

Characterisation of the covalent structure of proteins from biological material by MALDI mass spectrometry – possibilities and limitations

Martin Kussmann and Peter Roepstorff

Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark

Abstract. Matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS) has become a primary tool for the detailed characterisation of the covalent structure of proteins isolated from biological material, mainly because of its following potentials: high sensitivity and specificity, speed of analysis, appropriateness for mixture analysis, high tolerance towards contaminants, and compatibility with separation techniques, e.g., gel electrophoresis. These characteristics enable the structural analysis of proteins even if they are only available in limited amounts and/or in mixtures, and even if the protein preparations contain large amounts of salts, buffers, detergents and denaturants. Additionally, structural data can be generated within a relatively short time.

Whereas X-ray crystallography and multidimensional NMR techniques can provide “absolute” structural data, i.e., a three-dimensional “picture” of the protein of interest, MALDI-MS – especially in combination with selective protein chemistry – yields information on particular aspects of the entire protein structure, e.g., primary structure, active site(s), binding sites, and posttranslational modifications, all of which are often of crucial interest for the understanding of the protein function. Taking into account that protein crystallography and protein NMR studies require large quantities of highly purified sample, MALDI-MS can be even more regarded as a powerful complement in protein structure analysis.

This review aims at describing the state-of-the-art of MALDI-MS for characterisation of proteins from biological material by evaluating its potential and limitations.

1. Introduction

The invention and development of matrix-assisted laser desorption/ionisation (MALDI) [1] and electrospray ionisation (ESI) [2] as very efficient and soft ionisation techniques for mass spectrometric analysis of labile molecules has revolutionised traditional strategies for biomolecule characterisation [3–6]. ESI and MALDI are fundamentally different ionisation techniques, yet they achieve the same final result, namely, the generation of intact gas phase ions via a non-destructive ionisation process. This is a prerequisite to obtain molecular mass information on large involatile and thermally labile molecules such as peptides and proteins. Mass spectrometric protein analyses with ESI- and MALDI-MS – especially combined with selective protein chemistry – have now been refined to a state, in which they are unsurpassed by any other technique for protein structure determination with regard to sensitivity, speed, mass accuracy, and specificity [7,8]. ESI-MS is superior in terms of mass accuracy [9], MS/MS fragmentation studies [10,11], on-line compatibility with LC [12,13] and CZE [14,15], and with regard to the direct study of supramolecular complexes [16,17]. Meanwhile, MALDI-MS provides higher ultimate sensitivity [18,19], higher tolerance towards contaminants [20], and is

well suited for the analysis of complex mixtures [21]. Due to the latter potential, MALDI-MS is extensively used in protein interaction studies by limited proteolysis [22] (protein footprinting) [23,24], and protein surface topology probing [25,26]. Thus, ESI and MALDI represent complementary rather than competitive analytical tools. In the following, this review focuses on the principles, features and applications of MALDI-MS in peptide and protein characterisation.

The protein mass range accessible by MALDI-MS has been extended up to ca. 500,000 D and sensitivities in the low fmol range for peptides and low pmol range for proteins are routinely achieved [27]. The recent introduction of delayed extraction (DE), i.e., a two-stage ion extraction process contrasting the traditional, one-stage continuous extraction, has significantly enhanced the resolution and the accuracy of the mass determination [28]. When analysing peptides and small proteins up to 10,000 Da, resolutions of 5,000–15,000 and mass accuracies of 500–50 ppm are today's routine on high-end MALDI mass spectrometers with time-of-flight (TOF) analysers. By contrast, the resolution of protein molecular ions is dramatically reduced, mainly not because of worse instrument performance at high mass ranges, but due to the protein ion chemistry: large proteins tend to form multiple adduct ions, e.g., with alkali salts, and, upon desorption, they undergo fragmentation more extensively than peptide ions. While encompassing the same bond types as peptides, proteins include statistically more of the labile bonds accounting for possible fragmentation sites.

The combination of proteolytic digestion, MALDI-MS peptide mapping (i.e., the mass spectrometric analysis of the proteolytic mixture) and the use of this information in protein data bases have proven superior in fast and specific protein identification [29,30]. The compatibility of two-dimensional gel electrophoresis (2D-SDS-PAGE) with MALDI-MS analysis of in-gel generated proteolytic peptides is now routinely and efficiently used in protein identification from complex mixtures [31].

The complementary application of different proteases and the subsequent MALDI-MS peptide mapping analysis typically yields protein sequence coverages of 70–100% and thus provides a fast approach to direct protein primary structure characterisation. The approach of using specific proteases and the analysis of their cleavage products has been extended to the application of enzymes and reagents capable to cleave, transfer, or modify posttranslational modifications, making the latter also amenable to structural characterisation by mass spectrometry. This analytical concept has been successfully applied to the determination of location and structure of glycan moieties in glycoproteins [32] and to the characterisation of phosphorylation sites in regulatory proteins [33].

2. MALDI-MS characterisation of proteins from biological material

2.1. Principles and recent developments in MALDI-MS instrumentation

Figure 1 shows a schematic drawing of a MALDI mass spectrometer. The MALDI process generates gas phase ions by UV or IR laser vaporisation of a solid co-crystallised matrix/analyte mixture. The matrix, usually a small aromatic organic molecule, has a strong absorption at the laser excitation wavelength. The explosion-like formation of an ion plume effects the desorption of the sample into the gas phase and, thus, their readiness for mass analysis [34].

MALDI has been primarily used in combination with time-of-flight (TOF) analysers, in which ions are separated according to their travel time along a certain distance dependent on their mass : charge (m/z) ratio. The TOF analysers have the advantages of a theoretically unlimited m/z range, high ion transmission, simplicity of design and use, and relatively low costs [35]. A major limitation of conventional TOF analysers has been their poor resolving power, i.e., their ability to separate adjacent

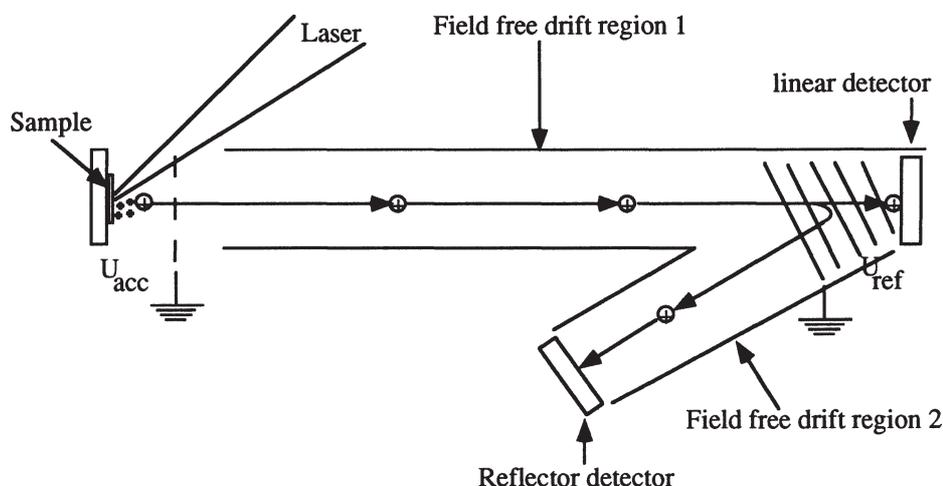


Fig. 1. Schematic representation of a MALDI mass spectrometer. Ions are generated by laser vaporisation of a solid co-crystallised matrix/analyte mixture. The explosion-like formation of an ion plume effects the desorption of the sample into the gas phase. The ions are accelerated by a 20–30 kV potential (U_{acc}). After having passed the first field-free drift region, they directly hit the linear detector (linear mode) or, alternatively, they pass the reflector ($U_{ref} = U_{acc} + \text{ca. } 10\%$) and travel to the reflector detector (reflector mode).

masses. Significantly improved mass resolution and an associated increase in mass accuracy have been demonstrated for MALDI-TOF-MS by implementation of reflector analysers [35], delayed extraction [28,36] (DE, also known as pulsed ion extraction [36] or time lag focusing [37]). More recently, ion trap (IT-MS) [38,39], Fourier-transform ion-cyclotron resonance (FT-ICR) analysers [40,41] and the introduction of ion trap-TOF hybrid-systems [42,43] have proven powerful MALDI mass analysers.

The implementation of the post-source decay (PSD) technology [44] enabled the study of ion fragmentation in MALDI-MS. The parent ion is “filtered” from the peptide ion mixture by the timed-ion selector, which “excises” a narrow time-of-flight window from the total TOF range. The fragment ion spectrum reflects the spontaneous metastable decay of the parent ion in the first field-free region. In linear TOF, the fragment ions are detected at the same “arrival time” as the molecular ions, but a reflecting TOF analyser has the capacity to resolve PSD-produced ions by stepwise reduction of the reflector voltage [44]. Recently, a curved-reflector instrument was introduced allowing the simultaneous acquisition of the entire PSD spectrum [45]. Post-source decay therefore allows for fragmentation and sequencing studies on a single peptide without previous separation of the analyte mixture. The implementation of a collision cell into the first field-free region, i.e., into the flight path prior to the reflector, significantly enhances the fragmentation yield in a PSD experiment on a MALDI-TOF instrument [46].

2.2. Sample preparation for MALDI-MS

The sample preparation is the crucial procedure in MALDI-MS analysis of peptides and proteins. Preparing a peptide or protein sample includes two steps: the first represents the isolation and the purification of a single component or a mixture, free of contaminants such as buffers, salts, detergents, or denaturants. The second step comprises the sample processing on the MALDI target, i.e., choice of matrix, matrix and analyte concentration, pH adjustment, crystallisation conditions, use of additives, and on-target sample clean-up.

Addressing sample preparation with in the context of this review, the authors would like to refer to an own study [47] encompassing a collection of experiences with regard to numerous MALDI-MS sample preparation techniques in terms of their suitability for different peptide and protein analytes. The importance of matrix selection, matrix and analyte concentration, pH adjustment, crystallisation conditions, and use of additives were evaluated. The conclusions of this study can be summarised as follows:

There is no universally applicable sample preparation for a broad variety of analytes. Rather, it is necessary to specifically adapt the sample preparation to the analyte properties. With regard to matrix selection, 2,5-dihydroxybenzoic acid [48] and a variety of cinnamic acid derivatives, e.g., ferulic acid, sinapic acid [49], and α -cyano-4-hydroxycinnamic acid (HCCA) [50] are frequently used for MALDI-MS analysis of peptides and proteins. HCCA is preferred for peptide mapping because it yields in general the best protein sequence coverage, i.e., its ionisation and desorption properties are suitable for a broad variety of peptides. The preparation of a thin, homogeneous layer of small HCCA crystals by the fast-evaporation or thin-layer method resulted in improved resolution and mass accuracy and enhanced sensitivity extended to the low attomole range [51]. However, the complementary use of sinapic acid (SA), 2,5-dihydroxybenzoic acid (DHB) and 2,4,6-trihydroxyacetophenone (THAP) [52] is recommended, since these matrices may generate additional peptide ions and thus may yield additional sequence coverage, possibly not obtainable by exclusive use of the matrix HCCA. Sinapic acid serves best for MALDI-MS analysis of proteins. DHB is the first choice for analysing glycopeptides and proteins with MALDI-MS, because – especially in contrast to HCCA – it does not cause fragmentation of Asn-N- and Ser/Thr-O-glycosidic bonds upon desorption and hence yields molecular ions of intact glycopeptides and proteins. A special feature of the matrix DHB applied for glycopeptides is the purification effect upon crystallisation [53,54]. DHB applied in a dried-droplet sample preparation forms large crystals on the sample rim and smaller crystals in the central sample area. The rim crystals predominantly yield abundant MH^+ ions, whereas less intense alkali adduct signals are obtained from the central sample area.

A substantial obstacle in the analysis of proteins and peptides from biological sources can be the contamination with high amounts of salts, buffers, synthetic polymers, detergents, and denaturants. This is in particular valid for samples derived from gel electrophoretic separations. In these cases, the “sandwich” preparation method [47] offers significant advantages due to its compatibility with extensive washing procedures on the MALDI target. Moreover, the inclusion of nitrocellulose [47,55] into the HCCA matrix solution efficiently suppresses the formation of salt adduct and polymer ions, which can severely deteriorate the MALDI spectrum quality. In addition, even the traditional dried-droplet sample preparation [47], e.g., with the matrices HCCA, SA, or DHB, is amenable to careful on-target washing steps. If on-target sample clean-up is not successful, a microscale purification procedure [47] prior to the MALDI-MS analysis is recommended. By loading the sample on reversed-phase material packed in pipette tips (i.e., disposable micro-columns) and by eluting the sample directly onto the MALDI target, the sample can be efficiently purified and almost quantitatively recovered.

In the case of very limited protein amounts and if a fast verification of preliminary structural data is desired, the performance of on-target reactions can be considered. For instance, the on-target proteolytic digestion with trypsin can be used as an “*in-situ* peptide mapping” procedure [56], and the on-target dithiothreitol (DTT) reduction of a protein or its proteolytic digest can provide a fast elucidation of disulfide bond patterns [56]. Alternatively to on-target digestions, the previously mentioned micro-columns can be packed with immobilised protease media and an on-column digestion can be performed, providing rapid and effective digestion and high sample recovery [57].

2.3. MALDI-MS of intact proteins

MALDI mass spectra of proteins are generally recorded in linear positive-ion mode, since proteins with molecular masses exceeding 10,000 Da fragment probably before passing the reflector and, therefore, may not yield intact molecular ions in reflector mode. Moreover, the positive-ion abundances are in general higher than those of the negative ions.

The MALDI mass spectrometric characterisation of intact proteins yields in fact much more information than “just” the molecular mass. In the case of a MALDI-MS analysis of a protein with an expected identity based on previous results like known cDNA, antibody recognition, or gelelectrophoretic analyses, the determined molecular mass may further consolidate or perhaps reject the assumed protein identity. On the other hand, if the protein of interest is unknown yet, the molecular mass can be used in a protein data base search (see Section 2.4.4) with a more restricted protein mass window to be searched in and it possibly reduces false positive protein identifications.

A further crucial information contained in a MALDI mass spectrum of an intact protein is the purity and homogeneity of the isolated material. The evaluation of purity and the detection of trace contaminants are important aspects of quality control. Due to the fact that ion abundances – especially of peptides and proteins – decrease with increasing molecular mass, the majority of low molecular weight contaminants, even when present in very low amounts, will be registered by direct MALDI-MS analysis of the protein sample. The presence of several proteins normally results in the detection of several molecular ions. In the case of protein isoforms, the molecular ions will fall into a narrow mass range around the molecular mass of the expected species. Therefore, the MALDI mass spectrum can provide the information, whether the analyte represents a protein mixture or one defined species. The simultaneous detection of several protein species may be, however, hampered by protein ion suppression effects resulting in failed detection of some species in the mixture. Protein truncation, degradation, or modification like oxidation, caused by biological processes or by preparation failures as insufficient addition of protease inhibitors or unsuitable storage of the protein, may also be revealed by the detection of additional molecular ions close to the expected mass of the intact species.

The detection of molecular ions of large and/or modified proteins by MALDI-MS is partially dependent on the detector sensitivity and on the ion transmittance in the MALDI mass spectrometer. Nonetheless, the signal-to-noise ratio is mainly determined by the level of chemical noise. Moreover, the ionisation and desorption efficiency for protein (and peptide) analytes is mainly influenced by the sample preparation technique (see Section 2.2).

Protein glycosylation often represents a heterogeneous modification meaning that instead of one defined protein species, rather a very complex mixture of differently glycosylated proteins is to be analysed. Glycoprotein populations often result in closely adjacent and low abundant protein ions according for the particular subtypes and sometimes they are difficult to be detected and/or resolved at all by MALDI-MS. Furthermore, protein ions are in general subjected to fragmentation and they readily form multiple adduct ions. The differentiation of proteins with closely coinciding molecular masses is a function of the instrument's resolution and of the protein-ion stability and ion chemistry. Assuming an approximate mass difference of 160 Da for a hexose residue and a glycoprotein molecular mass of 80 kDa, the differentiation between the glycoprotein subforms by direct MALDI-MS analysis theoretically requires an instrument resolution of 500. This value ranks far below the resolutions of up to 15,000 achievable today in MALDI-DE-reflector-TOF mass spectrometers, but, as mentioned above, they are of theoretical nature since the protein ion stability and chemistry influences much more the obtainable resolution. The determination of the peak width at-half-height partially circumvents the

limitation of reduced resolution of protein ions in MALDI-MS. This at least gives an approximation of the heterogeneity and/or the modification of the protein.

Figure 2 shows a typical example, in which the direct MALDI mass spectrometric analysis of intact proteins yields important structural information. In this case, MALDI mass spectra of the glycoprotein Glucoamylase (GA) expressed in different organisms are compared [58]. Figure 2(a) shows the recombinant GA purified from *P. pastoris*, (b) the recombinant GA from *S. cerevisiae*, (c) the recombinant GA from *A. niger*, and (d) the native GA from *A. niger*. The different molecular masses and the varying peak widths reflect different extents of glycosylation and different degrees of heterogeneity due to protein expression in different systems. The molecular masses determined by MALDI-MS corresponded closely to the calculated masses based on the amino acid sequence and the results of the quantitative analysis of neutral hexose [58].

Quantification in MALDI-MS is difficult, in fact, MALDI mass spectrometry *per se* is a non-quantitative analytical technique. The ion abundances and peak areas are highly dependent on the nature of the protein. Furthermore, the signal intensities in the MALDI-MS analysis of the same analyte vary significantly with the choice of matrix, with the sample preparation applied, and with the presence or absence of contaminants (salts, detergents, denaturants, etc.). However, using an internal standard, i.e., spiking the analyte with a protein of well-defined molecular mass and using it as an internal calibrant, can render a semi-quantitative estimation of the actual amount of protein present in the sample. The use of internal standardisation is often limited by possible suppression of the analyte ion(s). This can be partially circumvented by external standardisation, i.e., by preparing defined amounts of a protein standard exactly with the same sample preparation method as the analyte is prepared, and by acquiring standard and analyte spectrum at the same instrument settings. Moreover, a calibration file can be derived from the standard spectrum, which can be used to calibrate the analyte spectrum (external calibration). However, a recent study on the quantification of gliadins in gluten products by MALDI mass spectrometry has shown a certain potential of MALDI-MS for protein quantification [59].

2.4. MALDI-MS peptide mapping: Protein identification and primary structure characterisation

The origin of the protein sample predominantly determines the proteolytic digestion procedure, the performance of which is the basis for a following high-quality peptide mapping analysis. Detergent- and denaturant-free volatile buffers like ammonium bicarbonate are most compatible with MALDI-MS analysis. Exceptions are non-ionic detergents, e.g., *n*-octyl-glucoopyranoside (OGP), the latter often improving the quality of MALDI-MS peptide maps, probably by solubility enhancement of hydrophobic peptides [60]. For these reasons, it is recommended to exchange the original protein isolation buffers towards these volatile buffers, that, upon on-target sample acidification, can be directly used in MALDI-MS sample preparations.

2.4.1. Proteolytic digestion of proteins in-solution and on-column

If the protein isolated from biological tissue is available in solution, two possibilities of proteolytic processing open up: first, the direct addition of protease solution, perhaps after buffer exchange towards digestion-compatible buffers (especially, if the original solution contains stabilising and preserving chemicals possibly inhibiting the protease) and, second, submission of the protein solution to on-column digestion with immobilised enzyme media (mentioned in Section 2.2). The latter variant offers advantages like more rapid and more effective digestion, high sample recovery, as well as a reproducible digestion rate by control of the media : substrate ratio and the loading and elution times

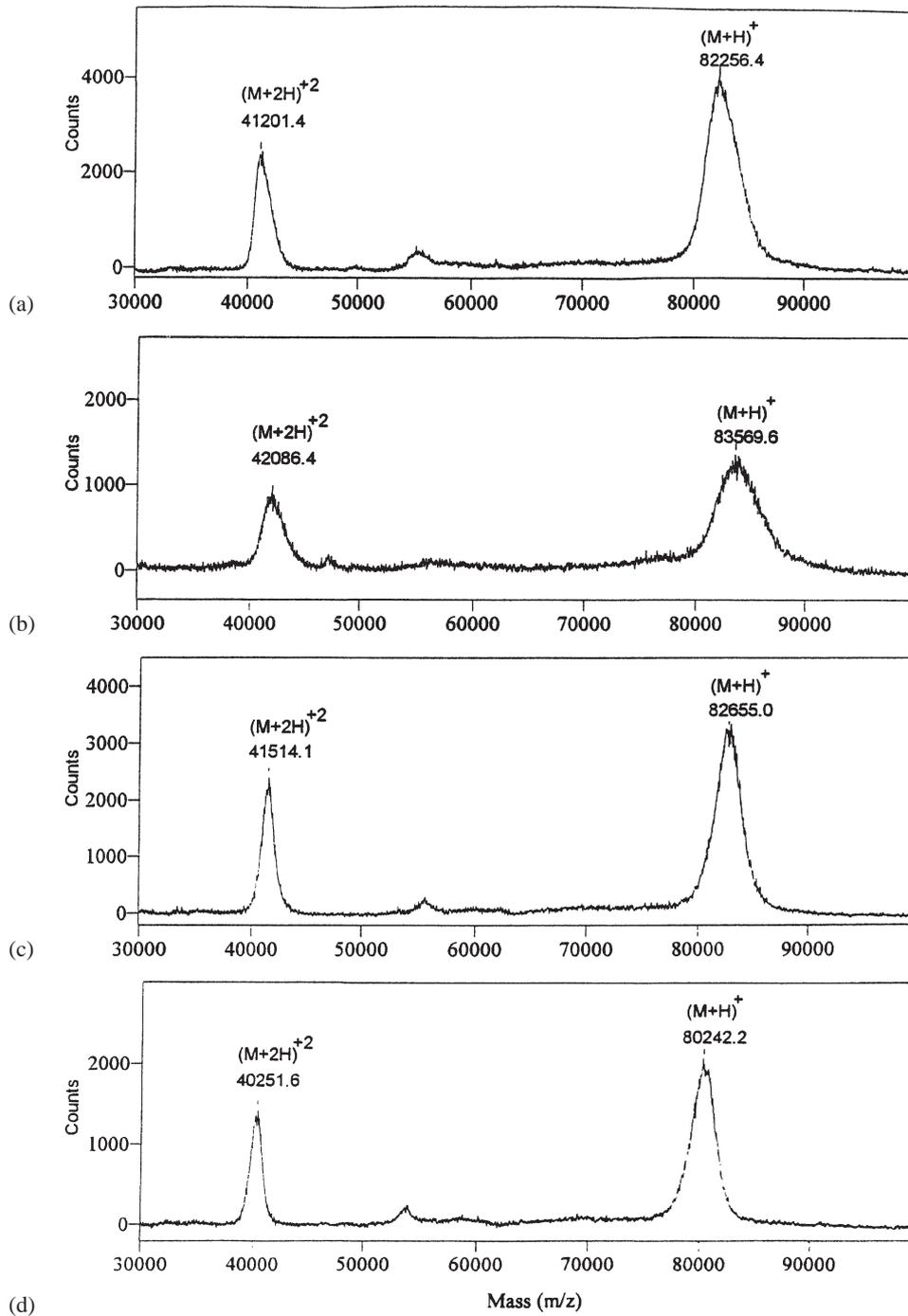


Fig. 2. MALDI mass spectra of the glycoprotein Glucoamylase (GA) expressed in different organisms. (a) Recombinant GA from *P. pastoris*, (b) rec. GA from *S. cerevisiae*, (c) rec. GA from *A. niger*, (d) native GA from *A. niger*. The different molecular masses and peak widths reflect different glycosylations and different degrees of heterogeneity, respectively. Reproduced from Fierobe et al. [58].

of the protein solution [61–63]. However, media with the specifically required immobilised enzyme are not always at hand.

There is a variety of endoproteases commercially available, all of them in sequencing grade quality and some of them with added inhibitors to prevent autoproteolysis. The last two quality features are important, because peptide contaminants, lack of specificity due to enzyme impurities, and autoproteolysis may severely interfere with the peptide maps derived from the protein sample. Commonly used specific proteases are trypsin, and the endoproteases Glu-C, Lys-C, Arg-C and Asp-N. Whereas trypsin and the endoproteases Glu-C, Lys-C and Arg-C effectively and specifically digest almost any protein, if it is amenable to proteolysis at all, the performance of Asp-N often turns out to be more substrate-dependent [64]. In order to make a protein susceptible to efficient proteolytic degradation, it is often necessary to first denature the protein with urea, guanidinium hydrochloride, or by short exposure to high temperature. Disulfide bond reduction preferably carried out with DTT might also be required. In these cases, it is important to check the tolerance of the protease towards additives and reagents. A buffer exchange after the pre-treatment of the protein might be necessary.

Sometimes, it is convenient to use less specific proteases like pepsin or chymotrypsin. Pepsin for example has a pH optimum around 2–3 and cleaves rather unspecifically at N- and C-termini of hydrophobic residues (i.e., amino acids with aliphatic or aromatic side chains). Hence, the digestion can be carried out under denaturing conditions (e.g., in 2% acetic acid) and no buffers and salts are required, rendering the procedure highly compatible with directly following MALDI-MS analysis. If the reaction conditions are kept constant, a high reproducibility of peptic digest patterns can be reached [65].

If partial sequence information is needed (e.g., for confirmation of peptide assignments) and if post-source decay data are either not available or ambiguous, exoprotease application and MALDI-MS time course monitoring of the digestion can be a powerful tool. The carboxypeptidases CPase A, B and Y are commercially available, each of them or as an “enzyme cocktail”, and they can serve as efficient sequencing reagents. Especially, if only small sample amounts are at hand or a fast sequence identification is needed, the exoproteases can be applied on-target, as previously mentioned for endoproteases (see Section 2.1) [66]. In general, the MALDI-MS analysis of sequence ladders generated either by exopeptidase digestion or by Edman degradation [67,68] can be a viable alternative to mass spectrometric fragmentation techniques. On the other hand, blocked N- or modified C-termini can impair Edman or exopeptidase sequencing and MS/MS may be the only access to peptide sequence information.

2.4.2. Proteolytic digestion of proteins separated by gelelectrophoresis

MALDI-MS is the method of choice for peptide mapping of proteins isolated by PAGE due to its sensitivity and, in particular, because of its superior tolerance towards contaminants encompassed by electrophoretic procedures, for example, stains, detergents, denaturants and salts. Even considering these qualities of MALDI-MS, it is important to minimise sample handling and transfer (e.g., by in-gel or on-membrane digestion) and, if possible, to use chemicals and solvents compatible with mass spectrometric analysis.

Given a protein embedded in an excised gel plug, that stems from a one or two-dimensional PAGE separation, there are three alternatives at hand to subject the protein to a proteolytic digest and a following mass spectrometric peptide mapping procedure:

1. The digestion can be directly carried out in the gel matrix and the peptides can be recovered by subsequent passive elution [31];

2. The protein can be first electroeluted and then subjected to conventional digestion in solution [69];

3. The protein can be blotted onto a membrane and subsequently digested on this membrane [70,71]. It is a general experience that the yield of electroeluted and blotted proteins varies a lot with the protein's size and hydrophobicity, i.e., the larger and the more hydrophobic the protein, the poorer the yields after electroelution or blotting. However, advantageously, both electroelution and blotting of the protein may enable the mass spectrometric characterisation of the intact protein [70,72], although the direct MALDI-MS analysis of electroeluted proteins is possibly impeded by broad, unresolved and/or low abundant molecular ions. This result can be attributed to alkylation with acrylamide, inefficient removal of SDS, remaining stain and contaminants introduced during blotting or electroelution. Therefore, in our laboratory, we prefer *in-situ* digestion of the protein in the gel matrix and subsequent passive elution of the generated peptides [73]. As a complement to MALDI-MS peptide mapping of coomassie stained protein bands, Mann et al. have furthermore developed a more sensitive in-gel digestion procedure even compatible with silver staining [74]. An intrinsic drawback of the *in-situ* digestion is the limited selection of proteases, that work in-gel as reliably as in solution. Trypsin is the most specific and reliable protease for in-gel digestions. Endoprotease Glu-C (V8 protease) often also exhibits a sufficient specificity when applied in a gel plug. By contrast, most of the other commonly used endoproteases like Asp-N, Lys-C, Arg-C, and chymotrypsin, that all exhibit effective and specific cleavage of most proteins in solution, reveal a substantially reduced activity and specificity when used in an in-gel digestion. This is partially due to the size of the enzymes. Larger molecules with less compact and robust tertiary structures, like endoprotease Asp-N, do probably not efficiently penetrate the gel matrix and thus do not have proper access to the substrate and, therefore, they do not exhibit accurate enzymatic work in the gel.

Another typical and often advantageous feature of *in-situ* digestions is limited proteolysis compared to in-solution procedures, meaning that a lot of the potential cleavage sites may remain unaffected. As a consequence, several larger proteolytic fragments will show up in the MALDI-MS peptide map, possibly giving rise – together with smaller fragments derived from quantitative cleavage – to improved protein sequence coverage due to the observation of overlapping and/or adjacent peptides.

2.4.3. Protein primary structure characterisation

Figure 3 resumes mass spectrometric approaches for protein primary structure characterisation. Given a protein of interest with a known or expected translated cDNA or amino acid sequence, a computer-based theoretical digest is first performed with different proteases in order to check for most convenient fragment generation with regard to the specific structural aspects to investigate. A number of studies [75] have been carried out using protein sequence databases and computer programmes [76–78] handling proteolytic procedures and other protein chemical (e.g., amino acid side chain modification) and mass spectrometric (e.g., MS/MS) techniques.

If the sample amount allows for it, it is advisable to use several proteases to confirm or perhaps discard a protein identification by different peptide maps. This is even more recommendable, if the protein sample is available in solution, since then, there is a broad variety of specific and complementing proteases at hand. Moreover, endoprotease arrays, i.e., sequential application of two or more proteases to the same sample, and the intermediate MALDI-MS analysis can be very useful in order to further cleave very large peptides possibly problematic to be assigned unambiguously, and to confirm peptide assignments deduced from previous digests.

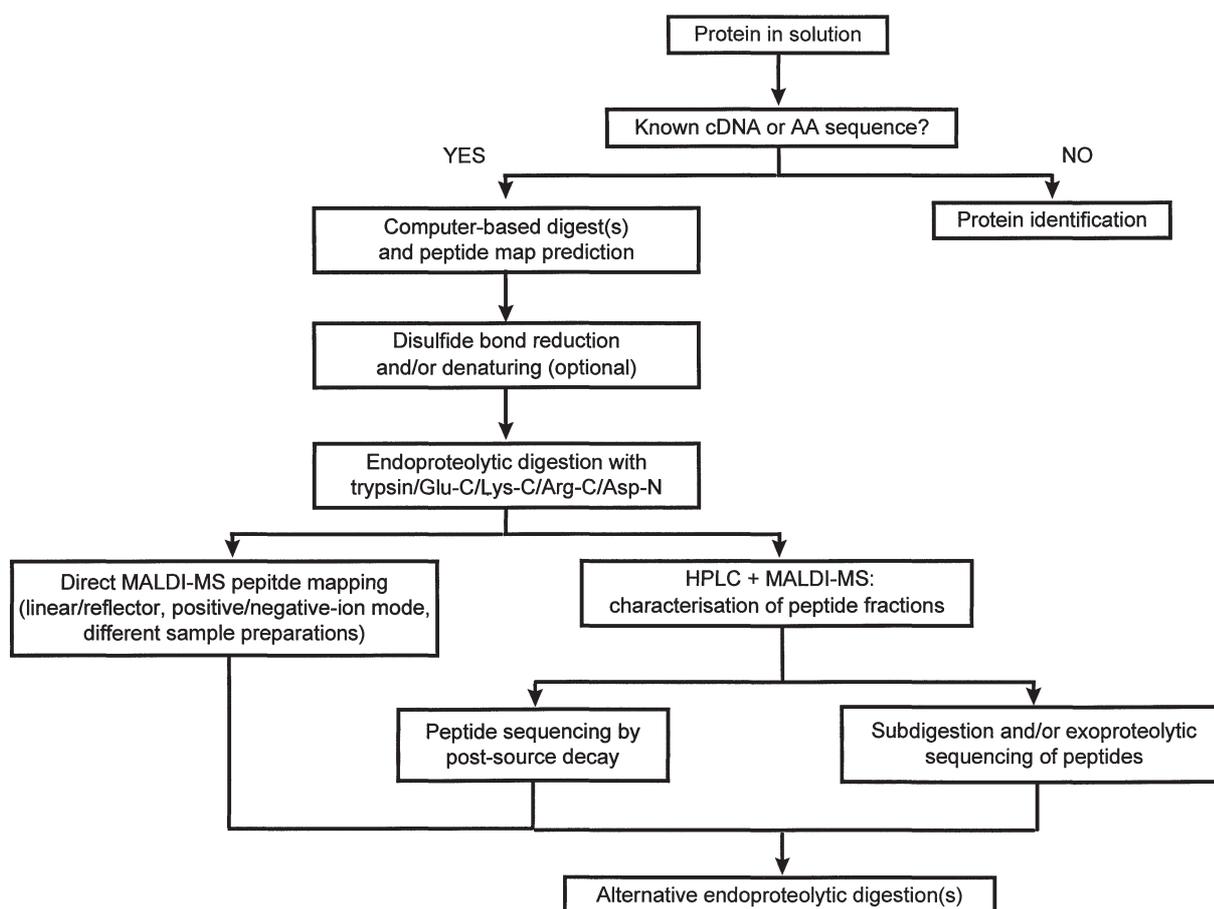


Fig. 3. Mass spectrometric approaches to protein primary structure characterisation.

As the use of different proteases supplies with complement peptide mapping data, so does the application of different matrices and sample preparations (see Section 2.2). The different mass spectrometric operation modes, i.e., positive- and negative-ion, linear and reflector mode, can as well generate complementary peptide mapping data. Most peptides yield positive and negative ions, the first being more intense than the latter. Some very acidic peptides may be, however, exclusively detected in negative-ion mode and especially under those circumstances, the acquisition of a peptide map in negative-ion mode pays off. The reflector mode in MALDI-MS provides highest resolution, typically monoisotopic at m/z 5,000, and, followed by external, internal, or instrumental default calibration, a very good mass accuracy (0.01% or better). The achievable mass accuracy in linear mode falls into the range of 0.05–0.01%. On the other hand, the reflector mode sensitivity in the higher peptide mass range (5,000–10,000 Da) is significantly lower than in linear mode, caused by a decreasing transmittance of the reflector for high mass ions and by an increased fragmentation of the latter.

Typically, 10–100 fmol of peptide mixture yield a good MALDI mass spectrum, i.e., a large number of abundant peptide molecular ions. A sequence coverage of 30–50% is in most cases sufficient in order to identify an unknown protein provided the species is registered in the protein data base (see Section 2.4.4). 20 pmol of protein are sufficient to test and optimise different proteolytic procedures.

By these means, the sequence coverage can be pushed from typically 30–70% achieved by one standard peptide mapping approach, e.g., tryptic digest analysed with HCCA as the matrix in positive-ion mode, to 90–100% for proteins with molecular masses below 100,000 D.

In Figs 4 and 5, the usefulness of complementary peptide mapping strategies, i.e., the advantage of applying different proteases and different MALDI-MS acquisition modes, is illustrated. Figure 4 shows MALDI peptide maps of Neurolin, a glycoprotein with an apparent molecular mass of 86 kDa [64]. All spectra were acquired with delayed extraction. The spectrum of a *tryptic* digest, recorded in *reflector positive-ion* mode is shown in Fig. 4(a). The full spectrum is complemented by an enlargement of the 2500 D range displaying monoisotopically resolved molecular ions of the native and the oxidised tryptic peptide 199–220, containing Met²⁰¹, as given in Fig. 4(b). Figure 4(c) shows the *negative-ion* MALDI mass spectrum of the same tryptic digest recorded in *linear* mode. Figure 4(d) finally shows a MALDI mass spectrum of a *V8-proteolytic* (endoprotease Glu-C) digest acquired in *linear positive-ion* mode. The results of these complementary peptide mapping procedures are summarised in Fig. 5, in which the primary structure of Neurolin and the total sequence coverage achieved by tryptic and V8-proteolytic mass spectrometric peptide mapping is given [64]. In this case, the entire sequence could be covered. The underlined stretches were identified by *both* tryptic and Glu-C peptide mapping; non-underlined partial sequences were *only* found as tryptic peptides. Tryptic and Glu-C cleavage sites identified by corresponding partial peptides are highlighted in bold face. Most of the peptides revealed positive and negative ions, but there were a few very acidic peptides exclusively detectable in negative-ion mode and thus giving additional peptide mapping data. Moreover, peptides in the high mass range (>4000 Da) could only be detected in linear mode, whereas the mass accuracy for peptides in the medium and low mass range (<4000 Da) was found superior in reflector mode providing monoisotopic resolution up to m/z 3000.

Peptide sequence data can be generated by fragmentation experiments with post-source decay (PSD, see Section 2.1). In the context of a detailed primary structure characterisation, partial sequence information can be efficiently used to verify peptide assignments, particularly in those cases, where several “theoretically matching” proteolytic peptides coincide very close in their mass. The PSD-generated fragment ions derived from peptides are found to be predominantly sequence ions [79] and hence PSD allows for sequencing of single peptides without previous separation of the analyte mixture. Due to the restriction on the spontaneity of the process, the yield of fragment ions is highly dependent on the nature of the parent ion. As mentioned in Section 2.1, this limitation can be partially overcome by the installation of a collision cell in the first field-free region, i.e., prior to the reflector [46]. Furthermore, as there is no control of the energetic input on the parent ion like in collision-induced dissociation (CID) [80–82] and ion trap [38,39] technique (most commonly coupled to electrospray sources), usually many different kinds of fragments (a-, b-, c-, x-, y- and z-ions [83]) occur simultaneously. The restriction to spontaneous ion decay and the complex fragmentation pattern render the PSD technology not very suitable for sequencing unknown peptides. However, in the cases where confirmation of expected peptide sequences or differentiation of a few possible candidates is needed, PSD is a powerful means in peptide sequencing. It is noteworthy, that an analytical approach including hydrogen/deuterium exchange facilitates the spectra interpretation [84].

As the peptide mixtures increase in complexity, the drawbacks of direct mixture analysis by MALDI-MS become more perceptible. Suppression effects may result in detection failure especially of modified peptides and can hence effect incomplete sequence coverage. Further reasons for reduced sequence coverage are unresolved peptides closely coinciding in mass and interference with matrix ions in the case of peptides below 700 D [85,86]. These limitations may even persist when using different sample preparation methods and MALDI-MS acquisition modes.

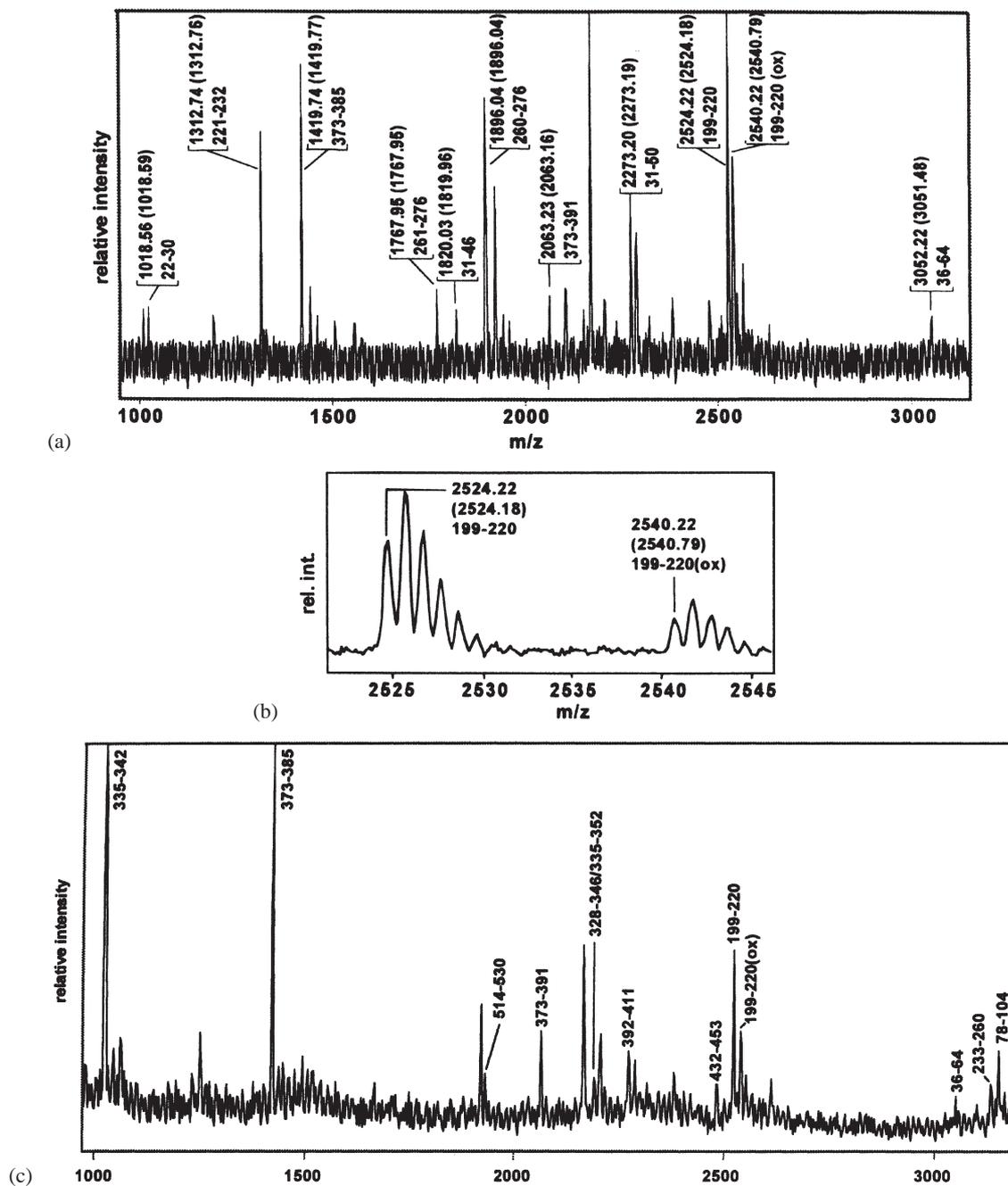


Fig. 4. MALDI mass spectrometric peptide mapping of Neurolin, a glycoprotein with an apparent molecular mass of 86 kDa. (a) MALDI mass spectrum of a tryptic digest of Neurolin, recorded with delayed extraction in reflector positive-ion mode. Peak assignment: found m/z first, calculated m/z second in brackets; sequence position third. (b) Enlargement of the 2500 D range: monoisotopically resolved molecular ions of native and oxidised tryptic peptide 199–220, containing Met²⁰¹. (c) Negative-ion MALDI mass spectrum of the same tryptic digest (linear mode, delayed extraction). Tryptic peptides found are assigned.

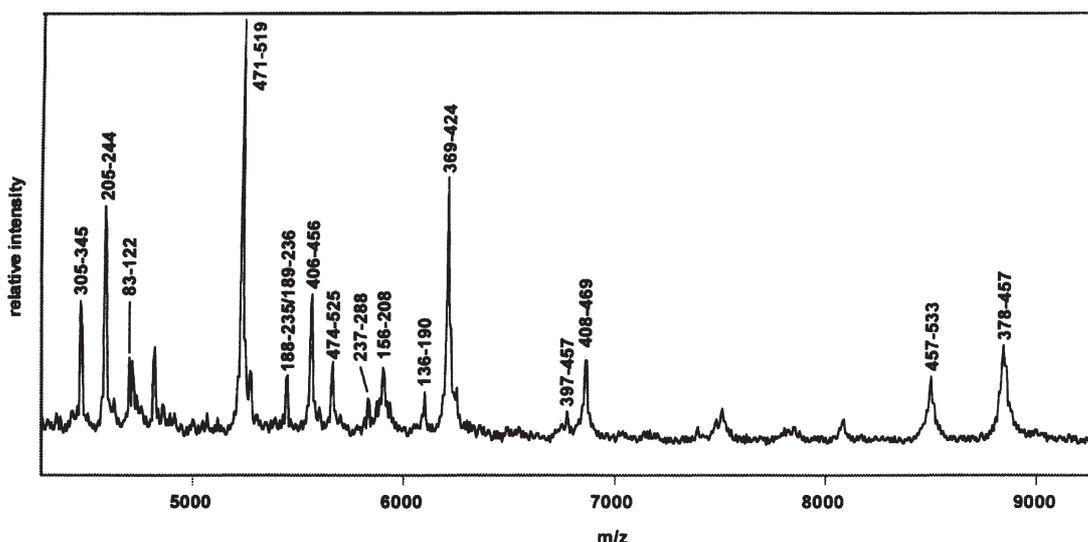


Fig. 4 (continued). (d) MALDI mass spectrum of a V8-proteolytic (endoprotease Glu-C) digest of Neurolin (linear positive-ion mode, delayed extraction). V8-proteolytic peptides found are assigned. Reproduced from Kussmann et al. [47].

VGTVIGLYGETIVVPCNDG**T**KKPDGLIF**T**KWKYVKDDGSPGDLLVKQAO
KDEATVSATDGYKSRVSI**A**ANSSLLI**A**RGLADORV**F**TCMVVS**F**TNLEE
YSVEVKVHKKPSAPVI**K**NNAKEL**E**NGKLTOLGECVVENANPPADLIWKK
 NNQTLVDDG**K**TIIITSTIT**K**DKITGLSSTSSRLOYTARKEDVESOFTCT
AKHVMGPDVSE**E**PE**S**FP**I**HYPTEKVSLQVVSQSP**I**REGEDVTLKCOADG
 NPPPTSFNFNIK**G**KKVTV**D**KDVYTLTGVT**R**ADSGIYKCSLLDNDVMES
 TOFVTVSFLDVSLTPTGKVLKNVGENLIVSLDKNASSE**E**AKVTWTKDNRK
LDKLPDFSKLTYSDAGLYVCDVSI**E**G**I**KRSL**S**FELTVEGIPKITS**L**TKH
RSSDGKHKVLT**C**E**A**EGSPKPDVQWSVNGT**N**DEVS**Y**NNGKATYKLT**V**VP**S**
KNLTVSCLVT**N**KL**E**DT**K**EISVFSQ**N**EDG**T**EOAKV**I**VGIVVGLLVAA**A**
LVGLIYWY**I**KK**T**ROGSW**K**T**G**E**K**EAGT**S**E**S**KK**L**E**N**NHKPDV⁵³³

Fig. 5. Primary structure of Neurolin and sequence coverage achieved by tryptic and V8-proteolytic mass spectrometric peptide mapping. The entire sequence could be covered. Underlined stretches were identified by *both* tryptic and Glu-C peptide mapping; non-underlined partial sequences were *only* found as tryptic peptides. Tryptic and Glu-C cleavage sites identified by corresponding partial peptides are highlighted in bold face. Reproduced from Kussmann et al. [47].

Alternatively to the direct MALDI mass spectrometric analysis of the proteolytic peptide mixture, the digest can be separated via HPLC in order to subject the individual fractions to MALDI-MS. This strategy can provide information on additional peptides, possibly not found in a mixture analysis. Furthermore, the individual treatment of isolated peptides with further proteases or other enzymes allows for more detailed structural characterisation. The PSD analysis of a peptide molecular ion is also facilitated in the case of an isolated species, since the creation of an abundant parent ion – effected by the timed-ion selector – from a complex peptide mixture with possibly closely adjacent other peptides can be difficult. In order to render the coupling of HPLC and MALDI-MS more convenient, there are now on-line elution/on-target sample preparation systems commercially available, in which the eluting peptide fraction is directly spotted onto the MALDI target and automatically mixed with matrix solution.

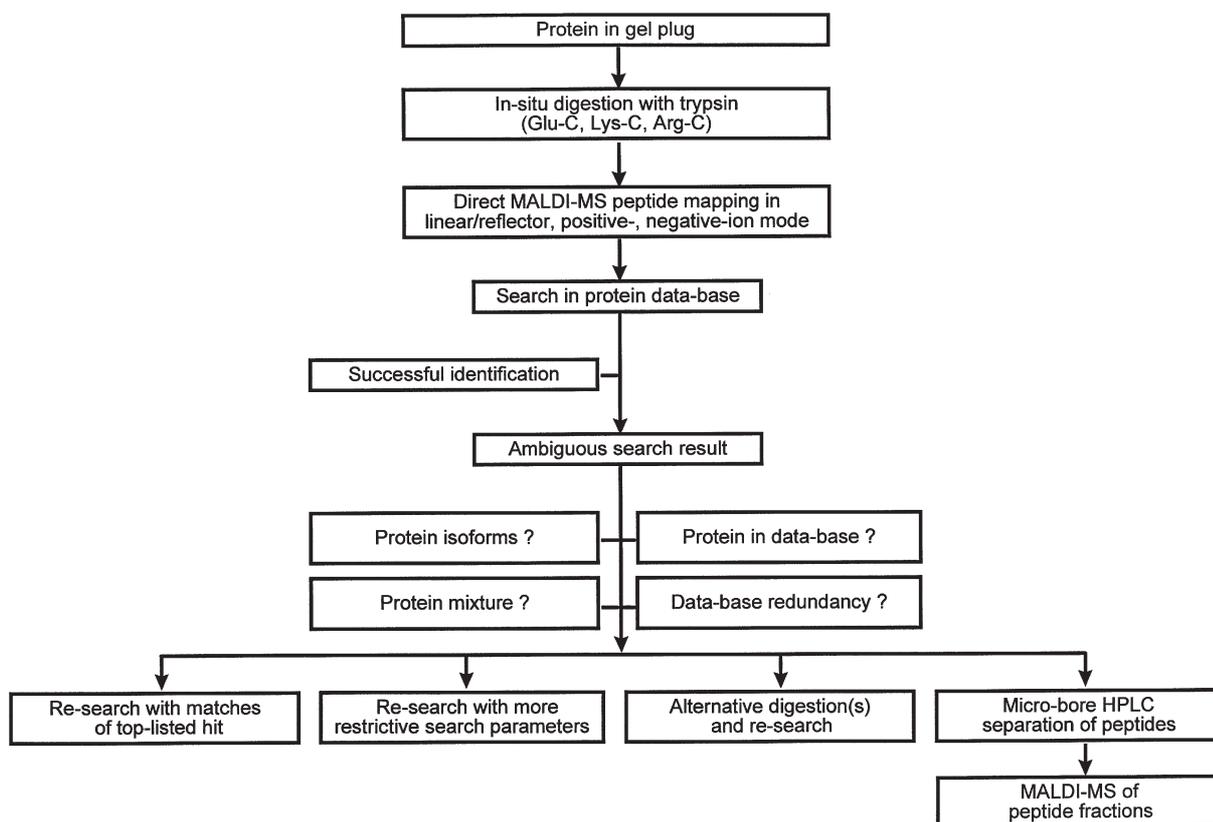


Fig. 6. Mass spectrometric strategy for the identification of a protein isolated by gel electrophoresis.

2.4.4. Protein identification by 2D-SDS-PAGE and MALDI-MS

The strategy for protein identification by MALDI-MS peptide mapping is illustrated in Fig. 6. These mass spectrometric approaches are increasingly combined with two-dimensional gel electrophoresis and by in-gel proteolytic digestion [87]. Two different gel bands in a 1D SDS-PAGE analysis typically represent two different protein species. In a two-dimensional gel-electrophoretic analysis, combining separation according to the molecular weight (PAGE dimension) and the isoelectric point (IEF dimension), two protein bands only differing in the IEF dimension likely stand for two isoforms of the same protein, e.g., for one protein in different phosphorylation states.

As mentioned in Section 2.4.3, a sufficient number of abundant peptide molecular ions accounting for a sequence coverage of 30–50% and enabling protein identification can be generated from low fmol amounts of pure proteins. With regard to gel samples, protein bands detectable with coomassie blue (i.e., minimum ca. 100 ng) provide enough material for protein identification. The silver-staining compatible in-gel digestion developed by Mann et al. (see Section 2.4.2) even enables the identification of low ng protein amounts in the gel matrix. In this approach the MALDI mass spectrometer has to be optimised *for* and exclusively dedicated *to* protein identification by peptide mapping.

For protein identification, the peptide mapping data and – if available – surplus information on the yet unknown protein (e.g., molecular mass based on PAGE) and on the intended search routine (search parameters, see below) are imported into a data-base search programme, e.g., Peptide Search [88–90], MS Fit/MS Tag [91], PAWS [92], or ProFound [93]. These programmes search in a protein data-base,

for instance, the SwissProt [94], which includes protein sequences as well as protein structures. The data-base search based on the peptide mapping results yields a hitlist of more or less likely candidates for the protein of interest, ordered according to the number of matching peptides, i.e., the number of proteolytically generated peptides, that correspond to theoretically expected proteolytic fragments of a given protein.

As a first search parameter, the expected molecular mass range of the candidate to be identified can possibly be diminished, for instance, based on a gel-electrophoretically determined molecular mass. However, the possibility of protein truncation should as well be taken into consideration. The peptide mass list obtained from a MALDI mass spectrum of a proteolytic digest can be directly imported into the search programme and the protease used for peptide mapping must be defined. Further critical search parameters are the mass deviation tolerance, dependent on the instrumentally achievable mass accuracy, and the number of missed protease cleavage sites. Searching with average peptide masses in the range of 2,500–10,000 D, e.g., obtained in linear mode on high-performance DE-MALDI mass spectrometers, a mass deviation tolerance of 0.05% is reasonable. If the MALDI-MS peptide analysis was performed with monoisotopic resolution, typically in reflector mode, the mass deviation tolerance in the search with monoisotopic peptide masses up to m/z 5,000 can be set to 0.01% instead in order to further reduce false positive peptide identifications. Logically, the higher the mass accuracy and the more narrow the mass deviation window, the more specific is the protein identification [95].

The number of missed protease cleavage sites, that is still accepted in the search programme, should be set according to the digestion conditions. Given the case of an *in-situ* digestion in a gel plug, it is recommended to accept 2–5 missed cleavage sites, since the digestion in the gel matrix is in general less effective than the one in solution.

Peptide sequence data can be obtained by PSD fragmentation studies as already described in the preceding section. The inclusion of “sequence tags”, i.e., short partial sequences, into the data base search, dramatically reduces the number of protein identities suggested by the search programme, and often it completely excludes false positive results [96]. A search on a single sequence tag can be sufficient to identify a protein unambiguously.

The secured identification of a protein by MALDI-MS peptide mapping is dependent on the number and abundance of the peptide molecular ions, the accuracy in mass determination, the mass deviation tolerance, and the availability of sequence tags. Independently of instrumentation, quality of the peptide mapping data and search parameters, the verification of a positively identified protein and the elimination of false positives can be promoted by evaluation of independent data: the suggested protein in the hitlist should be checked in terms of species of origin, molecular mass, and isoelectric point, if this additional information is known about the protein to be identified.

Ideally, the top-listed “protein hit” of a search result should reveal by far the most matching peptides. In the case of several protein candidates ranking close to each other in terms of their probability (i.e., with similar or identical number of matching peptides), these hits can either encompass a family of protein isoforms (e.g., stemming from different species), or the search programme suggests the same species under different names due to data base redundancy, or the hitlist is inconsistent in the way that several different and entirely unrelated proteins are proposed with almost equal probability. The last scenario represents either the fact, that the protein of interest is not registered in the data base, or that a protein mixture rather than one species has been mapped. In the latter case, the search results can be for instance improved by exclusively submitting the matching peptides of the highest-ranking candidate to a second search routine. If the first top-listed species becomes more outstanding in the second search, this can be interpreted as a more consolidated identification. Another possibility is a

more restricted mass deviation tolerance for peptide match searches that can lead to a more outstanding “protein hit” in the search result list. The most efficient way to consolidate the identification of several proteins from a mixture is to create sequence tags by PSD experiments (see above).

2.5. MALDI-MS peptide mapping: Characterisation of protein modifications

The ubiquity of protein modifications is well established. Protein glycosylation for instance occurs without exception in integral membrane proteins of higher organisms and is quite common with secretory proteins. In blood serum, almost all proteins are glycosylated as are those in hen egg white. Glycoproteins are also found in the cytoplasm and in the cell nucleus [97]. Phosphorylation is a common protein modification as well. Kinases encompass 2% of the eucaryotic genome and 1% of the latter accounts for phosphatases [98]. Approximately 30% of the intracellular proteins identified to date are found to be phosphorylated. Disulfide bonds, an elementary and almost omnipresent protein modification, determine and stabilise tertiary structures and define protein subunits and domains, e.g., immunoglobulin loop structures. Taking numerous other modifications into account like prosthetic groups (e.g., heme in myoglobin [99]), membrane anchors (e.g., fatty acids in the lung surfactant protein SP-C [100] or the GPI anchor on the neural cell adhesion protein axonin-1 [101]), sulfation (of serines, threonines and glycans), it becomes evident, that co- and post-translational protein modification is rather the rule than the exception. Many of the protein modifications can neither be studied by NMR spectroscopy, nor X-ray crystallography or by biochemical approaches. Carbohydrates on proteins for instance are rarely amenable to multi-dimensional NMR analysis because of their high structural complexity and the limited amounts available. Since they additionally represent very flexible structures, the investigation by means of X-ray crystallography is in most cases impaired. Therefore, there is a great demand for alternative analytical approaches to specific, sensitive and fast characterisation of protein modifications.

Consequently, the protein analysis approach of using specific proteases and the mass spectrometric analysis of their cleavage products has been extended to the application of enzymes and reagents capable to remove, transfer, or modify posttranslational modifications, making the latter amenable to structural characterisation by mass spectrometry [102,103]. This analytical concept has been successfully applied to the determination of location and structure of glycan moieties in glycoproteins [28] or to the characterisation of phosphorylation sites in regulatory proteins [104]. Within this review, these new strategies are exemplified by focusing on MALDI-MS characterisation of protein glycosylation, phosphorylation, and of disulfide bond positions.

2.5.1. Characterisation of protein glycosylations

Analytical properties of carbohydrates and glycoproteins. According to their multiple biological functions [105], carbohydrates on proteins occur in a broad variety of constitutional and conformational isomers [106]. The stereochemical differentiation of glycans, i.e., the specific characterisation of isomeric monosaccharides and of different glycosidic bond types, by pure mass spectrometric means is not possible. This limitation can be partially overcome by the combination of digestion with highly specific glycosidases and subsequent mass spectrometric analysis of their cleavage products. This strategy has to date become a powerful tool for the analysis of glycan constitution and sequence.

Furthermore, carbohydrates on proteins often reveal heterogeneous termini and often they are heterogeneously distributed throughout the protein primary structure, i.e., one carbohydrate species can be found at different glycosylation sites in the protein and a given glycosylation site does not necessarily have to be quantitatively glycosylated. Furthermore, it has recently been shown, that not

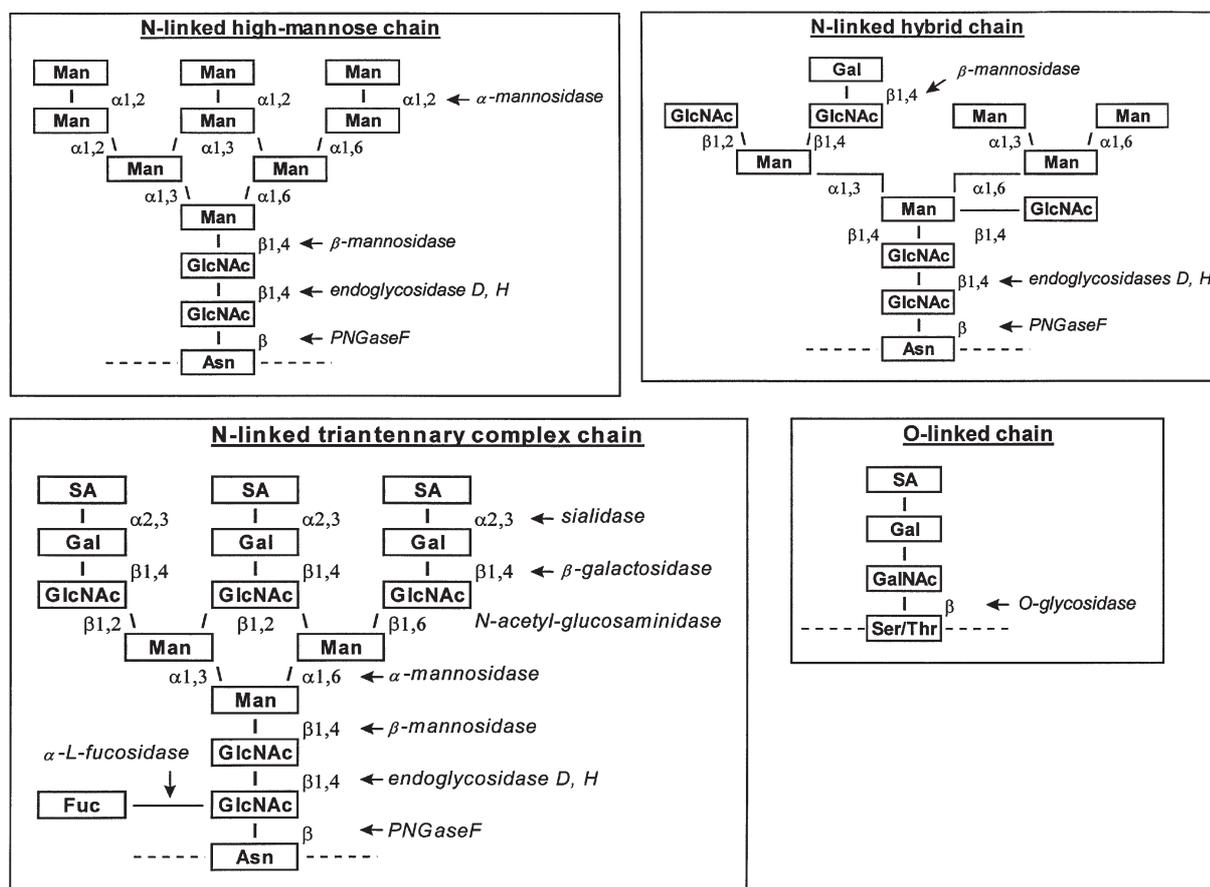


Fig. 7. Examples of typical N- and O-linked carbohydrate structures on proteins. Cleavage sites of specific exo- and endoglycosidases are indicated (Fuc = fucose, Gal = galactose, GlcNAc = N-acetyl-glucosamine, Man = mannose, SA = sialic acid).

only phosphorylation (see Section 2.5.2) but also glycosylation is involved in protein regulation and, therefore, these glycans are found to be attached *to* and removed *from* a glycoprotein within short time intervals [105,107].

In order to illustrate the analytical approach to the characterisation of protein glycosylation, the following chapter will focus on four typical categories of carbohydrate structures shown in Fig. 7: three Asn-N-linked glycan types, e.g., the high-mannose type, the triantennary complex structure, and the hybrid chain, as well as the Ser/Thr-O-linked chains. The high-mannose chain is characterised by three antennas of mannoses and two core GlcNAc residues. The complex chain encompasses three terminal sialic acids, bound to a complex core structure of alternating galactose, GlcNAc, and mannose residues. The hybrid variant combines structural features of both, the high-mannose and the complex chain, and bears one terminal galactose. The O-linked glycans are often less complex, e.g., consisting of the sequence sialic acid–galactose–N-acetyl galactoseamine.

Figure 8 summarises mass spectrometric strategies to the characterisation of protein glycosylation. When elucidating carbohydrate structures of a glycoprotein, there are two alternatives at hand in terms of the initial analytical step: firstly, the glycans can be released from the protein by N- and O-glycosidases and the carbohydrate pool can be analysed separately, or, secondly, the glycoprotein

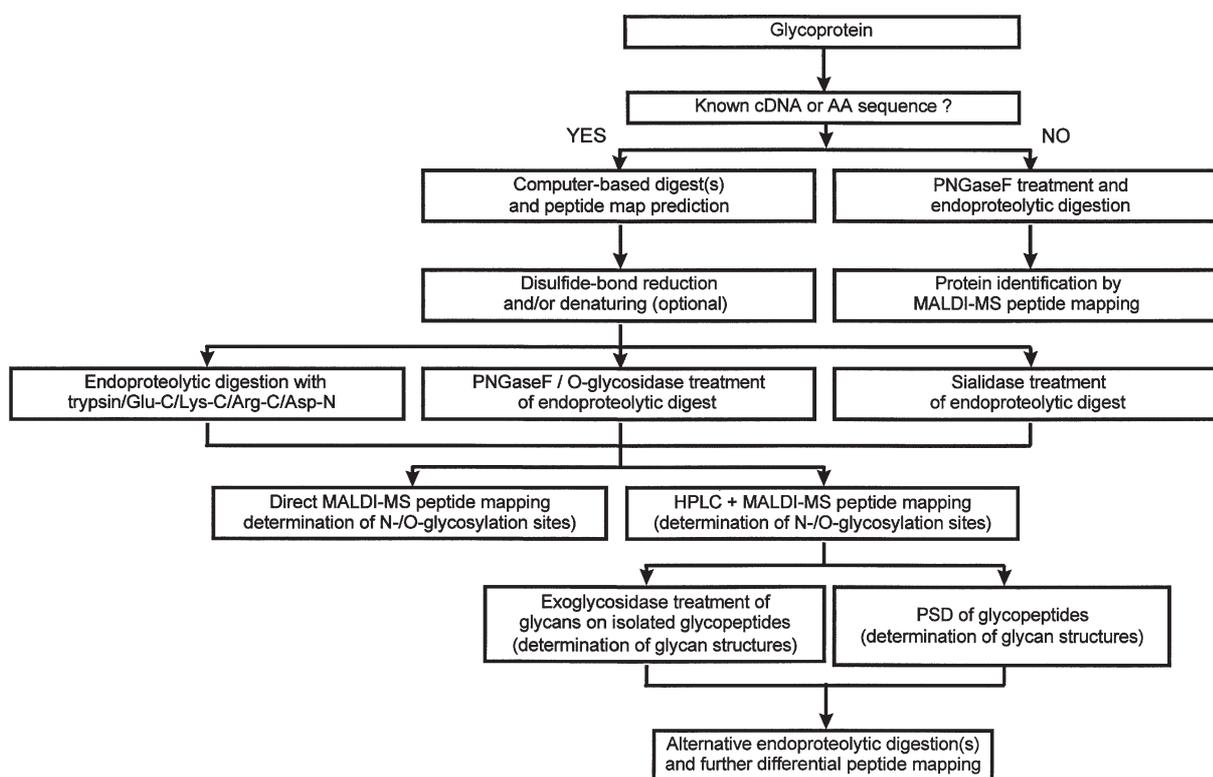


Fig. 8. Mass spectrometric approaches for the characterisation of protein glycosylations.

is directly subjected to mass spectrometric analysis and peptide mapping. The first approach is recommended for preliminary characterisation of extent and complexity of the protein glycosylation as well as to facilitate the characterisation of the protein primary structure (see Section 2.4.3). On the other hand, the peptide moieties of glycopeptides have turned out to be an efficient “handle” to characterise carbohydrates with MALDI-MS. The peptide-typical mass spectrometric properties in glycopeptides provide – relatively to pure glycans – elevated ion abundances due to more efficient ionisation and desorption [108].

Glycopeptides normally reveal a complex molecular ion pattern, in which the different species differ by one or several monosaccharide residues. This is especially valid for complex chains bearing terminal sialic acids. This micro-heterogeneity and the split into several subspecies results in reduced abundances of the single ions. It has been recently demonstrated, that this complex ion pattern actually reflects the species’ heterogeneity in nature and is not due to carbohydrate fragmentation upon the ionisation and/or desorption process [109]. An exception are the complex-type carbohydrate structures with terminal sialic acids, the latter possibly subjected to fragmentation at their glycosidic bond [109,110]. The consequence is a partial cleavage of the terminal sialic acids from the glycan upon MALDI ionisation and/or desorption.

Analysis of protein glycosylation by differential peptide mapping. There is a broad variety of endo- and exoglycosidases commercially available, the purities and specificities of which play a pivotal role in the combination with mass spectrometric characterisation of carbohydrate structures. In Fig. 7, the specificities of different glycosidases are indicated by their cleavage positions. The endoglycosidases

PNGaseF and O-glycosidase can be applied to specifically remove Asn-N- and Ser-/Thr-O-linked glycans, respectively. Sialidase, cleaving terminal sialic acids from complex-like glycan structures, α/β -mannosidase, and exo- β -galactosidase are examples for specific exoglycosidases cleaving O-glycosidic bonds between monosaccharides located in the outer antennas of branched carbohydrate structures. The purities of the enzymes vary to a certain extent and it is therefore recommended to perform a glycosidase incubation without the substrate as a control experiment. Doing so, contaminating peptides can be identified and excluded from misleading data interpretation.

The elucidation of N- and O-glycosylation sites requires the “differential peptide mapping” strategy. The glycoprotein is first digested with an appropriate endoprotease (e.g., trypsin, endoprotease Glu-C, Lys-C, Arg-C or Asp-N) without any treatment by glycosidases. The molecular ions matching to expected protonated proteolytic peptides obtained in this peptide map represent unmodified partial sequences. The mismatching peptide molecular ions are likely to bear modifications, in particular, if unspecific cleavage *by* and autodigestion *of* the applied protease can be ruled out. This peptide map has to be compared to the digest obtained with the same protease but with additional treatment by an endo- or exoglycosidase. Either the intact protein can be subjected to glycosidase digestion prior to the endoproteolytic treatment, or the proteolytic peptide mixture can be exposed to the glycosidase. The latter variant often requires less glycosidase and offers higher cleavage yields, since glycans on peptides are more readily accessible to the enzyme than they are in intact proteins. Peptides exclusively appearing in the peptide map after PNGaseF or O-glycosidase treatment and corresponding to expected proteolytic fragments are the candidates for N- and O-glycosylation, respectively. Ideally, the protein sequence coverage after application of protease(s) and endoglycosidase(s) equals 100% and closes the gaps, that could not be covered without removal of the glycans. An analogous differential peptide mapping study can be performed by use of sialidase in addition to the selected endoprotease. Sialidase – or neuraminidase – removes terminal sialic acids from complex-type carbohydrate structures and often enables the detection of these glycopeptides at all. As discussed in the previous paragraph, complex-type glycans are heterogeneously sialiated at their termini [109,110] and can furthermore undergo fragmentation at the glycosidic bond of the sialic acid resulting in a very complex ion pattern with small individual ion abundances. Therefore, sialic acid-carrying glycopeptides are often suppressed in MALDI-MS mixture analyses. Mismatching peptides, that are only detected after additional sialidase digestion of the protein or the digest, and heterogeneous molecular ion patterns with m/z spacings of 291 (i.e., the molecular mass of one sialic acid residue), that disappear after sialidase treatment, strongly indicate the presence of complex chain glycans with terminal sialic acids.

The separation of glycopeptides on reversed-phase HPLC columns prior to MALDI-MS analysis is often essential for detailed structural characterisation of protein glycosylations, in particular of the complex N-linked glycans. Differential peptide mapping using both RP-HPLC separation and subsequent MALDI-MS analysis can provide complete elucidation of all glycosylation sites and, additionally, *partial* glycosylation can be characterised. If a given peptide with a glycosylation consensus site is only partially (i.e., not quantitatively) glycosylated, meaning that it exists in both its unmodified and modified form, the latter is often not detected in the direct mixture analysis. After chromatographic separation, both species can be analysed individually with MALDI-MS.

In recent studies [64], there occurred at first glance contradictory results of protein glycosylation on the one hand, e.g., suggested by lectin-binding assays and mass spectrometric investigation of the intact protein, and a 100% sequence coverage found by peptide mapping without glycosidase treatment (i.e., every peptide is “at least once” found unmodified) on the other hand. Considering the fact that glycans are neither cleaved during MALDI sample preparation nor by the desorption/ionisation process, there

are two conclusions to be drawn: either the glycans are lost during protein purification and isolation prior to mass spectrometric sample preparation and analysis or the protein is heterogeneously and – with regard to single glycosylation sites – only partially modified. The recent finding of the spatially and temporarily regulated glycosylation patterns and the deduced regulatory function of glycosylations supports the hypothesis of heterogeneous and non-quantitative protein glycosylation. A complex blend of differently glycosylated proteins is an extremely demanding analytical case and has to date not been completely solved by mass spectrometric approaches.

Differential peptide mapping applying several enzymes, various sample preparations, and different MALDI-MS acquisition modes in order to get maximum data output resembles the “assembly of a complicated puzzle” rather than the “alignment of a few spectra”, in which the modification sites and types are immediately visible. This assessment is particularly valid for the often complex situation of protein glycosylation. However, to date there are several computer programmes available that accept protein and peptide modifications, enable the handling of complex data sets and, therefore, accelerate the data processing significantly (see Section 2.4.3).

Carbohydrate sequencing. Carbohydrate sequencing can be performed by sequential application of exoglycosidases (enzyme arrays) and by monitoring of the truncated glycans or, preferably, of the truncated glycopeptides with MALDI-MS. The sequential removal of monosaccharides will evolve in the mass spectrum as mass decrements between 146 and 291 D for the most common sugar residues. Due to the complexity of proteolytic digests of glycoproteins and the suppression effects on glycopeptides in mixture analyses, carbohydrate sequencing may not always work on glycopeptide mixtures but must rather be carried out on isolated glycans or glycopeptides.

If preliminary data like lectin-binding assays or first mass spectrometric results are not available or if these data are inconsistent so that a particular glycan structure can not be anticipated, the application of a glycosidase array, that first probes the termini of the different carbohydrates, is recommended. Treatment of the glycopeptide with α -mannosidase, sialidase, and *exo*- β -galactosidase carried out in different aliquots enables the differentiation between high-mannose, triantennary-complex, and hybrid chain. After the determination of the glycan type, the carbohydrate can be sequenced as described in the following.

The high-mannose chain requires the incubation of the glycopeptide with α -mannosidase in order to sequentially remove the mannose residues. The time course digestion monitored by MALDI-MS should reveal the length of the mannose antennas. Endoglycosidase D or, alternatively, N-acetylhexosaminidase cleave between the remaining two GlcNAc residues and PNGaseF removes the last GlcNAc from the asparagine residue. The sequential removal of the high-mannose core structure is also monitored by MALDI-MS.

For sequencing of complex-chain glycans, an enzyme array should be designed as follows: sequential application of (1) sialidase (see previous paragraph); (2) β -galactosidase (hydrolyses Gal- β -1,4-bonds); (3) N-acetyl- β -glucosaminidase (to remove GlcNAc- β -1,6-); (4) α -mannosidase (see above); (5) β -mannosidase; (6) N-acetylhexosaminidase or endoglycosidase D (cleave between two GlcNAc); (7) α -L-fucosidase (cleaves - α -1,6-fucose bonds); (8) PNGaseF (see previous paragraph). After every digestion step, the remaining glycopeptide is analysed by MALDI-MS.

The hybrid structures can be sequenced by a combined glycosidase array derived from the two preceding ones, starting by treating the glycopeptide with *exo*- β -galactosidase to remove the single terminal Gal- β -1,4- residue.

An example of a characterisation of a glycopeptide by a combination of MALDI-MS and *exo*-glycosidase treatment is given in Fig. 9. The glycopeptide was generated by digestion of Human

Urinary Erythropoietin with endoprotease Lys-C and it was isolated by RP-HPLC [110]. The MALDI mass spectrum in Fig. 9(a) shows three molecular ions differing by ca. 291 Da reflecting the terminal heterogeneity of the glycopeptide, i.e., the presence of none (m/z 2868.2), one (m/z 3158.6) and two (m/z 3448.8) terminal sialic acid residues. The native heterogeneity may be overlaid by a partial cleavage of the sialic acid residues from the core glycan-structure upon desorption and/or ionisation [109,110]. Spectrum (b) shows the glycopeptide after sialidase treatment, i.e., after removal of the terminal sialic acid residues, revealing only one molecular ion (m/z 2867.5) corresponding to the glycopeptide with the remaining core structure [110].

MALDI mass spectrometric fragmentation studies by post-source decay (PSD, see Section 2.1) can also be used for structural characterisation of glycopeptides. The limitation of this technique is, as already mentioned in the context of peptide sequencing, the high substrate dependency, and the lability of the Asn-N- and Ser/Thr-O-glycosidic bonds, the latter often causing the loss of the glycan as the major fragmentation reaction. As it is valid for peptide fragmentation studies, PSD should be employed for confirmation or further characterisation of an expected glycopeptide structure rather than for the analysis of an entirely unknown species.

2.5.2. Determination of protein phosphorylation sites

Probing protein phosphorylation by protein chemistry and mass spectrometric peptide mapping includes similar analytical approaches as applied to glycosylation site analysis. Due to the restriction to one type of modification, namely the phosphoryl residue with a mass increment of 80 D ($-\text{PO}_3\text{H}^-$ at pH 7), the experimental strategy and the data interpretation is easier. Nevertheless, taking the regulatory function and the spatial and temporal regulation of protein phosphorylation into account, the modification analysis can be still quite demanding.

In analogy to the previously described glycoprotein analysis, differential peptide mapping of phosphorylated proteins includes first the digestion with a suitable endoprotease and subsequently the use of the same protease plus the application of alkaline phosphatase, an enzyme, that specifically removes Ser-O- and Thr-O-bound phosphoryl groups. Serine- or threonine-containing mismatching peptides derived from the first digest, that, after alkaline phosphatase treatment, show a mass loss of (n times) 80 Da and yield a peptide match, are the candidates for carrying Ser-/Thr-O-phosphorylations. Although tyrosine phosphorylation only accounts for 0.01% of eucaryotic protein phosphorylation, it has also been investigated by mass spectrometry [111]. A tyrosine-O-phosphatase was used for differential peptide mapping. The HPLC separation of the proteolytic peptide mixture and the individual treatment of the putative phosphoryl peptides with (alkaline) phosphatase should confirm and supplement the results on modification sites obtained by direct peptide mapping and may be necessary in cases of a complex phosphorylation pattern.

In the PSD analysis of a phosphorylated peptide, the loss of the phosphoryl group is a common and characteristic fragmentation reaction indicated by a peptide mass reduction of 80 D and can therefore be interpreted as evidence for peptide phosphorylation.

2.5.3. Determination of disulfide bond positions

The strategy of differential peptide mapping, with or without HPLC application depending on sample amounts and complexity of the analytical problem, can also serve for the determination of disulfide bond positions.

The reduction of disulfide bonds is preferentially carried out with dithiothreitol (DTT), possibly under denaturing conditions by means of urea or guanidinium hydrochloride. Alternatively to conventional reduction in solution, the reaction can also be carried out on the MALDI-MS target. In

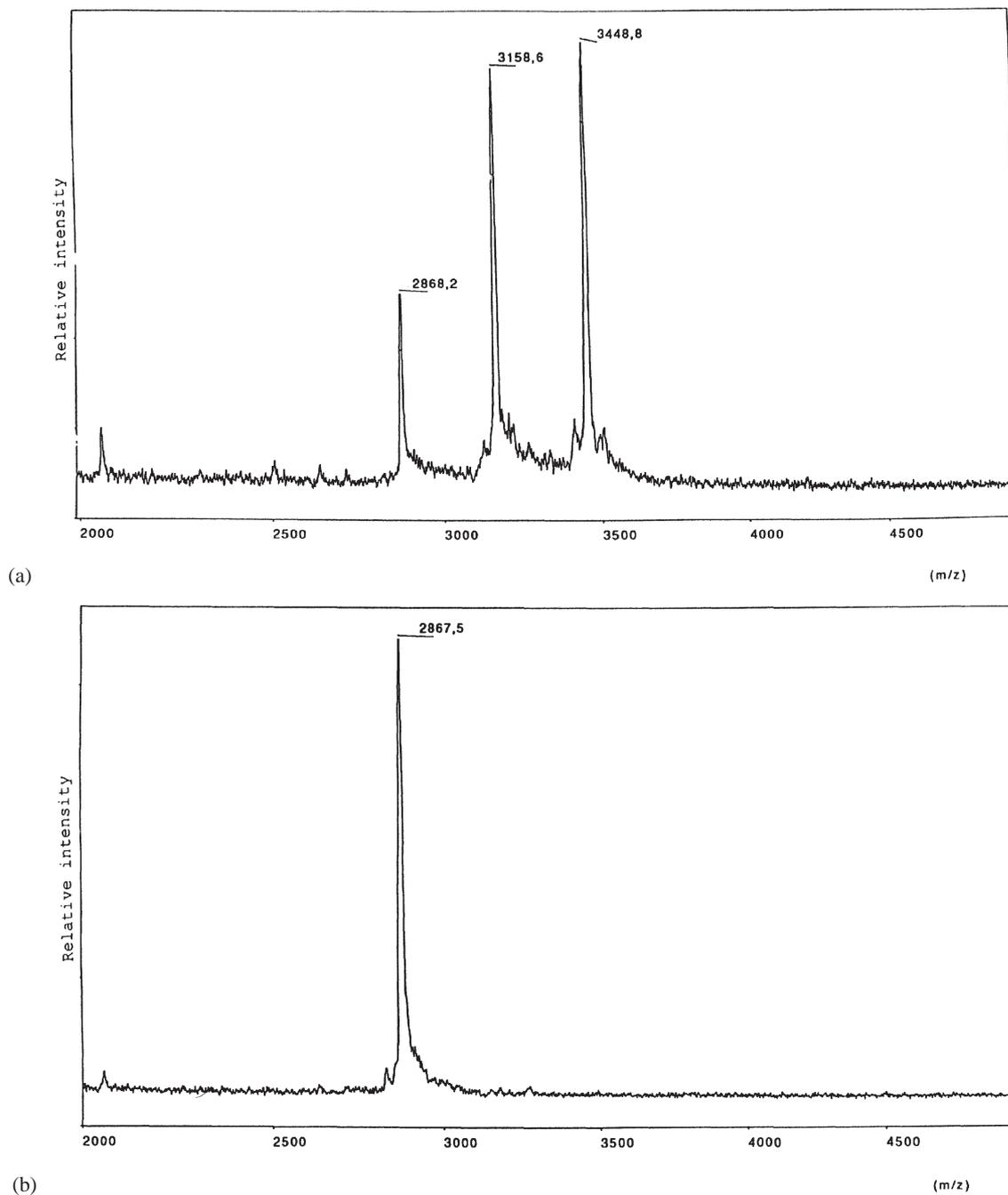


Fig. 9. MALDI mass spectrometric characterisation of a glycopeptide generated by digestion of Human Urinary Erythropoietin with endoprotease Lys-C and isolated by RP-HPLC. Spectrum (a) shows three molecular ions differing by 291 Da reflecting the terminal heterogeneity of the glycopeptide, i.e., the presence of none (m/z 2868.2), one (m/z 3158.6) and two (m/z 3448.8) terminal sialic acid residues. Spectrum (b) shows the glycopeptide after sialidase treatment, i.e., after removal of the terminal sialic acid residues, revealing only one molecular ion corresponding to the glycopeptide with the remaining core structure. Reproduced from Rahbek-Nielsen et al. [110].

the latter case, the protein/matrix sample is incubated with a DTT solution and the entire mixture is subsequently used in a HCCA/sandwich sample preparation, which is optimised for peptide mapping of contaminated proteolytic mixtures because of its compatibility with extensive on-target washing procedures (see Section 2.2). This variant is recommendable, if little sample amount suggests reuse of the protein/matrix sample and if a fast identification or confirmation of (a) disulfide bond(s) is needed.

According to the two preceding chapters, the protein of interest is first to be digested with a suitable endoprotease. This peptide map must be compared to that obtained by the same proteolytic digestion plus DTT reduction of the peptide mixture or protein. Given a few disulfide bonds distant from each other in the protein sequence and with endoproteolytic cleavage sites between the half-cystines, the unassigned cystine containing peptides derived from the digest of the unreduced protein or the unreduced proteolytic mixture, that split up after reduction into two smaller peptides corresponding to expected cysteinyl peptides, are likely to participate in disulfide bonds in the native protein structure. The situation becomes more complicated, if a protein possesses a complex disulfide pattern, possibly with disulfide bridges overlapping in terms of their half-cystine positions in the protein sequence or, if a mixture of intra- and interchain disulfide linkages is to be analysed. A crucial prerequisite for obtaining unambiguous peptide mapping data is the presence of proteolytic cleavage sites between two half-cystines. If there is no such cleavage position, a peptide with one cystinyl bridge will only exhibit a mass difference of +2 Da upon DTT reduction. In a MALDI-MS mixture analyses, such small mass differences may be difficult to assign. Consequently, in cases of proteins encompassing complex disulfide patterns, HPLC separation, individual reduction of potential cystinyl peptides and subsequent MALDI-MS analysis of the isolated peptides becomes inevitable.

PSD analysis of cystinyl peptides may yield disulfide bond breakage beside the generation of peptide sequence ions since metastable decay of disulfide-linked peptides has been demonstrated with LSI-MS [112] (liquid secondary-ion mass spectrometry, synonym for FAB-MS, i.e., fast-atom bombardment mass spectrometry). Even in normal MALDI-MS mode, disulfide bonds are susceptible to prompt fragmentation, if harsh desorption and ionisation conditions are used [113]. These fragmentations can be analytically exploited, especially, if there are no proteolytic cleavage sites available between to half-cystines.

The problem of analysing reduced proteins and their digests is the (possibly uncontrolled) refolding of the protein and the recombination of cysteinyl peptides effecting “disulfide bond scrambling” [114], i.e., the creation of protein or peptide artefacts with non-native disulfide linkages. This obstacle can be circumvented by acidic quenching of the reduction, since reduced thiols are stable at low pH. A commonly used alternative is the *in-situ* blocking of the freshly formed cysteinyl thiol groups. This irreversible protection can be achieved with excesses (relative to DTT) of thiol-alkylating reagents like iodoacetamide or 4-vinylpyridine.

The digestion of the unreduced protein can be hampered by very compact and stable protein tertiary structures that only give rise to limited proteolysis. If proteins have many disulfide bonds, even denaturing may not be sufficient to render the unreduced protein structure amenable to proteolysis.

In more recent studies, proteins were reduced under non-denaturing conditions and the reduced cysteines were probed with dithiol specific reagents, that exclusively cross-link spatially (not necessarily sequentially) vicinal thiol groups [114–116]. This protein chemical approach enables the spatial differentiation of cysteinyl thiol groups and provides a tool for probing protein folding and its intermediates.

3. Conclusions and perspectives

Matrix-assisted laser desorption/ionisation mass spectrometry is an essential and powerful tool in the characterisation of proteins in general, especially of limited protein amounts isolated from biological tissues. Due to the effective interplay and complement of MALDI- and ES-MS, the limiting factors in protein studies are no longer to be allocated to the mass spectrometric performance, but to the preceding isolation and purification procedures. Therefore, the development of improved separation and purification procedures tailored for subsequent mass spectrometric analysis is highly desirable and has recently shown impressive progress.

The potential and success of MALDI-MS are predominantly founded on sensitivity, specificity, and speed of the analysis as well as on the capability to handle complex mixtures and on the tolerance towards contaminations. A further useful feature is the compatibility with gel-electrophoretic separation procedures, in particular with two-dimensional PAGE, a method to date unsurpassed in terms of separation of complex protein mixtures.

The span of MALDI-MS applications covered in this review extends from protein identification by fast peptide mapping routines via the detailed primary structure investigation of proteins to the elucidation of complex protein modification patterns such as glycosylation. Especially, the combination of selective protein chemistry with MALDI mass spectrometric analysis of the protein derivatives has evolved to a primary analytical strategy for detailed characterisation of covalent protein structures.

MALDI-MS should be considered as one of several efficient techniques in an assembly of mass spectrometric methods, the applications and potentials all of which are by far exceeding those outlined in this review. The study of enzyme specificity and kinetics [117], of protein interactions [6], of protein tertiary [118] and supramolecular structures [16], and of immune specificity [119,120] should be mentioned in this context. Based on the combination of 2D-PAGE protein separation with MALDI-MS analysis, the proteome [121] research as a complement to the genome projects has been started recently and, since then, dramatically accelerated. Furthermore, first mass spectrometric investigations on oligonucleotide- and protein-DNA interactions have been carried out [122].

The growing use of MALDI- (and ES-) MS in biochemistry and molecular biology [123] is anticipated to inspire a further rapid development of improved instrumentation and sample preparation. Taking into account this speed of technological progress, especially the fast increasing compatibility of different kinds of ion sources with several mass analysers and various MS/MS techniques, today's limitations – like the complete analysis of heterogeneously modified proteins – can be tomorrow's routine.

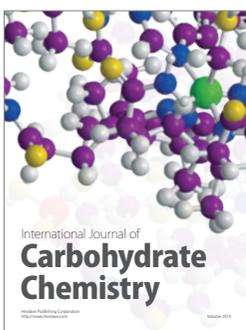
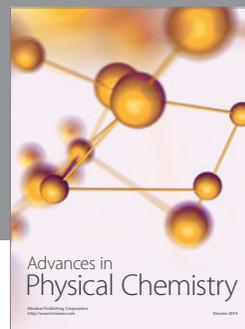
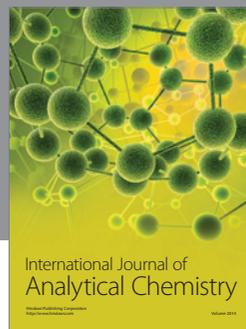
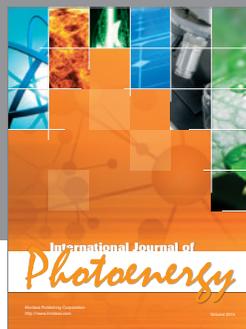
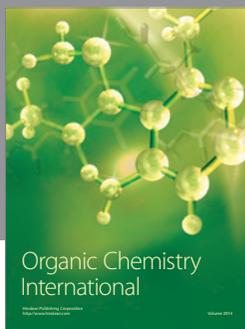
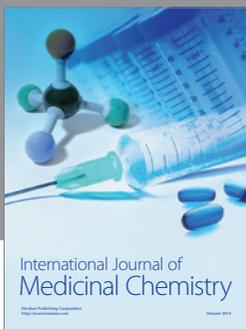
References

- [1] M. Karas and F. Hillenkamp, *Anal. Chem.* **60** (1988), 2299.
- [2] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong and C.M. Whitehouse, *Science* **246** (1989), 64.
- [3] A.L. Burlingame, T.A. Baillie and D.H. Russel, *Anal. Chem.* **64** (1992), 467R.
- [4] A.L. Burlingame, R.K. Boyd and S.J. Gaskell, *Anal. Chem.* **66** (1994), 634R.
- [5] A.L. Burlingame, R.K. Boyd and S.J. Gaskell, *Anal. Chem.* **68** (1996), 599R.
- [6] P. Roepstorff, *Curr. Op. Biotechnol.* **8** (1997), 6.
- [7] R. Wang and B. Chait, *Curr. Op. Biotechnol.* **5** (1994), 77.
- [8] M. Mann and G. Talbo, *Curr. Op. Biotechnol.* **7** (1996), 11.
- [9] M. Mann, *Org. Mass Spectrom.* **25** (1990), 575.
- [10] A. Shevchenko, M. Wilm, O. Vorm and M. Mann, *Anal. Chem.* **68** (1996), 850.
- [11] M. Wilm, G. Neubauer and M. Mann, *Anal. Chem.* **68** (1996), 527.

- [12] T. Matsuo, R.M. Caprioli, M.L. Gross and Y. Seyama, eds, *Biological Mass Spectrometry*, Wiley, New York, 1994.
- [13] M.R. Emmett and R.M. Caprioli, *J. Am. Soc. Mass Spectrom.* **5** (1994), 605.
- [14] R.D. Smith, J.H. Wahl, D.R. Goodlett and S.A. Hofstadler, *Anal. Chem.* **65** (1993), 574A.
- [15] D. Figeys, A. van Oostveen, A. Ducret and R. Aebersold, *Anal. Chem.* **68** (1996), 1822.
- [16] M. Przybylski and M.O. Glocker, *Angew. Chem. Int. Ed. Engl.* **35** (1996), 806.
- [17] N. Potier, P. Barth, D. Tritsch, J.-F. Biellmann and A. van Dorsselaer, *Eur. J. Biochem.* **243** (1997), 274.
- [18] S. Jespersen, W.M.A. Niessen, U.R. Tjaden, J. van der Greef, E. Litborn, U. Lindberg and J. Roeraade, *Rapid Comm. Mass Spectrom.* **8** (1994), 581.
- [19] O. Vorm, P. Roepstorff and M. Mann, *Anal. Chem.* **66** (1994), 3281.
- [20] O. Vorm, B.T. Chait and P. Roepstorff, in: *Proc. 41st ASMS Conf. Mass Spectrom. Allied Topics*, San Francisco, CA, USA, May 1993, p. 621.
- [21] R.C. Beavis and B.T. Chait, *Proc. Natl. Acad. Sci. USA* **87** (1990), 6873.
- [22] A. Fontana, M. Zamboni, P. Polverino de Laureto, V. De Filippis, A. Clementi and E. Scaramella, *J. Mol. Biol.* **266** (1997), 223.
- [23] D. Ellison, J. Hinton, S.J. Hubbard and R.J. Beaton, *Protein Sci.* **4** (1995), 1337.
- [24] D.A. Seilstad, K.E. Carlson, P.J. Kushner, G.L. Greene and J.A. Katzellenbogen, *Biochemistry* **34** (1995), 12 605.
- [25] D. Suckau, M. Mak and M. Przybylski, *Proc. Natl. Acad. Sci. USA* **89** (1992), 5630.
- [26] M. Ploug, H. Rahbek-Nielsen, V. Ellis, P. Roepstorff and K. Danø, *Biochemistry* **34** (1995), 12 524.
- [27] M. Mann, P. Højrup and P. Roepstorff, *Biol. Mass Spectrom.* **22** (1994), 338.
- [28] M.L. Vestal, P. Juhasz and S.A. Martin, *Rapid Comm. Mass Spectrom.* **9** (1995), 1044.
- [29] D.J.C. Pappin, P. Højrup and A.J. Bleasby, *Curr. Op. Biol.* **3** (1993), 327.
- [30] P. James, M. Quadroni, E. Carafoli and G. Gonnet, *Biochem. Biophys. Res. Comm.* **195** (1993), 58.
- [31] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley and C. Watanabe, *Proc. Natl. Acad. Sci. USA* **90** (1993), 5011.
- [32] A.L. Burlingame, *Curr. Op. Biotechnol.* **7** (1996), 4.
- [33] K.A. Resing, S.J. Mansour, A.S. Hermann, R.S. Johnson, J.M. Candia, K. Fukusawa, G.F. Van der Woude and N.G. Ahn, *Biochemistry* **34** (1995), 2610.
- [34] F. Hillenkamp, M. Karas, R.C. Beavis and B.T. Chait, *Anal. Chem.* **63** (1991), 1193A.
- [35] R.J. Cotter, *Anal. Chem.* **64** (1992), 1027A.
- [36] R.S. Brown and J.J. Lennon, *Anal. Chem.* **67** (1995), 1998.
- [37] W.C. Wiley and I.H. McLaren, *Rev. Sci. Instrum.* **26** (1955), 1150.
- [38] K.A. Cox, J.D. Williams, R.G. Cooks and R.E. Kaiser, Jr., *Biol. Mass Spectrom.* **21** (1992), 226.
- [39] R.G. Cooks, G.L. Glush, S.A. McLuckey and R.E. Kaiser, C&EN, *J. Am. Chem. Soc.* (1991).
- [40] M.V. Buchanan and R.L. Hettich, *Anal. Chem.* **65** (1993), 245A.
- [41] T. Robert, J. McYver, Y. Li and R.L. Hunter, *Proc. Natl. Acad. Sci. USA* **91** (1994), 4801.
- [42] M.G. Quian and D.M. Lubman, *Anal. Chem.* **67** (1995), 234A.
- [43] J. Quin, R.J.J.M. Steenvoorden and B.T. Chait, *Anal. Chem.* **68** (1996), 1784.
- [44] B. Spengler, D. Kirsch, R. Kaufmann and E. Jaeger, *Rapid Comm. Mass Spectrom.* **6** (1992), 105.
- [45] T.J. Cornish and R.J. Cotter, *Rapid Comm. Mass Spectrom.* **8** (1994), 781.
- [46] K.F. Medzihradzky, D.A. Maltby, Y. Qiu, Z. Yu, S.C. Hall, Y. Chen and A.L. Burlingame, *Int. J. Mass Spectrom. Ion Proc.* **160** (1997), 357.
- [47] M. Kussmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M. Røssel-Larsen, L. Jakobsen, J. Gobom, E. Mirgorodskaya, A. Kroll-Kristensen, L. Palm and P. Roepstorff, *J. Mass Spectrom.* **32** (1997).
- [48] M. Karas, U. Barr, A. Ingedoh, E. Nordhoff, B. Stahl, K. Strupat and F. Hillenkamp, *Anal. Chim. Acta* **241** (1990), 175.
- [49] R.C. Beavis and B.T. Chait, *Rapid Comm. Mass Spectrom.* **3** (1989), 432.
- [50] R.C. Beavis, T. Chaudhary and B.T. Chait, *Org. Mass Spectrom.* **27** (1992), 156.
- [51] O. Vorm, P. Roepstorff and M. Mann, *Anal. Chem.* **66** (1994), 3281.
- [52] U. Piele, W. Zuercher, M. Schaer and H.E. Moser, *Nucl. Acids Res.* **21** (1993), 3191.
- [53] B. Stahl, M. Steup, M. Karas and F. Hillenkamp, *Anal. Chem.* **63** (1991), 463.
- [54] A. Westman, T. Huth-Fehre, P. Demirev and B.U.R. Sundqvist, *J. Mass Spectrom.* **30** (1995), 206.
- [55] O.N. Jensen, A. Podtelejnikov and M. Mann, *Rapid Comm. Mass Spectrom.* **10** (1996), 1371.
- [56] B. Asgeirsson, S. Haebel, L. Thorsteinsson, E. Helgason, K.O. Gudmundsson, G. Gudmundsson and P. Roepstorff, *Proc. Natl. Acad. Sci. USA*, submitted.
- [57] J. Gobom, E. Nordhoff, R. Ekman and P. Roepstorff, *Int. J. Mass Spectrom. Ion Proc.* (1997, in press).
- [58] H.-P. Fierobe, E. Mirgorodskaya, T.P. Frandsen, P. Roepstorff and B. Svensson, *Protein Exp. Purification* **9** (1997), 159.
- [59] E. Camefeita, P. Alfonso, B. Acevedo and E. Mendez, *J. Mass Spectrom.* **32** (1997), 444.

- [60] S.L. Cohen and B.T. Chait, *Anal. Chem.* **68** (1996), 31.
- [61] J. Ronnenberg et al., *J. Chromatogr.* **656** (1994), 169.
- [62] W.-K. Chui and I.W. Wainer, *Anal. Biochem.* **201** (1992), 237.
- [63] R.D. Voyksner, D.C. Chen and H.E. Swaisgood, *Anal. Biochem.* **188** (1990), 72.
- [64] M. Kussmann, U. Laessing, C.A.O. Stuermer, M. Przybylski and P. Roepstorff, *J. Mass Spectrom.* **32** (1997), 483.
- [65] M. Kussmann, K. Bennett, A.H. Ruediger, B. Guerra, O.G. Issinger and P. Roepstorff, in: *Proc. 14th IMSC (Int. Mass Spectrom. Conf.)*, Tampere, Finland, 25–29 August 1997 (in press).
- [66] D.H. Patterson, G.E. Tarr, F.E. Regnier and S.A. Martin, *Anal. Chem.* **67** (1995), 3971.
- [67] B.T. Chait, R.Wang, R.C. Beavis and S.B.H. Kent, *Science* **262** (1993), 89.
- [68] M. Bartlett-Jones, W.A. Jeffrey, F. Hansen and D.J.C. Pappin, *Rapid Comm. Mass Spectrom.* **8** (1994), 737.
- [69] M. Schuhmacher, M.O. Glocker, M. Wunderlin and M. Przybylski, *Electrophoresis* **17** (1996), 848.
- [70] S. Haebel, C. Jensen, S.O. Andersen and P. Roepstorff, *Protein Sci.* **4** (1995), 394.
- [71] H.H. Rasmussen, E. Mørtz, M. Mann, P. Roepstorff and J.E. Celis, *Electrophoresis* **15** (1994), 406.
- [72] K. Strupat, M. Karas, F. Hillenkamp, C. Eckerskorn and F. Lottspeich, *Anal. Chem.* **66** (1994), 464.
- [73] J. Rosenfeld, J. Capdevielle, J.C. Guillemot and P. Ferrara, *Anal. Biochem.* **203** (1992), 173.
- [74] M. Mann and G. Talbo, *Curr. Op. Biotechnol.* **7** (1996), 11.
- [75] M. Mann, P. Højrup and P. Roepstorff, *Biol. Mass Spectrom.* **22** (1993), 338.
- [76] P. Højrup, General peptide mass analysis for Windows (GPMW), Lighthouse data, Odense, Denmark.
- [77] PE Sciex, MacBioSpec, Thornhill, ON, Canada.
- [78] Beckman Research Institute, MacProMass, City of Hope, USA.
- [79] R. Kaufmann, B. Spengler and F. Luetzenkirchen, *Rapid Comm. Mass Spectrom.* **7** (1993), 902.
- [80] S.A. McLuckey, *J. Am. Soc. Mass Spectrom.* **3** (1992), 599.
- [81] D. Despeyroux and K.R. Jennings, in: *Biological Mass Spectrometry*, T. Matsuo, R.M. Caprioli, M.L. Gross and Y. Seyama, eds, Wiley, New York, 1994, p. 227.
- [82] V. Katta, S.K. Chowdhury and B.T. Chait, *Anal. Chem.* **63** (1991), 174.
- [83] P. Roepstorff and J. Fohlman, *Biomed. Mass Spectrom.* **11** (1984), 601.
- [84] B. Spengler, F. Luetzenkirchen and R. Kaufmann, *Org. Mass Spectrom.* **28** (1993), 1482.
- [85] A. Tsarbopoulos, M. Karas, K. Strupat, B.N. Pramanik, T.L. Nagabhushan and F. Hillenkamp, *Anal. Chem.* **66** (1994), 2062.
- [86] J.S. Andersen, M. Sogaard, B. Svensson and P. Roepstorff, *Biol. Mass Spectrom.* **23** (1994), 547.
- [87] S.D. Patterson and R.S. Aebersold, *Electrophoresis* **16** (1995), 1791.
- [88] M. Mann, Peptide Search, Protein and Peptide Group, EMBL, Heidelberg, Germany.
- [89] M. Mann, P. Højrup and P. Roepstorff, *Biomed. Mass Spectrom.* **22** (1993), 338.
- [90] M. Mann and M. Wilm, *Anal. Chem.* **66** (1994), 4390.
- [91] K. Clauser and A.L. Burlingame, MS Fit and MS Tag, UCSF Mass Spectrometry Facility, University of California, San Francisco, CA, USA.
- [92] ProteoMetrics, PAWS.
- [93] W. Zhang and B. Chait, ProFound, Rockefeller University, New York, USA.
- [94] SwissProt Protein Data Base.
- [95] O.N. Jensen, A. Potelejnikov and M. Mann, *Rapid Comm. Mass Spectrom.* **10** (1996), 1371.
- [96] M. Mann and M. Wilm, *Anal. Chem.* **66** (1994), 4390.
- [97] G.W. Hart, R.S. Haltiwanger, G.D. Holt and W.G. Kelly, *Ann. Rev. Biochem.* **58** (1989), 841.
- [98] T. Hunter, *Cell* **80** (1995), 225.
- [99] M.C. Hsu and R.W. Woody, *J. Am. Chem. Soc.* **93** (1971), 3515.
- [100] J. Johansson, T. Curstedt, B. Robertson and J. Joernvall, *Biochemistry* **27** (1988), 3544.
- [101] R.A. Zuellig, C. Rader, A. Schroeder, M.B. Kalousek, F. Von Bohlen und Halbach, T. Osterwalder, C. Inan, E.T. Stoeckli, H.U. Affolter, A. Fritz, E. Hafen and P. Sonderegger, *Eur. J. Biochem.* **204** (1992), 453.
- [102] J.S. Andersen, B. Svensson and P. Roepstorff, *Nature Biotechnol.* **14** (1996), 449.
- [103] A.L. Burlingame and S.A. Carr, eds, *Mass Spectrometry in Biological Sciences*, Human Press, Totowa, NY, 1996.
- [104] M.F. Bean, R.S. Annan, M.E. Hemling, M. Mentzer, M.J. Huddleston and S.A. Carr, in: *Techniques in Protein Chemistry*, J. Crabbe, ed., Academic Press, San Diego, CA, 1995, p. 107.
- [105] A. Varki, *Glycobiology* **3**(2) (1993), 97.
- [106] H. Lis and N. Sharon, *Eur. J. Biochem.* **218** (1993), 1.
- [107] L.K. Kreppel, M.A. Blomberg and G.W. Hart, *J. Biol. Chem.* **272** (1997), 9308.
- [108] U. Rapp, A. Resemann, F.J. Mayer-Posner, W. Schäfer and K. Feichtinger, in: *Techniques in Glycobiology*, R.R. Townsend and A.T. Hotchkiss, Jr., eds, M. Dekker, New York, 1997, p. 53.
- [109] E. Mørtz, T. Sarena, I. Julkunen and P. Roepstorff, *J. Mass Spectrom.* **31** (1996), 1109.

- [110] H. Rahbek-Nielsen, P. Roepstorff, H. Reischl, M. Wozny, H. Koll and A. Haselbeck, *J. Mass Spectrom.* (1997, in press).
- [111] A. Weijland, G. Neubauer, S.A. Courtneidge, M. Mann, R.K. Wierenga and G. Superti-Furga, *Eur. J. Biochem.* **240** (1996), 756.
- [112] A.G. Craig and J.E. Rivier, *Org. Mass Spectrom.* **27** (1992), 549.
- [113] S.D. Patterson and V. Katta, *Anal. Chem.* **66** (1994), 3727.
- [114] M.O. Glocker, B. Arbogast and M.L. Deinzer, *J. Am. Soc. Mass Spectrom.* **6** (1995), 638.
- [115] M. Kussmann and M. Przybylski, in: *Methods in Enzymology*, Vol. 251, B.B. Buchanan, E. Cadenas, C. Gitler, A. Holmgren, A. Meister and H. Sies, eds, Academic Press, Orlando, FL, 1995, p. 430.
- [116] P. Happersberger, C. Cowgill, M.L. Deinzer and M.O. Glocker, in: *Proc. 44th ASMS Conf. Mass Spectrom. Allied Topics*, Portland, USA, May 1996, p. 1050.
- [117] M. Moniatte and A. van Dorsselaer, in: *3rd EU Network Meeting on Peptide and Protein Structure Analysis by Mass Spectrometry*, Vienna, Austria, 31 March–2 April 1996.
- [118] S.L. Cohen, *Structure* **4**(9) (1996), 1013.
- [119] M. Macht, W. Fiedler, K. Kuerzinger and M. Przybylski, *Biochemistry* **35** (1996), 15 633.
- [120] T.-T. Yip, J. Van de Water, M.E. Gershwin, R.L. Coppel and T.W. Hutchens, *J. Biol. Chem.* **271** (1996), 32 825.
- [121] M.R. Wilkins, C. Pasquali, R.D. Appel, K. Ou, O. Golaz, J.-C. Sanchez, J.X. Jan, A.A. Gouley, I. Humphrey-Smith, K.L. Williams and E.F. Hochstrasser, *Biotechnology* **14** (1996), 61.
- [122] E. Nordhoff, A. Krogsdam, C. Soenksen, I. Madsen, K. Kristiansen and P. Roepstorff, in: *Proc. 45th ASMS Conf. Mass Spectrom. Allied Topics*, Palm Springs, USA, May 1997 (in press).
- [123] D.N. Nguyen, G.W. Becker and R.M. Riggan, *J. Chrom. A* **705** (1995), 21.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

