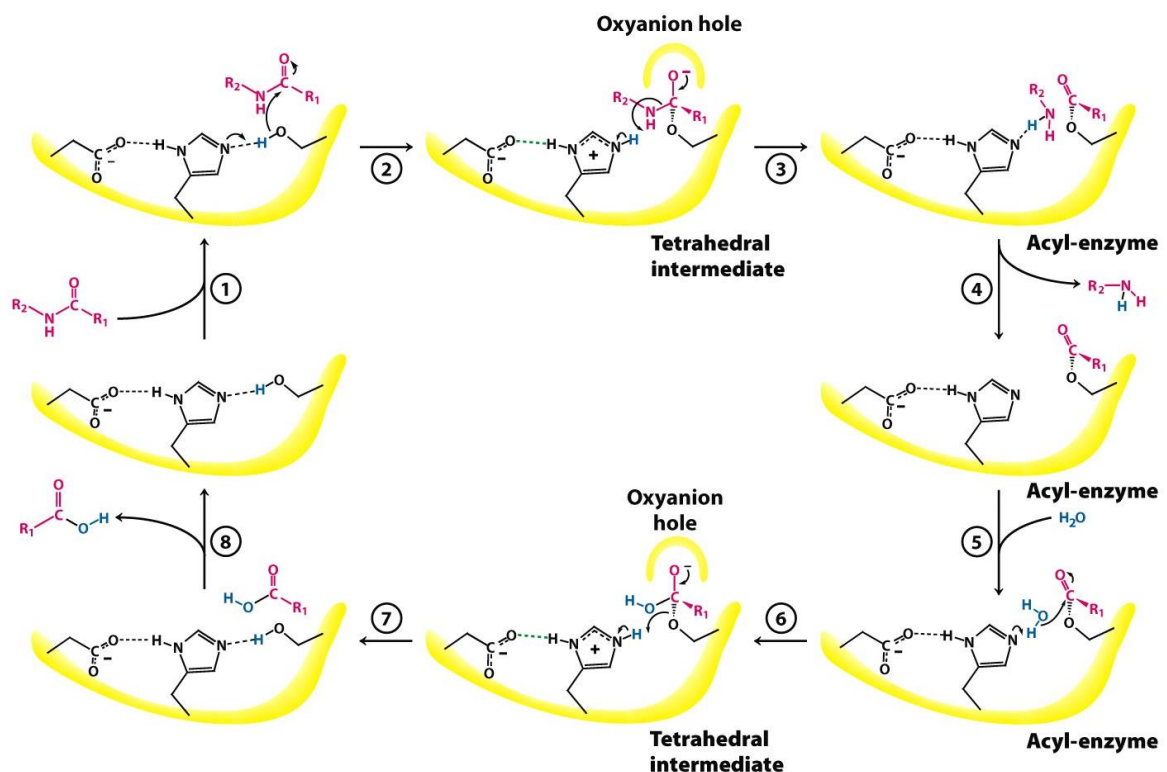
Figure 9.8
Biochemistry, **Seventh Edition**
© 2012 W. H. Freeman and Company

Figure 2: An example of a mechanism of the cleavage of a peptide bond by the active site of a serine protease [15].

Aspartic proteases use the negatively charged side chain of aspartic acid as the active site. The side chain forms a hydrogen bond with a water molecule, which is used as a nucleophile for the hydrolysis. Metalloproteases contain a metal ion, which is usually a zinc ion. The zinc ion is present in a complex with two positively charged, histidine amino acid side chains, a negatively charged amino acid side chain and a water molecule at the active site. Also here, the water molecule is used as the nucleophile for the hydrolysis [12,13].

Table 1: Overview of amino acids and the specificities of their side chains [11,17].
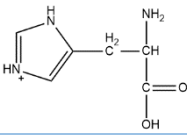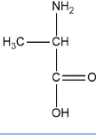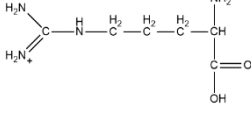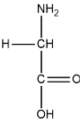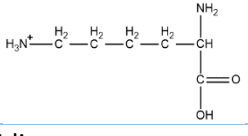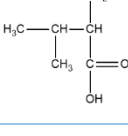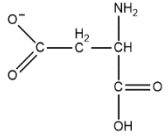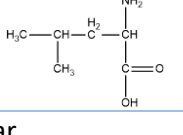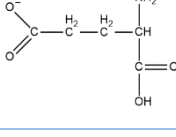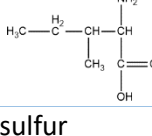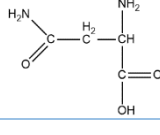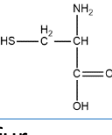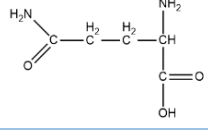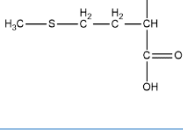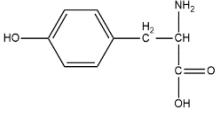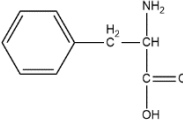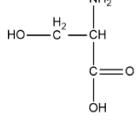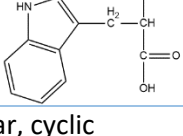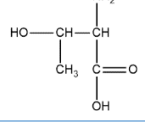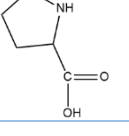
Hydrophilic		Hydrophobic	
Amino Acid	**Side Chain**	**Amino Acid**	**Side Chain**
Histidine (H)	Basic	Alanine (A)	Nonpolar
Arginine (R)	Basic	Glycine (G)	Nonpolar
Lysine (K)	Basic	Valine (V)	Nonpolar
Aspartic acid (D)	Acidic	Leucine (L)	Nonpolar
Glutamic acid (E)	Acidic	Isoleucine (I)	Nonpolar
Asparagine (N)	Polar -NH_2	Cysteine (C)	Nonpolar, sulfur
Glutamine (Q)	Polar -NH_2	Methionine (M)	Nonpolar, sulfur
Tyrosine (Y)	Polar -OH, aromatic	Phenylalanine (F)	Nonpolar, aromatic
Serine (S)	Polar -OH	Tryptophan (W)	Nonpolar, aromatic
Threonine (T)	Polar -OH	Proline (P)	Nonpolar, cyclic

Table 2: The most commonly used enzymes, the family they belong to, the cleavage site and exceptions to the cleavage, the optimal pH and the optimal temperature.

Enzyme[a]	Family[a]	Site of cleavage[a]	Exceptions	pH[a]	Temp. (°C)[b]
Trypsin	Serine protease	C-terminal R and K	Not when followed by P, R and K Not when followed by X* located next to phosphorylated S or T[c]	7-9	37
Chymotrypsin	Serine protease	C-terminal Y, W, L and F	Can cleave C-terminal M, but many missed cleavages[a]	7-9	25
Lys-C	Serine protease	C-terminal K	Not when followed by E, K or P[d]	7-9	37
Glu-C	Serine protease	C-terminal E and D	Only after D when present in phosphate buffer[e]	4-9	25
Arg-C	Cysteine protease	C-terminal R	Can cleave C-terminal of K, but less efficient[b]	7.6-7.9	37
Asp-N	Metalloprotease	N-terminal D	Can cleave N-terminal of E, when present in buffer containing detergents[b]	4-9	37
Pepsin	Aspartic protease	C-terminal F, Y, W[b] and L[f]	Can also cleave C-terminal of L, but this is pH dependent[b]	1-3[g]	37[g]

[a] [3,18], [b] [18], [c] [19–21], [d] [3,21], [e] [3,18,22], [f] [23], [g] [24]
* X can be any amino acid.

2.2.2 Denaturation, Reduction & Alkylation

Before a protein can be digested by an enzyme, it is a possibility to go through a few steps to reveal all the possible cleavage sites present in the protein. The first step is the denaturation of the protein, which can be performed with chemicals or by applying heat. By denaturation of the protein, the interactions that form the tertiary structure of the protein are broken, thereby unfolding the protein [12]. One of the most used denaturing agents is urea, with other common options being sodium dodecyl sulfate (SDS), organic solvents or chaotropic agents such as guanidium hydrochloride [5,7,19,25]. Deoxycholate (DOC) has become a more prevalently used denaturing agent due to its benefits: using DOC increases the activity of trypsin [26] and DOC is, compared to other denaturing agents, easier to remove from the sample, by acidifying the sample resulting in precipitation of DOC [27].

As mentioned, heat is an alternative to using chemicals for the denaturation of proteins. With this method, the sample is heated between 55 and 95 °C to denature the proteins [3,5,28,29]. Denaturation with heat prevents the use of chemicals and the additional steps that are required for removing the chemicals [30]. Using temperatures around 55 °C takes 1 to 2 hours for denaturation. However, when heating to 95 °C, denaturation can be completed in 2 minutes [29].

The second step is reduction of the disulfide bonds between cysteine amino acids contributing to the tertiary structure of a protein. Chemicals used for reduction are tris(2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DTT). TCEP is reported to have a better stability and stronger reducing capabilities than DTT, but DTT has a neutral pH, making DTT a more straightforward choice and thus more commonly used. When the disulfide bonds are reduced, the reduced cysteine residues are alkylated

to prevent formation of the disulfide bonds. Alkylation is usually performed in the dark by using an alkylating agent like iodoacetamide [5,7,31].

2.2.3 Optimization of Digestion

To obtain optimal results for the protein digestion, several experimental factors can be altered. Optimization could lead to a shorter digestion time, which has positive effects on achieving only the desired products. Furthermore, it could help prevent deamidation and oxidation, which will be discussed later on, and the formation of other side products [25].

The pH is of great influence on the functioning of the enzyme and the stability of digestion products. As shown in the overview in table 2, most of the enzymes have a pH range that indicates an alkaline environment for optimal digestion conditions. However, deamidation of asparagine and glutamine amino acids occurs more frequently within this optimal, alkaline pH range [19]. A study on trastuzumab tested the effect of pH on digestion with trypsin and deamidation in four samples, with pH 7.0, 7.5, 8.0 and 8.5. The sample with pH 7.0 showed the least deamidation within three hours after starting the digestion while still giving a reasonably fast digestion [32]. The pH can be maintained by using a buffer that dilutes the sample, before the digestion is started. Examples for buffers that are frequently used for tryptic digestion are Tris buffers and ammonium bicarbonate (ABC or AMBIC) buffers. Both of these buffers facilitate a pH between 7 and 9 [3,23,30,33]. For each enzyme the optimal pH should be considered and the influence of the buffer that is chosen. An example for this is Glu-C, which cleaves at different sites with different buffers, making it a more diverse enzyme. In an ammonium acetate or bicarbonate buffer, Glu-C cleaves with more specificity at the C-terminal end of glutamic acid. However, when using a phosphate buffer, Glu-C cleaves C-terminally of both glutamic acid and aspartic acid [3,18,22].

Raising the digestion temperature can shorten the digestion time, depending on the protease. With a higher temperature, the activity of trypsin is enhanced, but the activity of Glu-C would not increase [3]. Whether the temperature can be increased during digestion depends on the thermostability of the enzyme. The thermostability can be improved by modifications to the enzyme, which improves the inability to change the shape of the secondary structure. For example, the optimal temperature for tryptic activity can be shifted from 37 °C to between 50 and 60 °C with reductive methylation [23,34]. Another technique is using microwave irradiation for heat during the digestion. With a microwave, the temperature can be sufficiently controlled between 45 to 55 °C. Around these temperatures digestion with trypsin could be performed in 10 to 20 minutes [3,23,35].

The actual digestion step often is time consuming, considering it can take as long as 18 hours, leading to digestion overnight to make sure the digestion is completely finished [3,4,19,23,28,30,35]. As stated before, the digestion time can be shortened by raising the digestion temperature, whereas other methods are using organic solvents, higher concentrations of the enzyme, using pressure, infrared or ultrasound radiation and immobilized enzymes [3,19,23]. A recent study on nivolumab in samples of human serum, used immobilized trypsin for the digestion. Due to trypsin being immobilized on a solid surface, it cannot digest itself, leading to a higher efficiency in digestion and a shorter digestion time. Here, the digestion step only took 20 minutes [36].

Another aspect of efficient and rapid enzymatic digestion is the enzyme to protein ratio. The rule of thumb for overnight digestion with trypsin is 1:20 (trypsin:protein), which can be changed to fit the needs of the experiment [19]. One of the reasons for this ratio is to prevent the auto-digestion of trypsin, which is less abundant with this ratio. However, a study using bovine serum albumin, cytochrome C and human chorionic gonadotropin reported, after denaturation with heat, reduction and alkylation, that a 1:1 trypsin to protein ratio resulted in decreased digestion times when comparing

with overnight digestion at a 1:40 ratio. The exact digestion time was lower (45 minutes) for bovine serum albumin and human chorionic gonadotropin, than for cytochrome C (4 hours), to obtain the highest yield [28].

2.3 Evaluation Enzymes

To illustrate an evaluation of the most commonly used enzymes, an *in silico* digestion using mMass [37] was performed on the human growth hormone isoform 1 (hGH 1, 22 kDa), with no missed cleavages allowed and using a mass range between 100 and 5000 Da. However, it should be noted that missed cleavages do occur during experiments [18]. The protein sequence was theoretically digested with trypsin, chymotrypsin, Lys-C, Glu-C in ammonium bicarbonate, Glu-C in phosphate buffer, Arg-C, Asp-N and pepsin. The evaluation is based on the performance of the enzymes: length, charge and hydrophobicity, and the availability and price of the enzymes.

2.3.1 Pros & Cons Generated Peptides

Every enzyme has its pros and its cons, which can be determined by several aspects. Here, the length, charges and hydrophobicity of the peptides are discussed. The length of the peptides depends on where the enzymes cleave (table 2) and on the composition and sequence of amino acids in the protein. The optimal length for the quantification of the peptides with LC-MS/MS is between 7 and 20 amino acids [3,38]. The composition of amino acids also determines the charge and the polarity of the peptides. Polarity has influence on how well the peptides can be separated by LC. With the Sequence Specific Retention Calculator (SSRCalculator) [39] the relative hydrophobicity (SSRC score) can be calculated and predicted. The hydrophobicity index (HI) was calculated for a 100 Å C18 RP-HPLC column with a mobile phase of 0.1% formic acid, the version from 2010, HI = -2.6687 + 0.4954 * hydrophobicity. The SSRC score gives an indication of the interaction between the peptides and the column, which can help with the identification and quantification of the peptides. The hydrophobicity should not be too low or too high, since it results in a too short or too long retention time, which both make the quantification more complex. Therefore, a SSRC score around the range of 10 to 45 can be used as a general rule for selecting on hydrophobicity [40]. When looking at the peptides from all the enzymes, in general the peptides with a suitable length have also a suitable SSRC score.

Peptides that contain amino acids with charges are a pro for the quantification by mass spectrometry. The charges that are counted in the tables below are positive charges resulting from positively charged amino acids and the positively charged N-terminus, when the peptides are present in an acidic mobile phase, like formic acid. In such a mobile phase, the acidic amino acids and the C-terminus are neutral.

2.3.1.1 *Trypsin*

Tryptic digestion of hGH 1 (table 3) resulted in 21 peptides. Here, 10 of the peptides have a length between 7 and 20 amino acids. Trypsin cleaves after lysine (K) and arginine (R), which both are polar, basic and positively charged amino acids and giving most of the tryptic peptides two positive charges [11,18,20], which is also observed for the tryptic peptides here. Only one peptide had a single charge on the N-terminal. The length, charge and hydrophobicity of the tryptic peptides, makes trypsin a good candidate for enzymatic digestion of proteins for LC-MS/MS analysis.

Table 3: Digestion of hGH1 (22 kDa) with trypsin; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[168-168]	147.1128	K	1	+2	0.7
[17-19]	383.215	AHR	3	+3	0
[39-41]	404.214	EQK	3	+2	0
[169-172]	508.2072	DMDK	4	+2	3.91
[179-183]	618.3392	IVQCR	5	+2	5.46
[141-145]	626.3144	QTYSK	5	+2	1.97
[135-140]	693.393	TGQIFK	6	+2	16.21
[65-70]	762.3628	EETQQK	6	+2	-0.39
[173-178]	764.4301	VETFLR	6	+2	21.45
[128-134]	**773.3788**	**LEDGSPR**	**7**	**+2**	**5.73**
[184-191]	**785.3134**	**SVEGSCGF**	**8**	**+1**	**17.86**
[71-77]	**844.4887**	**SNLELLR**	**7**	**+2**	**26.5**
[1-8]	**930.5407**	**FPTIPLSR**	**8**	**+2**	**31.6**
[9-16]	**979.503**	**LFDNAMLR**	**8**	**+2**	**31.22**
[159-167]	**1148.556**	**NYGLLYCFR**	**9**	**+2**	**38.88**
[116-127]	**1361.673**	**DLEEGIQTLMGR**	**12**	**+2**	**33.85**
[146-158]	**1489.692**	**FDTNSHNDDALLK**	**13**	**+3**	**21.43**
[78-94]	**2055.2**	**ISLLLIQSWLEPVQFLR**	**17**	**+2**	**66.42**
[95-115]	2262.129	SVFANSLVYGASDSNVYDLLK	21	+2	49.01
[20-38]	**2342.134**	**LHQLAFDTYQEFEEAYIPK**	**19**	**+3**	**48.23**
[42-64]	2616.24	YSFLQNPQTSLCFSESIPTPSNR	23	+2	44.32

2.3.1.2 Lys-C

Lys-C has cleavage after lysine in common with trypsin. Digestion with Lys-C (table 4) resulted in 9 peptides, most of which are too short or too long. Only three of these peptides had a length between 7 and 20 amino acids, while each contained three positive charges. The peptides contain more positive charges than the tryptic peptides, which can be explained by Lys-C not cleaving after arginine like trypsin, thereby retaining the positively charged arginine within the peptides. Since Lys-C generates fewer and longer peptides than trypsin when digesting hGH 1, Lys-C would in this case most likely not be preferred over trypsin. However, it could be used together with or after trypsin to reduce missed cleavages of trypsin after lysine.

Table 4: Digestion of hGH1 (22 kDa) with Lys-C; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[39-41]	404.214	EQK	3	+2	0
[169-172]	508.2072	DMDK	4	+2	3.91
[141-145]	626.3144	QTYSK	5	+2	1.97
[159-168]	**1276.651**	**NYGLLYCFRK**	**10**	**+3**	**32.68**
[146-158]	**1489.692**	**FDTNSHNDDALLK**	**13**	**+3**	**21.43**
[173-191]	**2130.047**	**VETFLRIVQCRSVEGSCGF**	**19**	**+3**	**38.83**
[116-140]	2790.409	DLEEGIQTLMGRLEDGSPRTGQIFK	25	+4	43.35
[42-70]	3359.585	YSFLQNPQTSLCFSESIPTPSNREETQQK	29	+3	39.86
[1-38]	4578.339	FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPK	38	+6	62.64

2.3.1.3   Arg-C

Digestion of hGH 1 with Arg-C (table 5) resulted in 11 peptides, from which 7 had a length between 7 and 20 amino acids. Arg-C cleaves after the positively charged amino acid arginine, which gives most of the peptides two or more positive charges. Since Arg-C only cleaves after arginine, there are fewer peptides and some relatively long peptides, compared to tryptic digestion. However, when comparing with Lys-C, Arg-C produces more peptides with a suitable length, that also have sufficient positive charges for quantification. This does not mean that Arg-C is generally preferable over Lys-C; this strongly depends on the actual number and distribution of lysines and arginines along the protein chain.

Table 5: Digestion of hGH1 (22 kDa) with Arg-C; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[17-19]	383.215	AHR	3	+3	0
[179-183]	618.3392	IVQCR	5	+2	5.46
[128-134]	**773.3788**	**LEDGSPR**	**7**	**+2**	**5.73**
[184-191]	**785.3134**	**SVEGSCGF**	**8**	**+1**	**17.86**
[1-8]	**930.5407**	**FPTIPLSR**	**8**	**+2**	**31.6**
[9-16]	**979.503**	**LFDNAMLR**	**8**	**+2**	**31.22**
[168-178]	**1381.714**	**KDMDKVETFLR**	**11**	**+4**	**25.15**
[65-77]	**1587.834**	**EETQQKSNLELLR**	**13**	**+3**	**25.45**
[78-94]	**2055.2**	**ISLLLIQSWLEPVQFLR**	**17**	**+2**	**66.42**
[95-127]	3604.784	SVFANSLVYGASDSNVYDLLKDLEEGIQTLMGR	33	+3	61.66
[135-167]	3900.901	TGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFR	33	+6	44.37

2.3.1.4 Chymotrypsin

Digestion of hGH 1 with chymotrypsin (table 6) resulted in 47 relatively short peptides, from which 9 have a length between 7 and 20 amino acids. Chymotrypsin cleaves after aromatic and non-polar amino acids, except leucine (not aromatic) [11,18], so it will generally create more and smaller peptides for proteins with relatively many hydrophobic parts. In the case of hGH 1, the digestion even resulted in 12 peptides only consisting of amino acid and a few peptides consisting of only two or three amino acids. Three of the peptides with a suitable length only had one positive charge, on the N-terminus, the other 6 peptides had two or more positive charges. The fact that chymotrypsin does not cleave after a charged amino acid means that peptides can be formed with no other charge than the N-terminus. This, in combination with the fact that in general, chymotryptic cleavage results in many missed cleavages [3,18], makes it a less suitable candidate for LC-MS/MS analysis.

Table 6: Digestion of hGH1 (22 kDa) with chymotrypsin; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[45-45]	132.1019	L	1	+1	0
[76-76]	132.1019	L	1	+1	0
[81-81]	132.1019	L	1	+1	0
[82-82]	132.1019	L	1	+1	0
[87-87]	132.1019	L	1	+1	0
[93-93]	132.1019	L	1	+1	0
[114-114]	132.1019	L	1	+1	0
[157-157]	132.1019	L	1	+1	0
[163-163]	132.1019	L	1	+1	0
[177-177]	132.1019	L	1	+1	0
[10-10]	166.0863	F	1	+1	0
[164-164]	182.0812	Y	1	+1	0
[161-162]	189.1234	GL	2	+1	0
[24-25]	237.1234	AF	2	+1	0
[112-113]	247.1288	DL	2	+1	0
[43-44]	253.1183	SF	2	+1	0
[74-75]	261.1445	EL	2	+1	0
[53-54]	269.0954	CF	2	+1	0
[165-166]	269.0954	CF	2	+1	0
[102-103]	281.1496	VY	2	+1	0
[115-117]	375.2238	KDL	3	+2	0.06
[7-9]	375.235	SRL	3	+2	0.99
[144-146]	381.2132	SKF	3	+2	1.02
[21-23]	397.2194	HQL	3	+2	0
[26-28]	398.1558	DTY	3	+1	0
[98-101]	404.214	ANSL	4	+1	7.21
[29-31]	423.1874	QEF	3	+1	0
[158-160]	424.2191	KNY	3	+2	1.11
[125-128]	476.265	MGRL	4	+2	1.32
[77-80]	488.3191	RISL	4	+2	3.26
[94-97]	508.2878	RSVF	4	+2	2.6
[32-35]	511.2035	EEAY	4	+1	6.15

[83-86]	533.2718	IQSW	4	+1	19.64
[140-143]	539.2824	KQTY	4	+2	0.7
[11-15]	563.2494	DNAML	5	+1	17.59
[88-92]	619.3086	EPVQF	5	+1	18.55
[16-20]	652.4002	RAHRL	5	+4	0.77
[1-6]	687.4076	FPTIPL	6	+1	31.32
[46-52]	**787.3945**	**QNPQTSL**	**7**	**+1**	**9.83**
[118-124]	**789.3989**	**EEGIQTL**	**7**	**+1**	**16.95**
[104-111]	**812.3421**	**GASDSNVY**	**8**	**+1**	**13.99**
[36-42]	**905.5091**	**IPKEQKY**	**7**	**+3**	**1.83**
[147-156]	**1101.444**	**DTNSHNDDAL**	**10**	**+2**	**9.15**
[129-139]	**1206.575**	**EDGSPRTGQIF**	**11**	**+2**	**19.39**
[167-176]	**1268.63**	**RKDMDKVETF**	**10**	**+4**	**19.44**
[178-191]	**1540.736**	**RIVQCRSVEGSCGF**	**14**	**+3**	**20.64**
[55-73]	**2145.042**	**SESIPTPSNREETQQKSNL**	**19**	**+3**	**18.85**

2.3.1.5   Glu-C in Bicarbonate Buffer

As discussed before, Glu-C has the possibility, and thus advantage, of different specificities in ammonium or phosphate buffers [3,18,22]. In ammonium bicarbonate buffers, Glu-C cleaves after glutamic acid, which is negatively charged. Digestion of hGH 1 with Glu-C (table 7) resulted in 14 peptides, where 6 peptides have a length between 7 and 20 amino acids. All 6 of these peptides have two positive charges, which is an advantage for quantification with MS/MS. The results make Glu-C in bicarbonate buffer a good candidate for the digestion of hGH 1, but trypsin is still a better candidate.

Table 7: Digestion of hGH1 (22 kDa) with Glu-C in bicarbonate buffer; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[33-33]	148.0604	E	1	+1	0
[66-66]	148.0604	E	1	+1	0
[119-119]	148.0604	E	1	+1	0
[31-32]	295.1288	FE	2	+1	0
[187-191]	470.1704	GSCGF	5	+1	11.39
[34-39]	720.3927	AYIPKE	6	+2	8.72
[67-74]	**947.4793**	**TQQKSNLE**	**8**	**+2**	**4.99**
[57-65]	**1000.506**	**SIPTPSNRE**	**9**	**+2**	**10.22**
[120-129]	**1117.603**	**GIQTLMGRLE**	**10**	**+2**	**31.45**
[175-186]	**1450.784**	**TFLRIVQCRSVE**	**12**	**+2**	**25.86**
[75-88]	**1697.036**	**LLRISLLLIQSWLE**	**14**	**+2**	**62.65**
[40-56]	**2019.948**	**QKYSFLQNPQTSLCFSE**	**17**	**+2**	**36.01**
[89-118]	3359.716	PVQFLRSVFANSLVYGASDSNVYDLLKDLE	30	+3	56.87
[1-30]	3600.853	FPTIPLSRLFDNAMLRAHRLHQLAFDTYQE	30	+6	53.21

2.3.1.6 Glu-C in Phosphate Buffer

In phosphate buffers, Glu-C cleaves after both glutamic and aspartic acid, which are both negatively charged. As expected, cleavage in phosphate buffer (table 8) resulted in more and a few shorter peptides than with digestion in the bicarbonate buffer. The digestion resulted in 11 of the 26 formed peptides having a length between 7 and 20 amino acids. Almost all of the peptides with a suitable length have more than one positive charge, because of the presence of a lysine or arginine. Since there are more suitable peptides to choose from than after digestion with Glu-C in the bicarbonate buffer, digestion in the phosphate buffer would have the preference over the bicarbonate buffer for the digestion of hGH 1.

Table 8: Digestion of hGH1 (22 kDa) with Glu-C in phosphate buffer; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[130-130]	134.0448	D	1	+1	0
[154-154]	134.0448	D	1	+1	0
[33-33]	148.0604	E	1	+1	0
[66-66]	148.0604	E	1	+1	0
[119-119]	148.0604	E	1	+1	0
[117-118]	261.1445	LE	2	+1	0
[170-171]	265.0853	MD	2	+1	0
[31-32]	295.1288	FE	2	+1	0
[172-174]	375.2238	KVE	3	+1	3.01
[187-191]	470.1704	GSCGF	5	+1	11.39
[113-116]	488.3079	LLKD	4	+2	-2.70
[27-30]	540.23	TYQE	4	+1	3.77
[108-112]	597.2515	SNVYD	5	+1	8.46
[148-153]	687.2693	TNSHND	6	+2	-3.41
[34-39]	720.3927	AYIPKE	6	+2	8.72
[67-74]	**947.4793**	**TQQKSNLE**	**8**	**+2**	**4.99**
[57-65]	**1000.506**	**SIPTPSNRE**	**9**	**+2**	**10.22**
[120-129]	**1117.603**	**GIQTLMGRLE**	**10**	**+2**	**31.45**
[1-11]	**1305.72**	**FPTIPLSRLFD**	**11**	**+1**	**41.74**
[175-186]	**1450.784**	**TFLRIVQCRSVE**	**12**	**+3**	**25.86**
[75-88]	**1697.036**	**LLRISLLLIQSWLE**	**14**	**+2**	**62.65**
[12-26]	**1792.939**	**NAMLRAHRLHQLAFD**	**15**	**+5**	**26.09**
[155-169]	**1816.978**	**ALLKNYGLLYCFRKD**	**15**	**+4**	**39.48**
[131-147]	**1959.992**	**GSPRTGQIFKQTYSKFD**	**17**	**+4**	**22.29**
[40-56]	**2019.948**	**QKYSFLQNPQTSLCFSE**	**17**	**+2**	**36.01**
[89-107]	**2070.066**	**PVQFLRSVFANSLVYGASD**	**19**	**+2**	**46.84**

2.3.1.7 Asp-N

Asp-N cleaves at the N-terminal of aspartic acid, which has a negatively charged side chain. The digestion of hGH 1 with Asp-N (table 9) resulted in 11 relatively short peptides. Five of these peptides have a length between 7 and 20 amino acids, all of them have one positive charge or more. Comparing Asp-N with Arg-C, which also resulted in 11 peptides, Asp-N resulted in only 5 peptides that have a suitable length, whereas Arg-C resulted in 7 peptides with a suitable length.

Table 9: Digestion of hGH1 (22 kDa) with Asp-N; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[153-153]	134.0448	D	1	+1	0
[169-170]	265.0853	DM	2	+1	0
[112-115]	488.3079	DLLK	4	+2	17.23
[107-111]	597.2515	DSNVY	5	+1	10.32
[147-152]	687.2693	DTNSHN	6	+2	-2.03
[1-10]	**1190.693**	**FPTIPLSRLF**	**10**	**+1**	**44.58**
[116-129]	**1603.8**	**DLEEGIQTLMGRLE**	**14**	**+2**	**41.88**
[11-25]	**1792.939**	**DNAMLRAHRLHQLAF**	**15**	**+5**	**28.71**
[154-168]	**1816.978**	**DALLKNYGLLYCFRK**	**15**	**+4**	**37.75**
[130-146]	**1959.992**	**DGSPRTGQIFKQTYSKF**	**17**	**+3**	**17.94**
[171-191]	2373.169	DKVETFLRIVQCRSVEGSCGF	21	+4	39.62

2.3.1.8 Pepsin

Pepsin has less specific proteolytic activity than, for example, trypsin. Due to pepsin being less specific, it also generates more peptides [23], which was also observed here. A total of 36 peptides was obtained by peptic digestion of hGH 1 (table 10), from which 10 peptides had a length between 7 and 20 amino acids. Pepsin cleaves after the nonpolar amino acids phenylalanine, leucine and tryptophan, and the uncharged polar amino acid tyrosine. The named amino acids do not have a charge, which leaves some peptides with only the positive charge from the amino group from the N-terminal. All of the peptic peptides are relatively short, with most of them even having a length below 10 peptides. Furthermore, pepsin is less specific in cleavage [5], resulting in a peptide mixture that is complex to analyze [18].

Table 10: Digestion of hGH1 (22 kDa) with pepsin; showing the slice, m/z, sequence, length, charge and SSRC score of the peptides.

Slice	m/z	Sequence	Length	Charge	SSRC Score
[45-45]	132.1019	L	1	+1	0
[76-76]	132.1019	L	1	+1	0
[81-81]	132.1019	L	1	+1	0
[82-82]	132.1019	L	1	+1	0
[93-93]	132.1019	L	1	+1	0
[114-114]	132.1019	L	1	+1	0
[157-157]	132.1019	L	1	+1	0
[163-163]	132.1019	L	1	+1	0
[177-177]	132.1019	L	1	+1	0
[1-1]	166.0863	F	1	+1	0
[10-10]	166.0863	F	1	+1	0
[74-75]	261.1445	EL	2	+1	0
[53-54]	269.0954	CF	2	+1	0
[115-117]	375.2238	KDL	3	+2	0.06
[7-9]	375.235	SRL	3	+2	0.99
[164-166]	432.1588	YCF	3	+1	0
[125-128]	476.265	MGRL	4	+2	1.32
[77-80]	488.3191	RISL	4	+2	3.26
[2-6]	540.3392	PTIPL	5	+1	19.48
[11-15]	563.2494	DNAML	5	+1	17.59
[158-162]	594.3246	KNYGL	5	+2	13.73
[21-25]	615.3249	HQLAF	5	+2	15.95
[88-92]	619.3086	EPVQF	5	+1	18.55
[83-87]	646.3559	IQSWL	5	+1	25.8
[16-20]	652.4002	RAHRL	5	+4	0.77
[46-52]	**787.3945**	**QNPQTSL**	**7**	**+1**	**9.83**
[118-124]	**789.3989**	**EEGIQTL**	**7**	**+1**	**16.95**
[26-31]	802.3254	DTYQEF	6	+1	18.61
[140-146]	**901.4778**	**KQTYSKF**	**7**	**+3**	**13.43**
[147-156]	**1101.444**	**DTNSHNDDAL**	**10**	**+2**	**9.15**
[129-139]	**1206.575**	**EDGSPRTGQIF**	**11**	**+2**	**19.39**
[167-176]	**1268.63**	**RKDMDKVETF**	**10**	**+4**	**19.44**
[178-191]	**1540.736**	**RIVQCRSVEGSCGF**	**14**	**+3**	**20.64**
[32-44]	**1631.795**	**EEAYIPKEQKYSF**	**13**	**+3**	**23.99**
[55-73]	**2145.042**	**SESIPTPSNREETQQKSNL**	**19**	**+3**	**18.85**
[94-113]	**2177.051**	**RSVFANSLVYGASDSNVYDL**	**20**	**+2**	**40.21**

2.3.2 Availability & Price

The availability and price of enzymes can play a role in choosing an enzyme for the digestion step. The availability and prices were investigated at three suppliers: Promega [41], Sigma-Aldrich [42] and Thermo Fisher Scientific [43]. Table 11 presents an overview of the availability and prices of commonly used enzymes in the biggest size available at the three suppliers. Since the amount of enzyme that is necessary for the digestion depends on how much protein is present in the sample and to be able to compare prices for the enzymes, the prices were calculated to price/mg in table 12.

The least expensive enzymes are pepsin and untreated chymotrypsin and trypsin [41–43]. Pepsin is easy to obtain and is widely used for defining disulfide bonds, especially in antibodies [42,44]. Trypsin is able to hydrolyze itself (auto-digestion), giving it chymotryptic activity, resulting in cleavage at sites that chymotrypsin normally cleaves. To prevent chymotryptic activity, chymotrypsin inhibitors can be used. An example of such an inhibitor is N-p-Tosyl-L-phenylalanine chloromethyl ketone (TPCK) [16,30,34]. This makes it useful to use modified trypsin, meaning it is treated with TPCK. Although modified trypsin is more expensive than unmodified trypsin, they are both readily available. Another way of using trypsin is with TPCK treated immobilized trypsin, which shortens the digestion time and makes it easier to separate trypsin from the resulting peptide solution [41–43]. Furthermore, a protease mix containing TPCK treated trypsin and Lys-C is available. Trypsin has a slight preference for cleaving after arginine, resulting in more missed cleavages at lysine [18,20,44,45]. The added Lys-C then covers for missed cleavages at lysine, to obtain complete digestion [41,43].

Chymotrypsin is also an inexpensive enzyme. However, it is less specific, like pepsin, and chymotryptic digestion results in lots of missed cleavages [3]. Since chymotrypsin can show tryptic activity, it is possible to treat chymotrypsin with N-Tosyl-L-lysine chloromethyl ketone (TLCK), which is a trypsin inhibitor [16], or use highly purified, sequencing grade chymotrypsin. As expected, TLCK treated chymotrypsin is more expensive than non-treated chymotrypsin [41–43]. However, as observed in the overview in table 11, it appears that TPCK treated trypsin and TLCK treated chymotrypsin are becoming more standard.

The enzymes with a higher specificity; Glu-C, Arg-C, Lys-C and Asp-N, are also the most expensive enzymes, respectively. These enzymes are commercially available [41–43] and become more popular as alternatives for trypsin [18,44].

Table 11: Overview of availability and prices of enzymes from three suppliers: Promega [41]*, Sigma-Aldrich* [42] *and Thermo Fisher* [43]*.*

Enzyme	Promega		Sigma-Aldrich		Thermo Fisher	
Pepsin	250 mg	€28.00	1 kg	€1,100.00	-	-
Chymotrypsin (untreated)	-	-	10 g	€691.00	-	-
Trypsin (untreated)	-	-	10 g	€539.00	-	-
Trypsin (TPCK treated)	100 µg	€77.00	5 g	€1,930.00	500 µg	€352.00
Chymotrypsin (TLCK treated)	100 µg*	€240.00*	100 mg	€364.00	100 µg	€256.00
Trypsin (TPCK treated)/Lys-C	100 µg	€142.00	-	-	100 µg	€153.00
Glu-C	50 µg	€142.00	50 µg	€175.00	50 µg	€253.00
Arg-C	10 µg	€199.00	15 µg	€363.00	-	-
Lys-C	20 µg	€414.00	15 µg	€465.00	100 µg	€894.00
Asp-N	2 µg	€177.00	6 µg	€431.00	2 µg	€272.00

*Highly purified, sequencing grade chymotrypsin, not TLCK treated.

Table 12: Overview of availability and prices/mg of enzymes from three suppliers: Promega [41], Sigma-Aldrich [42] and Thermo Fisher [43].

Enzyme	Promega	Sigma-Aldrich	Thermo Fisher
Pepsin	€0.11	€0.001	-
Chymotrypsin (untreated)	-	€0.07	-
Trypsin (untreated)	-	€0.05	-
Trypsin (TPCK treated)	€770.00	€0.39	€704.00
Chymotrypsin (TLCK treated)	€2,400.00*	€3.64	€2,560.00
Trypsin (TPCK treated)/Lys-C	€1,420.00	-	€1,530.00
Glu-C	€2,840.00	€3,500.00	€5,060.00
Arg-C	€19,900.00	€24,200.00	-
Lys-C	€20,700.00	€31,000.00	€8,940.00
Asp-N	€88,500.00	€71,833.33	€136,000.00

*Highly purified, sequencing grade chymotrypsin, not TLCK treated.

2.3.3 Choosing an Enzyme

Based on the performance, the availability and prices of the enzymes, an enzyme with the best characteristics can be chosen for the digestion of hGH 1. TPCK treated trypsin would be the preferred enzyme for the digestion, considering the relatively low price, the wide availability, the specificity and the characteristics of the obtained peptides compared with the other enzymes. Almost all the tryptic peptides have at least two positive charges, which is ideal for the quantification with MS. Furthermore, trypsin generated enough peptides with a suitable length. Although the situation for other proteins obviously depends on the number, nature and distribution of the amino acids in their structures, it can be speculated that the same considerations apply and that the popularity of trypsin can be explained by the fact that it generally creates peptides of a suitable size and polarity with sufficient charges, and for a reasonable price.

3. Selection of Peptides

After the digestion of the protein, the sample contains peptides. When selecting peptides from the protein sample, there are certain criteria that the peptide should meet for successful protein quantification. Table 13 shows an overview of the criteria, which can be used for selecting a signature peptide. One of the most important aspects is that the sequence of the peptide should be unique, to make sure that only the target protein is quantified and not other proteins present in the matrix. This unique peptide is then called the signature peptide [3,38].

There are several programs for *in silico* digestion, that simulate the digestion with several enzymes. An example of such programs is mMass [37]. After the simulated digestion, the obtained peptide sequences can be checked in databases for the uniqueness of the peptide sequences, for example with BLAST [40,46]. For the sequence to be unique, there should be no sequence matches with other peptides or proteins, but only with the target protein.

To select a unique peptide, the length of the signature peptide is of great influence. Different lengths are reported and most of these lengths differ between 7 and 20 amino acids, with an average around 14 amino acids [7,35,40]. When the peptide length becomes too short, there is a smaller chance for the peptide to be unique [3,5,23]. Selecting a peptide that is 'too long', can increase the charge state distributions, making analyses harder [40].

Furthermore, the characteristics of the amino acids present in the peptides should be considered. Such characteristics are the stability, the charge, hydrophobicity and how they act together [7,35,40]. The

stability of an amino acid indicates how reactive and prone the amino acid is to modifications and thus how likely it is that degradation occurs during storage or analysis. Two modifications that will be discussed due to their prevalence are oxidation and deamidation, which both do not only happen during storage, but can also occur during the experiment [29,35,40,47–50].

3.1   Oxidation

A very prevalent modification of amino acids is oxidation. Several amino acids that are aromatic or contain a sulfur atom are prone to oxidation and one should try to avoid selecting a signature peptide containing those: cysteine (C), tryptophan (W), tyrosine (Y), histidine (H), phenylalanine (F) and especially methionine (M) [32,35,40]. There are many factors that have an influence on oxidation, for example; the temperature, the pH, the buffer, light, oxidizing agents like hydrogen peroxide, and several metals [51]. Oxidation occurs via reactive oxygen species (ROS) that are induced by the named factors [29,47,48,51]. Oxidized amino acids are more hydrophilic than the non-oxidized amino acids [3], which could lead to unfolding of the protein and thus to changes in its shape. This results in more unstable peptides and proteins, which can give rise to decreased therapeutic activity. Furthermore, the oxidized amino acids have a mass difference of +15.9949 Da compared to the non-oxidized amino acids [3,29,32,40,50].

Oxidation of amino acids can be prevented by taking measures. For example, use packing that keeps light away from the sample, storage of the sample at a low temperature of 4 °C or frozen between -20 and -80 °C [40], make sure there are no metals present that induce ROS formation and minimize the time that the sample has contact with solvents minimize the amount of oxygen [29,40,51].

However, sometimes it is not possible to select a signature peptide without an amino acid that is prone to oxidation. A study on quantifying rhTRAIL in serum from mouse and human used signature peptides, both containing methionine, to be able to differentiate between two variants of rhTRAIL that differed only in one amino acid. To avoid uncontrolled degradation of the signature peptides, they were fully oxidized with hydrogen peroxide, making it possible to quantify the peptides in a reliable manner [52].

3.2   Deamidation

Another prevalent modification is deamidation. Deamidation is the removal of amide groups from glutamine (Q) and asparagine (N), which eventually leads to the formation of the other amino acids glutamic acid (E) and aspartic acid (D) via intermediates [32,50]. The deamidated products have a mass difference of +0.9840 Da and they are more hydrophobic, compared to the non-deamidated amino acids. Deamidation is most pronounced for asparagine (N) and especially when followed by a small amino acid, such as glycine (G) or serine (S). At different pH values, the deamidation reaction takes place in a different manner. The first type of deamidation is the hydrolysis of the amide group, which occurs at pH values lower than 3. This deamidation is mostly seen for asparagine and less for glutamine. The second type of deamidation is cyclization within glutamine or asparagine, which occurs at pH values higher than 6 [35,50]. The cyclization leads to a five-membered ring structure, also named succinimide. This intermediate is unstable at a pH of 6 or higher and is hydrolyzed. For example, the amide group of asparagine cyclized to the succinimide intermediate, which then hydrolyzes in aspartic acid and an isoform of aspartic acid, isoaspartic acid [32,50].

Due to deamidation, there is a possibility of instability and loss of protein activity, depending on the location of the amino acid and on whether the polypeptide chain is folded into a protein or digested into peptides [32,35]. That peptides are faster deamidated compared to intact proteins was shown in a study on trastuzumab where the deamidation of a particular asparagine in intact trastuzumab in plasma was compared to deamidation of the same asparagine in the signature peptide when trastuzumab was added to a digestion buffer containing trypsin [32].

When it is not possible to select signature peptides without asparagine or glutamine, there are ways to prevent deamidation as much as possible. For example, keep the digestion time as short as possible and preferably under an hour [35], using a low concentration buffer with a pH between 3 and 6 and, here the same as for oxidation, storage of the samples at a low temperature [32,35,40].

3.3   Post-translational Modifications

Post-translational modification (PTM) is the addition of a functional group to an amino acid, which occurs after the synthesis of (biopharmaceutical) proteins [53]. Not only influences a PTM the function, the structure, activity or inactivity and many more aspects of proteins [54], but also the digestion and quantification of proteins [23,54]. Glycosylation, for example, is the enzymatic addition of glycan groups to the amide group of asparagine (N-glycosylation) and the amino acid side chains of serine and threonine (O-glycosylation) [35]. Phosphorylation is the addition of a phosphoryl group to the amino acid side chains of threonine, serine and tyrosine [54]. These glycan and phosphoryl groups can cause steric hindrance and thereby prevent the enzymatic cleavage of peptide bonds [23]. Signature peptides including such PTMs are therefore to be avoided. Other common PTMs are disulfide bond formation, pyroglutamic acid formation and glycation [35,40,49].

Table 13: Overview of criteria for choosing a signature peptide with comments.

Criteria	Comments
Length peptide	Between 7 and 20 amino acids[a].
Charge	Peptides containing positive charges are ideal for analysis with MS/MS[b].
Hydrophobicity	An SSRC score around the range of 10 to 45[a].
Amino acids	Choose peptides without amino acids that are prone to the following modifications[c]: • Oxidation: M, C, W • Deamidation: Q or N, followed by G • Disulfide bridge formation: C • Known PTMs: glycosylated N, S, T; phosphorylated T, S, Y
Uniqueness	The signature peptide should be unique for the analyte protein, which can be checked by using BLAST[d].

[a] [40], [b] [23], [c] [5,40], [d] [46]

3.4   Peptide Selection Example hGH 1 & 2

To show an example on how signature peptides can be selected for quantification and how two structurally similar proteins can be distinguished by means of their signature peptides, two isoforms of hGH with a mass of 22 kDa were digested in mMass, while allowing no missed cleavages and using a mass range of 100 to 5000 Da. Based on the evaluation of the different enzymes in the previous chapter, trypsin was chosen as protease for the digestion. The obtained peptides were selected by using the criteria that are shown in table 13.

First, a selection was made of peptides that have a length between 7 and 20 amino acids, which are shown in table 14, the other peptides were left out. Next, the SSRC score of the peptides should be around the range of 10 to 45. So, the peptides that were far below or far above this range were crossed out to keep the peptides with a suitable hydrophobicity. Peptides that contained the amino acids methionine, cysteine, tryptophan, or glutamine or asparagine followed by glycine were also crossed out. This resulted in hGH 1 having four possible peptides and hGH 2 having five possible peptides (table 15).

Table 14: Peptides after tryptic digestion of hGH 1 and hGH 2 with a mass of 22 kDa, the slice, m/z, sequence, length, charge and SSRC scores. The peptides that do not fulfill the criteria for a signature peptide are crossed out.

hGH 1 22 kDa	Slice	m/z	Sequence	Length	Charge	SSRC Score
	[128-134]	773.3788	LEDGSPR	7	+2	5.73
	[184-191]	785.3134	SVEGSCGF	8	+1	17.86
	[71-77]	844.4887	SNLELLR	7	+2	26.5
	[1-8]	930.5407	FPTIPLSR	8	+2	31.6
	[9-16]	979.503	LFDNAMLR	8	+2	31.22
	[159-167]	1148.556	NYGLLYCFR	9	+2	38.88
	[116-127]	1361.673	DLEEGIQTLMGR	12	+2	33.85
	[146-158]	1489.692	FDTNSHNDDALLK	13	+3	21.43
	[78-94]	2055.2	ISLLLIQSWLEPVQFLR	17	+2	66.42
	[20-38]	2342.134	LHQLAFDTYQEFEEAYIPK	19	+3	48.23
hGH 2 22 kDa						
	[128-134]	773.38	LEDGSPR	7	+2	5.73
	[184-191]	785.31	SVEGSCGF	8	+1	17.86
	[71-77]	844.49	SNLELLR	7	+2	26.5
	[1-8]	930.54	FPTIPLSR	8	+2	31.6
	[9-16]	979.5	LFDNAMLR	8	+2	31.22
	[150-158]	1012.5	SHNDDALLK	9	+3	13.93
	[159-167]	1148.6	NYGLLYCFR	9	+2	38.88
	[135-145]	1272.6	TGQIFNQSYSK	11	+2	22.01
	[116-127]	1490.7	DLEEGIQTLMWR	12	+2	43.82
	[95-112]	1948.9	SVFANSLVYGASDSNVYR	18	+2	38.96
	[78-94]	2021.2	ISLLLIQSWLEPVQLLR	17	+2	65.88
	[20-38]	2400.2	LYQLAYDTYQEFEEAYILK	19	+2	56.32

Table 15 shows the peptides that are left after removing the peptides that do not fulfill the criteria for a signature peptide. All peptides contain two or three positive charges, making MS/MS analysis easier. Slice [1-8] is not unique for hGH and is therefore unsuitable as a signature peptide. When comparing the other peptides of hGH 1 and hGH 2, for both isoforms slices [71-77] is the same. The other peptides occur in one isoform of the protein, but not in the other and can therefore in principle be used as a unique signature peptide for the specific isoform of hGH. Especially, slices [146-158] for hGH 1 and [135-145] for hGH 2 are interesting, because they have quite similar hydrophobicities and are likely to elute close together in the LC chromatograms.

Table 15: Tryptic peptides of hGH 1 and hGH 2 (22 kDa) left after crossing out peptides that do not fulfill the criteria, the slice, m/z, sequence, length, charge and SSRC score. The peptides are checked with BLAST if they are unique for hGH and if they are unique for the isoforms hGH 1 and hGH 2.

hGH 1 22 kDa	Slice	m/z	Sequence	Length	Charge	SSRC Score	Unique for hGH	Unique for hGH 1
	[71-77]	844.4887	SNLELLR	7	+2	26.5	Yes	No
	[1-8]	930.5407	FPTIPLSR	8	+2	31.6	No	No
	[146-158]	**1489.692**	**FDTNSHNDDA LLK**	**13**	**+3**	**21.43**	**Yes**	**Yes**
	[20-38]	**2342.134**	**LHQLAFDTYQE FEEAYIPK**	**19**	**+3**	**48.23**	**Yes**	**Yes**
hGH 2 22 kDa								Unique for hGH 2
	[71-77]	844.49	SNLELLR	7	+2	26.5	Yes	No
	[1-8]	930.54	FPTIPLSR	8	+2	31.6	No	No
	[150-158]	**1012.5**	**SHNDDALLK**	**9**	**+3**	**13.93**	**Yes**	**Yes**
	[135-145]	**1272.6**	**TGQIFNQSYSK**	**11**	**+2**	**22.01**	**Yes**	**Yes**
	[95-112]	**1948.9**	**SVFANSLVYGA SDSNVYR**	**18**	**+2**	**38.96**	**Yes**	**Yes**

4. Sample Preparation

The preparation of biological samples for the quantification of proteins is an important process. It includes choosing and using internal standards, the extraction of proteins, the actual digestion step and the extraction of peptides. Below, these subjects will be discussed and how they influence the digestion step and the final quantification.

4.1 Internal Standard

Internal standards (IS) are chemically and physically alike to the analyte that is to be quantified. However, small differences between the internal standard and the analyte make it possible to still recognize the individual responses in LC chromatograms and MS/MS mass spectra. The usage of an internal standard adds more accuracy and precision to the quantification, compared to not using an internal standard [25]. There are several types of internal standards, roughly divided in protein internal standards and peptide internal standards, which can be added to the sample during different steps in the quantification of the target protein [5,25]. Table 16 shows an overview of the different internal standards and what they correct for during the sample preparation.

To start, protein internal standards are added to the sample, before digestion takes place. The advantage of using protein internal standards is that they go through the whole process and, therefore, can possibly correct for all variability during the process. There are two variants of protein internal standards: stable-isotope labeled (SIL) proteins and structural analogue proteins. SIL-protein internal standards have the same structure as the analyte, but the polypeptide chain contains one or more amino acids labeled with a stable isotope, such as ^{13}C and/or ^{15}N. The labelling makes it possible to distinguish between the internal standard and analyte by mass spectrometry, while they still have the same physicochemical characteristics [5,7]. The structural analogue protein internal standards can also be used, however, they do not have the exact same structure as the analyte. Since the structure is different, the digestion of the analogue and the analyte will give different peptides [5,25]. Whereas the SIL-proteins correct for extraction of the protein, the digestion and extraction of the peptides; analogues do not correct for extraction of the protein, but do correct to a lesser extent for the digestion

and the extraction of the peptides. On the other hand, analogues have a better availability than SIL-proteins, since SIL-proteins have to be produced by modified organisms or by wheat germ or *E. coli* lysate [25].

The second group are the peptide internal standards, which can be used when there is no possibility of obtaining a protein internal standard. Also here, different options are available for peptide internal standards: SIL-peptides, SIL-peptides with a group or groups that are cleavable by the used enzyme, structural analogue peptides and differential derivatization [5,25]. SIL-peptides are usually the isotope labeled versions of the signature peptide of the analyte and they can be added to the sample before [3] the digestion step and after the digestion step [5,25,33,55]. Since it is already a peptide, it is not digested and, therefore, SIL-peptides do not correct for digestion of the analyte. However, there is a possibility, to some extent, for correcting for digestion by using the cleavable SIL-peptides as internal standard [3,5,25,33]. This type of internal standard is made of the SIL-peptide of the signature peptide of the internal standard, with groups attached that will be cleaved during the digestion step [3,7,25]. Analogue peptides can also be used, even though it is not commonly used anymore. The analogue peptides corrects the least compared to the other internal standard options [5,25,33]. The last type of internal standards are differential derivatized peptides. The sample with the analyte is digested, while at the same time a solution containing the protein at a known concentration is digested. The sample, now containing peptides and signature peptide(s), is derivatized with, for example, formaldehyde. The other solution with the, now also digested, protein is derivatized with deuterated formaldehyde. The deuterated formaldehyde then 'labels' the peptides, making it possible to distinguish the peptides of the sample from the solution. Next, the solution is added to the sample, so it can be analyzed [3,5,25,33].

Table 16: Overview of the different options for an internal standard and what they correct for. [7,25]

Internal Standard / Corrects for	SIL protein	Analogue protein	SIL peptide	SIL peptide (cleavable)	Analogue peptide	Differential derivatization peptide
Protein Extraction	✓	-	-	-	-	±
Digestion	✓	±	-	±	-	±
Peptide Extraction	✓	±	✓	✓	-	✓

4.2 Protein Extraction, Before Digestion

There are different biological sample types, that all need a different approach for the extraction of the protein analyte. Three common sample types are blood (which is usually processed to plasma or serum by removing blood cells), urine and cerebrospinal fluid. The remaining sample without the target proteins is called matrix. Matrix from biological samples contains proteins, lipids, salts and other small molecules, for example. The exact composition of the matrix differs per sample and per sample type. Not removing the matrix from the sample could lead to interference of the analysis of the analyte and could thus decrease the sensitivity of the quantification [5,19,56]. In particular, protein-rich matrices such as plasma and serum will generate an extremely high number of background peptides upon digestion, released from endogenous matrix proteins. Extraction of the protein analyte before digestion is then helpful to reduce interferences. Typical approaches for extraction of the target protein from the sample are protein precipitation, solid phase extraction (SPE), liquid-liquid extraction (LLE) and immunocapture [56].

Protein precipitation is a technique that uses the solubility of the proteins and the analyte. The addition of certain solvents makes the proteins in the matrix or the analyte less soluble and lets them

precipitate, so they can be removed. It is especially useful for quantification of smaller proteins which are not precipitated themselves. Solid-phase extraction (SPE) can be used for the extraction of the analyte from the sample and is based on hydrophobicity or charge, for example. The sample is run through a column or another instrument that contains absorbing compounds, which is the solid stationary phase. The analyte will interact with and remain in the column, while the compounds from the matrix will be washed out. By using a solution that the analyte has a stronger interaction with than the column, the analyte is also washed out. SPE is useful to extract protein analytes that have a clearly different hydrophobicity or charge than most of the matrix proteins. Liquid-liquid extraction (LLE) is the addition of a solvent to the sample to create a two-layer system. The solvent has a different polarity than the sample, which will lead to the analyte transferring between the samples and thus be extracted [56]. Most unmodified proteins are insoluble in organic solvents, so LLE is not used a lot. It can be applied to proteins containing a large hydrophobic modification such as polyethylene glycol (PEG) moiety. The most widely applied protein extraction method is immunocapture. This method is based on the very selective binding of the protein analyte to an antibody or other target for which the analyte has affinity and that is present on, for example, magnetic beads. After capturing the protein on the beads, all other matrix proteins are removed by washing, the beads are then isolated by a magnet and the protein is eluted by acidification for further analysis [19].

4.3   Peptide Extraction, After Digestion

After extraction and digestion of the protein analyte, signature peptides can also be extracted to further increase the sensitivity. However, this step is not always executed within sample preparation [3]. The extraction of peptides from the digested sample can be more complicated, due to the high similarity of the peptides. Reported techniques are immunocapture and SPE. Immunocapture of signature peptides has a similar principle as the immunocapture of proteins, only now antibodies with affinity for the signature peptide are used [3,19]. Reversed-phase SPE (RP-SPE) [3,36], strong-cation exchange SPE (SCX-SPE) or strong anion-exchange SPE (SAX-SPE) have also been used for peptide extraction. SCX-SPE is very effective for the extraction of tryptic peptides, considering these peptides contain at least two positive charges [3,57]. Still, to properly discriminate between signature peptides and other tryptic peptides from the matrix, it is helpful if signature peptides have more positive charges so that they are more strongly retained on the SPE phase than the average matrix peptide.

5.   Conclusion

In this overview, the enzymatic digestion of proteins in biological samples for quantification with LC-MS/MS was discussed. The proteases used for the digestion step have the same mechanism of action, but with different active groups that function as the nucleophile in the binding site. The shape of the binding site leads to the enzymes cleaving at different sites of the polypeptide chain and in some cases the cleavage site can also be influenced by the pH. In general, the protein is denatured, reduced and alkylated before the digestion. Optimization of the enzymatic digestion to fit the specificities of the enzyme and the protein and thereby reducing the digestion time could result in more sensitive quantification with LC-MS/MS. Factors for optimization are the pH, temperature, the enzyme:protein ratio and using immobilized enzymes.

As an illustrative example, an evaluation was made by *in* silico digestion of hGH 1 (22 kDa) with the most commonly used enzymes in protein quantification with LC-MS/MS, based on the length, charge, SSRC score, and availability and prices of the enzymes. Trypsin came out as the enzyme with the most suitable peptides that is widely available at a relatively low price. These results were used to illustrate how the process of selecting a signature peptide works for quantification of two very similar isoforms

of hGH, when taking several criteria into account. Several peptides were found to be unique for hGH 1 and hGH 2 and also fulfilled all other criteria.

Furthermore, sample preparation was discussed. Internal standardization is an important tool to correct for analytical variability in the different steps within the whole sample preparation process, to obtain more accuracy and precision to the quantification. The only internal standards that correct for the whole sample preparation are SIL proteins. Also for sample preparation, the extraction of proteins from the matrix of the biological sample keeps the sensitivity and selectivity of the quantification high, which can be accomplished by using techniques like immunocapture and, to a lesser extent, SPE and LLE. For the extraction of peptides SPE and immunocapture can be used.

6. Bibliography

[1]   Beccaria M, Cabooter D. Current developments in LC-MS for pharmaceutical analysis. Analyst 2020;145:1129–57. https://doi.org/10.1039/c9an02145k.

[2]   van de Merbel NC. Advances in Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)-Based Quantitation of Biopharmaceuticals in Biological Samples. LC-GC North America 2016;34:28–34.

[3]   de Jong KAM, van Breugel SJ, Hillebrand MJX, Rosing H, Huitema ADR, Beijnen JH. Bottom-up sample preparation for the LC-MS/MS quantification of anti-cancer monoclonal antibodies in bio matrices. Bioanalysis 2020;12:1405–25. https://doi.org/10.4155/bio-2020-0204.

[4]   van de Merbel NC. Protein quantification by LC-MS: A decade of progress through the pages of Bioanalysis. Bioanalysis 2019;11:629–44. https://doi.org/10.4155/bio-2019-0032.

[5]   van den Broek I, Niessen WMA, van Dongen WD. Bioanalytical LC-MS/MS of protein-based biopharmaceuticals. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 2013;929:161–79. https://doi.org/10.1016/j.jchromb.2013.04.030.

[6]   Robotham AC, Kelly JF. LC-MS characterization of antibody-based therapeutics. Approaches to the Purification, Analysis and Characterization of Antibody-Based Therapeutics, Elsevier; 2020, p. 1–33. https://doi.org/10.1016/b978-0-08-103019-6.00001-1.

[7]   el Amrani M, Donners AAM, Hack CE, Huitema ADR, van Maarseveen EM. Six-step workflow for the quantification of therapeutic monoclonal antibodies in biological matrices with liquid chromatography mass spectrometry – A tutorial. Analytica Chimica Acta 2019;1080:22–34. https://doi.org/10.1016/j.aca.2019.05.076.

[8]   Harris DC, Lucy CA. Quantitative Chemical Analysis. Ninth Edition, New York: W. H. Freeman & Company; 2016, p. 559–91, 674–700.

[9]   McDonald AG, Boyce S, Tipton KF. ExplorEnz: The primary source of the IUBMB enzyme list. Nucleic Acids Research 2009;37. https://doi.org/10.1093/nar/gkn582.

[10]   Barrett AJ, Rawlings ND, Woessner JF, editors. Handbook of Proteolytic Enzymes. vol. Volume 1. Third Edition, San Diego: Elsevier Science & Technology; 2012, p. Ii–Iiii.

[11]   Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, et al. Molecular Biology of the Cell. 6th edition, New York: Garland Science, Taylor & Francis Group; 2015, p. 94, 95, 110–3, 134–6, 140–1.

[12]   Bruice PY, Robinson JK, McMurry J. Organic Chemistry. In: Zalesky J, editor. vol. Global Edition. 8th edition, 2016, p. 1061, 1082–7.

[13]   Erez E, Fass D, Bibi E. How intramembrane proteases bury hydrolytic reactions in the membrane. Nature 2009;459:371–8. https://doi.org/10.1038/nature08146.

[14]   Ahern K, Rajagopal I, Tan T. Biochemistry. vol. 1.3, Oregon State University; 2018, p. 382–5.

[15]   Kevin Ahern's Biochemistry (BB 450/550) at Oregon State University n.d. http://oregonstate.edu/instruct/bb450/fall14/lecture/catalyticstrategiesoutline.html (accessed March 15, 2021).

[16]   Punekar NS. ENZYMES: Catalysis, Kinetics and Mechanisms. 1st ed., Singapore: Springer Singapore; 2018, p. 58, 242–3. https://doi.org/10.1007/978-981-13-0785-0.

[17]   ChemDraw Professional n.d.

[18]   Giansanti P, Tsiatsiani L, Low TY, Heck AJR. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. Nature Protocols 2016;11:993–1006. https://doi.org/10.1038/nprot.2016.057.

[19]   van de Merbel N. Sample Preparation for LC-MS Bioanalysis of Proteins. In: Li W, Fu Y, editors. Sample Preparation in LC-MS Bioanalysis. 1st edition, John Wiley & Sons; 2019, p. 304–18.

[20]   Šlechtová T, Gilar M, Kalíková K, Tesařová E. Insight into Trypsin Miscleavage: Comparison of Kinetic Constants of Problematic Peptide Sequences. Analytical Chemistry 2015;87:7636–43. https://doi.org/10.1021/acs.analchem.5b00866.

[21]   Gershon PD. Cleaved and missed sites for trypsin, Lys-C, and Lys-N can be predicted with high confidence on the basis of sequence context. Journal of Proteome Research 2014;13:702–9. https://doi.org/10.1021/pr400802z.

[22]   Liu S, Moulton KR, Auclair JR, Zhou ZS. Mildly acidic conditions eliminate deamidation artifact during proteolysis: Digestion with endoprotease Glu-C at pH 4.5. Amino Acids 2016;48:1059–67. https://doi.org/10.1007/s00726-015-2166-z.

[23]   Switzar L, Giera M, Niessen WMA. Protein digestion: An overview of the available techniques and recent developments. Journal of Proteome Research 2013;12:1067–77. https://doi.org/10.1021/pr301201x.

[24]   Enzyme Explorer - Pepsin | Sigma-Aldrich n.d. https://www.sigmaaldrich.com/life-science/metabolomics/enzyme-explorer/analytical-enzymes/pepsin.html (accessed March 4, 2021).

[25]   Bronsema KJ, Bischoff R, van de Merbel NC. Internal standards in the quantitative determination of protein biopharmaceuticals using liquid chromatography coupled to mass spectrometry. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 2012;893–894:1–14. https://doi.org/10.1016/j.jchromb.2012.02.021.

[26]   Leon IR, Schwammle V, Jensen ON, Sprenger RR. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. Molecular and Cellular Proteomics 2013;12:2992–3005. https://doi.org/10.1074/mcp.M112.025585.

[27]   Li D, Farchone A, Zhu Q, MacChi F, Walker DE, Michels DA, et al. Fast, Robust, and Sensitive Identification of Residual Host Cell Proteins in Recombinant Monoclonal Antibodies Using Sodium Deoxycholate Assisted Digestion. Analytical Chemistry 2020;92:11888–94. https://doi.org/10.1021/acs.analchem.0c02258.

[28]   Egeland SV, Reubsaet L, Halvorsen TG. The pros and cons of increased trypsin-to-protein ratio in targeted protein analysis. Journal of Pharmaceutical and Biomedical Analysis 2016;123:155–61. https://doi.org/10.1016/j.jpba.2016.02.011.

[29]   Singh DB, Tripathi T, editors. Frontiers in Protein Structure, Function, and Dynamics. Frontiers in Protein Structure, Function, and Dynamics. 1st ed., Singapore: Springer Singapore; 2020, p. 17, 34–6. https://doi.org/10.1007/978-981-15-5530-5.

[30]   Zheng YZ, DeMarco ML. Manipulating trypsin digestion conditions to accelerate proteolysis and simplify digestion workflows in development of protein mass spectrometric assays for the clinical laboratory. Clinical Mass Spectrometry 2017;6:1–12. https://doi.org/10.1016/j.clinms.2017.10.001.

[31]   Bronsema KJ, Klont F, Schalk FB, Bischoff R, Kema IP, van de Merbel NC. A quantitative LC-MS/MS method for insulin-like growth factor 1 in human plasma. Clinical Chemistry and Laboratory Medicine 2018;56:1905–12. https://doi.org/10.1515/cclm-2017-1042.

[32]   Bults P, Bischoff R, Bakker H, Gietema JA, van de Merbel NC. LC-MS/MS-Based Monitoring of in Vivo Protein Biotransformation: Quantitative Determination of Trastuzumab and Its Deamidation Products in Human Plasma. Analytical Chemistry 2016;88:1871–7. https://doi.org/10.1021/acs.analchem.5b04276.

[33]   Bronsema KJ, Bischoff R, van de Merbel NC. High-sensitivity LC-MS/MS quantification of peptides and proteins in complex biological samples: The impact of enzymatic digestion and internal standard selection on method performance. Analytical Chemistry 2013;85:9528–35. https://doi.org/10.1021/ac4015116.

[34]   Heissel S, Frederiksen SJ, Bunkenborg J, Højrup P. Enhanced trypsin on a budget: Stabilization, purification and high-temperature application of inexpensive commercial trypsin for proteomics applications. PLoS ONE 2018;14. https://doi.org/10.1371/journal.pone.0218374.

[35]   Lee MS. Protein analysis using mass spectrometry : accelerating protein biotherapeutics from lab to patient. In: Lee MS, Ji QC, editors. First Edition, Hoboken: John Wiley & Sons; 2017, p. 24, 25, 65, 111, 152, 153, 203–16, 232–3.

[36]   Abe K, Shibata K, Naito T, Karayama M, Hamada E, Maekawa M, et al. Quantitative LC-MS/MS method for nivolumab in human serum using IgG purification and immobilized tryptic digestion. Analytical Methods 2019;12:54–62. https://doi.org/10.1039/c9ay02087j.

[37]   Strohalm M, Hassman M, Košata B, Kodíček M. mMass data miner: an open source alternative for mass spectrometric data analysis. Rapid Communications in Mass Spectrometry 2008;22. https://doi.org/10.1002/rcm.3444.

[38]   Rifai N, Horvath AR, Wittwer CT, Hoofnagle AN, editors. Principles and Applications of Clinical Mass Spectrometry: Small Molecules, Peptides, and Pathogens. sixth edition. San Diego: Elsevier; 2018.

[39]   Sequence Specific Retention Calculator 2010. http://hs2.proteome.ca/SSRCalc/SSRCalcX.html (accessed March 20, 2021).

[40]   Hoofnagle AN, Whiteaker JR, Carr SA, Kuhn E, Liu T, Massoni SA, et al. Recommendations for the generation, quantification, storage, and handling of peptides used for mass spectrometry-based assays. Clinical Chemistry 2016;62:48–69. https://doi.org/10.1373/clinchem.2015.250563.

[41]   Mass Spec Proteases | Trypsin | Protein Digestion | Promega n.d. https://nld.promega.com/products/mass-spectrometry/proteases-and-surfactants/ (accessed March 16, 2021).

[42]   Enzyme Explorer Key Resource: Proteolytic Enzymes Index | Sigma-Aldrich n.d. https://www.sigmaaldrich.com/life-science/metabolomics/enzyme-explorer/learning-center/proteolytic-enzymes.html (accessed March 16, 2021).

[43]   Protein Digestion for Mass Spectrometry | Thermo Fisher Scientific - NL n.d. https://www.thermofisher.com/nl/en/home/life-science/protein-biology/protein-mass-spectrometry-analysis/sample-prep-mass-spectrometry/protein-digestion-mass-spectrometry.html (accessed March 16, 2021).

[44]   Tsiatsiani L, Heck AJR. Proteomics beyond trypsin. FEBS Journal 2015;282:2612–26. https://doi.org/10.1111/febs.13287.

[45]   Chiva C, Ortega M, Sabidó E. Influence of the digestion technique, protease, and missed cleavage peptides in protein quantitation. Journal of Proteome Research 2014;13:3979–86. https://doi.org/10.1021/pr500294d.

[46]   Protein BLAST: search protein databases using a protein query n.d. https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (accessed March 2, 2021).

[47]   Badgett MJ, Boyes B, Orlando R. The Separation and Quantitation of Peptides with and without Oxidation of Methionine and Deamidation of Asparagine Using Hydrophilic Interaction Liquid Chromatography with Mass Spectrometry (HILIC-MS). Journal of the American Society for Mass Spectrometry, vol. 28, Springer New York LLC; 2017, p. 818–26. https://doi.org/10.1007/s13361-016-1565-z.

[48]   Krause ME, Sahin E. Chemical and physical instabilities in manufacturing and storage of therapeutic proteins. Current Opinion in Biotechnology 2019;60:159–67. https://doi.org/10.1016/j.copbio.2019.01.014.

[49]   Grassi L, Cabrele C. Susceptibility of protein therapeutics to spontaneous chemical modifications by oxidation, cyclization, and elimination reactions. Amino Acids 2019;51:1409–31. https://doi.org/10.1007/s00726-019-02787-2.

[50]   Zapadka KL, Becher FJ, Gomes dos Santos AL, Jackson SE. Factors affecting the physical stability (aggregation) of peptide therapeutics. Interface Focus 2017;7. https://doi.org/10.1098/rsfs.2017.0030.

[51]   Donev R. Therapeutic Proteins and Peptides. vol. Volume 112, San Diego: Elsevier; 2018, p. 4–11.

[52] Wilffert D, Reis CR, Hermans J, Govorukhina N, Tomar T, de Jong S, et al. Antibody-free LC-MS/MS quantification of rhTRAIL in human and mouse serum. Analytical Chemistry 2013;85:10754–60. https://doi.org/10.1021/ac4017902.

[53] Tran DT. Engineering Proteases for Mass Spectrometry-Based Post Translational Modification Analyses. Proteomics 2019;19. https://doi.org/10.1002/pmic.201700471.

[54] Woods AG, Darie CC, editors. Advancements of Mass Spectrometry in Biomedical Research. vol. 1140. Second Edition, Cham: Springer International Publishing; 2019, p. 180–6, 589, 620–1. https://doi.org/10.1007/978-3-030-15950-4.

[55] Grangeon A, Clermont V, Barama A, Gaudette F, Turgeon J, Michaud V. Development and validation of an absolute protein assay for the simultaneous quantification of fourteen CYP450s in human microsomes by HPLC-MS/MS-based targeted proteomics. Journal of Pharmaceutical and Biomedical Analysis 2019;173:96–107. https://doi.org/10.1016/j.jpba.2019.05.006.

[56] Li W, Jian W, Fu Y. Basic Sample Preparation Techniques in LC-MS Bioanalysis. Sample Preparation in LC-MS Bioanalysis. First Edition, 2019, p. 3–30.

[57] Wilffert D, Donzelli R, Asselman A, Hermans J, Govorukhina N, ten Hacken NH, et al. Quantitative antibody-free LC–MS/MS analysis of sTRAIL in sputum and saliva at the sub-ng/mL level. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 2016;1032:205–10. https://doi.org/10.1016/j.jchromb.2016.04.041.