ELSEVIER

Biophysical Chemistry 68 (1997) 173–188

Biophysical Chemistry

Time, life . . . and mass spectrometry New techniques to address biological questions

Catherine E. Costello *

Mass Spectrometry Resource, Dept. of Biophysics, Boston, Univ. School of Medicine, Boston, MA USA

Received 9 January 1997; revised 21 February 1997; accepted 21 February 1997

Abstract

Within the last ten years, startling new developments in two ionization methods—matrix-assisted laser desorption (MALDI) and electrospray (ESI)—have been described by Karas et al. [M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Proc., 78 (1987) 53.] and by Fenn et al. [J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science, 246 (1989) 64.], respectively. Their work demonstrated that these techniques, under appropriate experimental conditions, have high sensitivity and wide mass range, extending to hundreds of thousands of daltons and beyond, and thus can be extremely effective for the study of biopolymers. The result has been a revolution in the way that mass spectrometry experiments are carried out, a widening of the range of investigators who employ mass spectrometry in their own laboratories and a penetration of mass spectrometry into the investigation of biological phenomena that exceeds any previous expectations. Progress in improving mass spectral ionization and mass analysis methods and in interpreting and understanding the spectra is actively being pursued and exploited in many laboratories, to capitalize even further upon these advances. The results should facilitate understanding of structure–activity relationships pertinent to biology and medicine. In our laboratory, the focus of research is on oligosaccharide and glycoconjugate structural determinations, and on the improvement of methods for these important classes of compounds that relate to development, immune response, signalling, lipid and protein transport and disease. Representative examples of applications of MALDI and ESI mass spectrometry to these and other biological questions are provided herein. © 1997 Elsevier Science B.V.

Keywords: Mass spectrometry; Electrospray ionization; Matrix-assisted laser desorption/ionization; Glycoproteins; Glycolipids; Oligosaccharides; Muramyl peptides; Proteins

1. Introduction

During the last decade, dramatic progress has been made in methods for the structural determination and analysis of biological molecules, largely through exploitation of the advantages of two ionization methods that increase both sensitivity and accessible mass range: matrix-assisted laser desorption/ionization (MALDI) [1] and electrospray ionization (ESI) [2]. In conjunction with the availability of these ionization methods, advances in instrumentation, data processing and information retrieval have also been rapid and significant. It is now possible to determine molecular weight profiles with low femto-

* Corresponding author. Mass Spectrometry Resource, Dept. of Biophysics, 80 E. Concord Street, R-806, Boston, MA 02118-2394. Tel.: +1-617-638-6490; fax: +1-617-638-6491; e-mail: cecmsms@bu.edu

0301-4622/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved.
PII S0301-4622(97)00033-1

mole sensitivity and to use tandem mass spectrometry (MS/MS) to elucidate structural details at the low- or sub-picomole level for individual compounds or the components of mixtures. The spectra can be compared to internationally available databases, or interpreted with the guidance of rules being developed for standard sets of compounds. Biophysical data can be obtained regarding interactions with ligands and cations, ideally under conditions that approach or mimic the physiological norm. As the reach of mass spectrometry is extended to higher molecular weight compounds—now up to hundreds of thousands of daltons and beyond—and to trace components in biological systems, new methods for sample isolation and handling are required. Concomitantly, a new generation of 'user-friendly' instruments is being developed to bring the techniques directly into the workplaces of biological and clinical investigators.

The range of questions whose answers are now amenable to mass spectral pursuit is illustrated here with a few examples from our laboratory. Some collaborative studies involve proteins, peptide receptors and oligonucleotides, but our primary research emphasis is placed on oligosaccharides and glycoconjugates. We seek to improve the mass spectral sensitivity for these compounds and to increase the information content in the spectra, in order to facilitate investigations of their structure–activity relationships in biological systems. Because of the complexity of the structures and the regular occurrence of mixtures of related compounds with varying activities, the analytical problems for carbohydrates exceed those of linear biopolymers such as proteins and oligonucleotides. The challenge must be met, however, because increasing evidence demonstrates that oligosaccharides and their conjugates play many diverse and important roles [3]. Our areas of interest include immunology, nervous system growth and development, transplantation, carcinogenesis, infectious diseases, and parasitology.

Matrix-assisted laser desorption / ionization time-of-flight mass spectrometry (MALDI-TOF MS) is illustrated in Fig. 1(a). The sample is mixed with a thousand-fold excess of a material which absorbs at the wavelength of the laser (337 nm for the nitrogen laser employed for most of the examples described herein). By mechanisms that are as yet incompletely

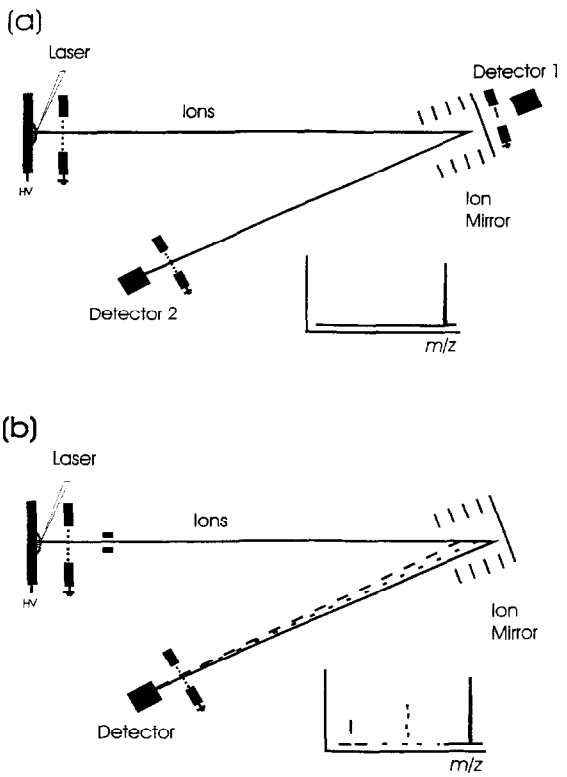


Fig. 1. (a) Scheme of matrix-assisted laser ionization/desorption and mass analysis in a reflectron time-of-flight mass spectrometer. Not to scale: the ion flight path is usually 1–3 m, whereas dimensions in the ion source and detector regions are less than a few cm. (b) Ion trajectories in a reflectron time-of-flight mass spectrometer as a function of reflector voltages. Ion transmission pathways during various stages of the post-source decay experiment are indicated.

understood, the energy transfer to the matrix results in desorption of both matrix and analyte from the surface and the formation of ions within the plume. It is convenient to measure the spectra of the desorbed ions with a time-of-flight mass analyzer, because this device is immediately compatible with the pulsed generation of high-energy ions and the mass range is theoretically unlimited. Linear mode spectra are acquired when the ion currents are measured at Detector 1. The use of a reflectron, as illustrated here with signal recorded at Detector 2, increases the

resolution through energy focusing of the ions in a mirror [4]. When set to transmit the ions that have full energy (stable molecular ions and fragments generated from very unstable molecular ions that underwent 'prompt' fragmentation before acceleration), the reflectron rejects lower energy fragmentations that have arisen during flight. In an experiment that utilizes stepwise decrease of the reflector potential, these lower energy ions may be transmitted to the detector to capture the structural information contained in the fragmentation pattern, as illustrated in Fig. 1(b). This type of data acquisition is now designated as the post-source decay (PSD) mode [5–7] and corresponds to the multistage (MS/MS) measurement of the metastable ion spectrum on other instruments. Sensitivity for PSD is in the sub-picomole range. The performance of even a simple linear device may be significantly improved by allowing a brief (nanoseconds) pause between ion generation and acceleration (time-delayed extraction) [8–10]. The delay results in extraction of ions within a narrower energy range. Other types of analyzers may also be utilized, with varying degrees of cost and requirements for operator expertise. Double-focusing instruments with integrating detectors and ion traps have been used (*mutatis mutandi*) in conjunction with MALDI, and have yielded promising results.

Electrospray ionization mass spectrometry (ESI-MS) and tandem mass spectrometry (ESI-MS/MS) are illustrated in Fig. 2(a) and (b), respectively. In this mode, the sample is introduced in a flowing stream, driven from a syringe pump or eluting from an HPLC or capillary zone electrophoresis system. Maximum sensitivity is observed when the sample solution is admitted directly to the ion source through an extremely finely drawn tip (1–3 μM) that delivers 1 μL over about 30 min [11]. The flow contains an electrolyte, such as sodium acetate, and may include other additives, buffers or 'electron traps', e.g., chlorinated hydrocarbons. As the flow discharges from the tiny nozzle held at high voltage with respect to the ion source lens, droplets emerge bearing the electrolyte on the surface and carrying the sample within. As solvent is stripped away in the vacuum, the droplet size decreases, and sample-rich particles result, often bearing high numbers of charges. The charged molecular ions enter the analyzer region with little excess energy and are therefore quite stable; ESI mass spectra contain mostly molecular ions in a distribution of charge states. Fragmentation may be induced by manipulation of the ion source repeller lens voltage, or by collision of mass-selected ions in the region between MS-1 (Q1) and MS-2 (Q3). The latter approach to ESI-MS/MS is usually preferable, because it allows precursor ion selection and better control over the decomposition process. Although single and triple quadrupole instruments dominate the ESI literature, ESI may also be performed with ion traps, time-of-flight and multisector mass analyzers. With ion traps, whether quadrupolar or Fourier transform ion cyclotron resonance (FT-ICR), multiple stages of mass selection and decomposition are possible, as well as studies of ion-molecule reactions. ESI investigations of noncovalent complexes (e.g., receptor–ligand binding) and differential rates of deuterium exchange [12,13] are particularly relevant for biological studies.

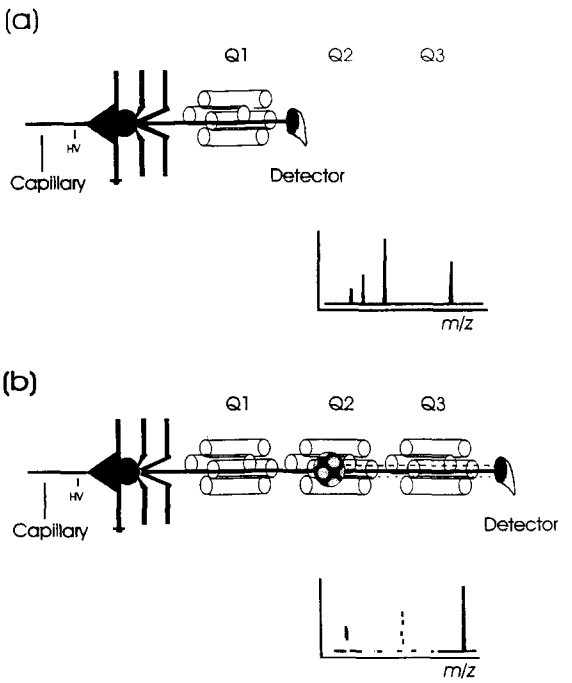


Fig. 2. (a) Electrospray ionization with mass separation in a single-stage quadrupole MS. (b) Ionization and mass separation during an ESI-MS/MS experiment in a triple quadrupole mass spectrometer. The central quadrupole, Q2, serves as the collision cell. Precursor ions selected by Q1 are decomposed in Q2; the product ions are separated in Q3 and detected.

2. Experimental

2.1. MALDI-TOF MS

MALDI-TOF MS was carried out using a Finnigan MAT (now Thermo BioAnalysis, Franklin, MA) Vision 2000 reflectron time-of-flight instrument operated at ±30 kV (linear mode) or ±5–10 kV (reflectron and PSD modes). In the ultraviolet (UV) range, a nitrogen laser (Laser Science, Newton, MA) was used to generate the primary beam, at 337 nm, with 3–5 ns pulse width. In the infrared (IR) range, an Er-YAG laser (finYAG 20-2, Spektrum, Berlin) generated a beam at 2.94 μm, with 50–100 ns pulse width. PSD mass spectra were acquired in 10–15 segments, by stepping down the reflector voltage and were reassembled using the instrument's software package. Samples were applied to the stainless steel target as 10^{-5}–10^{-6} M solutions (1 μL) mixed with an equal volume of the 10^{-4} M matrix solution (usually 2,5-dihydroxybenzoic acid for UV-MALDI, succinic acid for IR-MALDI). The applied drop was allowed to air dry and crystallize before the probe was introduced into the mass spectrometer ion source.

2.2. ESI-MS and ESI-MS/MS

ESI-MS and ESI-MS/MS experiments were carried out with a VG (now Micromass, Beverly, MA) Quattro II triple quadrupole mass spectrometer or a Finnigan (now Thermo BioQuest, Sunnyvale, CA) TSQ 700 triple quadrupole tandem mass spectrometer. In either case, collision-induced decomposition took place with Ar (2 mTorr) in the second (rf-only) quadrupole. Sample solutions (ca. 0.1 pmol/μL in methanol/water (6:4 v/v) containing 0.25 mM NaOH) were introduced with a Harvard syringe pump, at a flow rate of 0.85 μL/min.

2.3. Sample preparation

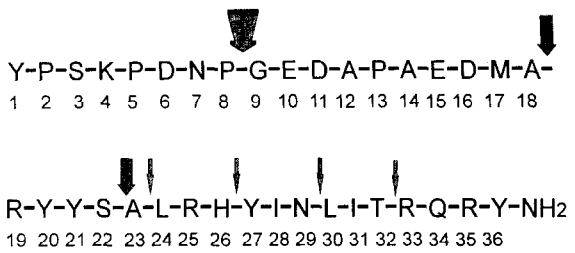
For MALDI-TOF MS, the 2,5-dihydroxybenzoic acid (Aldrich, Milwaukee, WI) matrix was recrystallized or subjected to ion exchange to control cation content. Native samples of oligosaccharides, glycoconjugates, proteins, peptides or oligonucleotides were desalted by passage through a C-18 column (HPLC or Sep-Pak®), microdialysis on a floating membrane [14,15] or ion-exchange [16] until salt content was eliminated or minimized, as judged from the mass spectra. For both MALDI-TOF MS and ESI-MS, oligosaccharides and glycolipids were permethylated using the solid NaOH/methyl iodide procedure of Ciucanu and Kerek [17], adapted for microscale [18], or peracetylated in capillaries using a gas/solid reaction [19].

3. Results and discussion

We move easily back and forth between MALDI and ESI techniques, and frequently use both approaches during a single analysis. This provides information on the unique properties (and potential biases) of the two methods, and allows the selection of optimum conditions for determinations related to samples which are available only in very small amounts because of their biological occurrence at low levels in microscopic tissues or extraordinary requirements for isolation and purification that preclude preparation of larger quantities. These may still be contaminated and may require microscale cleanup, derivatization or degradation to complete the determinations. As an illustration of the approaches, a few recent examples from our own laboratory and collaborative projects are presented here.

3.1. Studies of neuropeptides

Biologically active neuropeptides are generated from much larger preprohormones in response to various stimuli; the initial products from the preprohormones may undergo several transformations to the active form, before they are degraded and inactivated [20]. The timing and course of these processes are controlled by normal and pathological conditions. The pattern of neuropeptides found in a given tissue at any moment thus represents a window on a number of interrelated metabolic pathways. While radioimmunoassay (RIA) often is seen as a convenient method for estimating neuropeptide levels, it can suffer from lack of specificity when presented with a series of closely related compounds, and loss of utility when the recognized functionality is substantially modified or deleted [21]. With the advent of MALDI and ESI techniques, MS sensitivity equals



Scheme 1. Structure of neuropeptide Y with arrows indicating the cleavage sites observed for incubation in human CSF. Reprinted (with permission) from Ref. [25].

or exceeds that of radioimmunoassay [22–24], provides specific molecular weight information and can furnish structural details when required.

3.2. Neuropeptides in human cerebrospinal fluid (CSF)

The fact that neuropeptides may be analyzed by MALDI-TOF MS with little sample preparation has been demonstrated [22,23]. One of the most abundant mammalian CNS neuropeptides is neuropeptide Y (NPY), a 36-residue peptide with a C-terminal amide (Scheme 1) [26,27]. Neuropeptide Y has been shown to control, through hypothalamic mechanisms, the release of several pituitary hormones, to stimulate the appetite, especially for carbohydrates [28], and to affect emotions [29–31]. There is evidence for the existence of several subtypes of neuropeptide Y receptors which individually have a specific requirement for either the intact peptide or a truncated fragment [32]. Information on complete peptide processing patterns has not yet been available. To address this need, in collaboration with R. Ekman et al. [25], Univ. Göteborg, we recently followed and compared the processing of neuropeptide Y in the CSF of normals and depressed patients. CSF samples lyophilized in phosphate buffer (10 mM pH 7.0) were reconstituted to their original volumes with water, spiked with NPY solution to a final concentration of 2 fmol → 1 pmol/μL and incubated at 37°C. For MALDI-TOF MS kinetic profiles, 1-μL aliquots were withdrawn and mixed with 5–10 volumes of matrix solution; spectra were acquired for 1 μL of the resulting mixture. We found that the major products were formed by cleavage of single peptide bonds in NPY to yield N-terminal or C-terminal fragments (Scheme 1). Although the product distribution was fairly constant among individuals, the degradation rates varied sharply. Within the small sample examined thus far, depressed patients showed rates faster than normals. Further MALDI-TOF MS studies will explore the apparent correlation, define processing pathways in different groups or in individuals over time, and may lead to clinical assays and a better understanding of the biochemical basis of some types of depression.

3.3. Glycoforms on rat pineal gland glycopeptide

The sensitivity of MALDI-TOF MS and ESI-MS are such that studies may be carried out on single cells or small biopsy samples. In many cases, the lowest practical limit is one of sample-handling rather than mass spectral sensitivity. Our collaborators in the Vrie Universität, Amsterdam, have demonstrated that peptides in single snail neurons can be profiled by mixing individual cells with the MALDI matrix, applying the mixture to the MS target, and irradiating with a nitrogen laser [23]. In a recent study, they undertook profiling the peptides in the neurointermediate lobe of the rat pituitary gland under normal conditions and compared the pattern to that observed after salt-loading [33]. While some peptide levels remained fairly constant, the levels of the suite of peptides that are released from the preprohormones propressophysin and prooxyphysin were distinctly lowered. Included in this group was a component whose partial amino acid sequence (AREQS?ATQL...) was determined by microscale Edman degradation and corresponded to the C-terminal peptide [A(129)REQSNATQL...Y(168)] of propressophysin (CPP), but whose apparent molecular weight, 5529 daltons, indicated a posttranslational modification. The peptide sequence contained a potential *N*-glycosylation site and the Edman results suggested that the Asn residue was modified. We received an HPLC fraction containing this candidate glycopeptide obtained from extraction of three cells, an amount estimated to be about 200 pmol. The sample was examined by MALDI-TOF MS and a portion (about 150 pmol) was treated with *N*-glycanase to release the carbohydrate. The carbo-

hydrate was purified by passage through a C18 Sep-pak® and permethylated by a modification of the solid NaOH/methyl iodide method described by Ciucanu and Kerek [17] and Linsley et al. [18]. The permethylated oligosaccharide fraction was dissolved in methanol/water (6:4 v/v) containing 0.25 mM NaOH and analyzed by ESI-MS and ESI-MS/MS. This procedure has been developed in our laboratory for profiling glycoprotein oligosaccharides at high sensitivity [34].

In the ESI mass spectrum, Fig. 3(a), the doubly-charged sodiated molecular ion $[M + 2Na]^{2+}$ of the major component was observed at m/z 1052.0, and a minor component had $[M + 2Na]^{2+}$ at m/z 1241.0. These correspond to permethylated molecular weights of 2059.3 and 2437.7, respectively, and imply native molecular weights of 1666.5 and 1974.5 for the released oligosaccharides. The molecular weights calculated for the terminal CPP peptide with these glycoforms would be 5930 and 6238 daltons. These results were consistent with the MALDI-TOF mass spectrum, wherein the observed $[M + H]^+$ values were m/z 5930 and 6239. The mass spectral measurements thus suggested the oligosaccharide compositions $Hex_3HexNAc_5Fuc$ and $Hex_4HexNAc_5Fuc_2$.

In order to investigate details of the peptide glycosylation, the $[M + 2Na]^{2+}$ ion from each re-

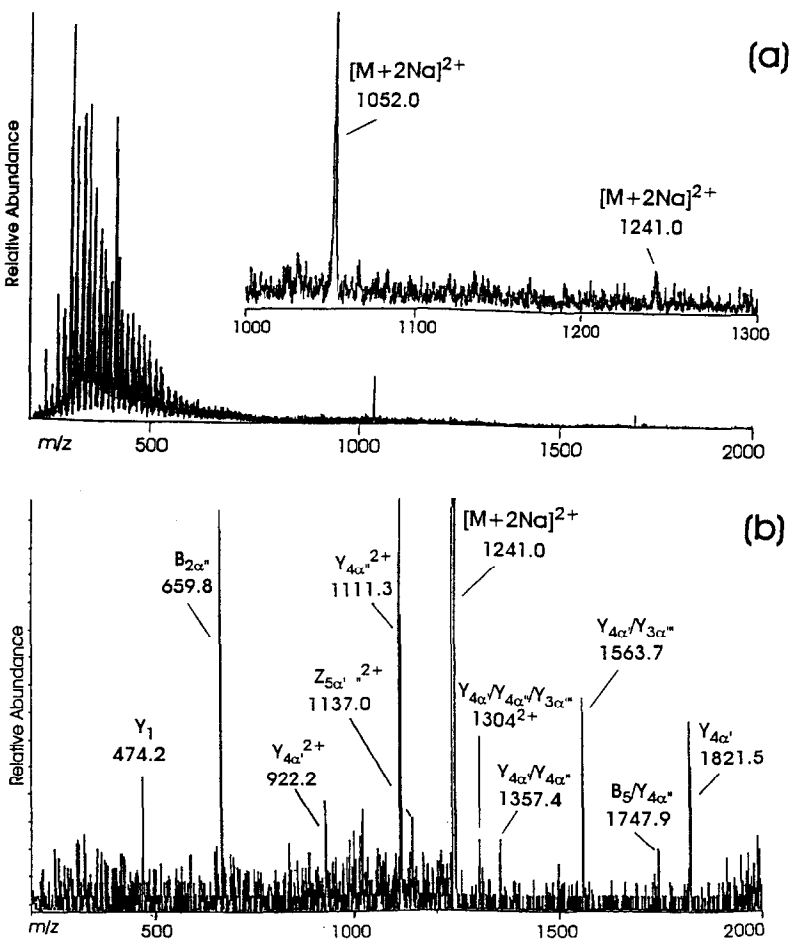
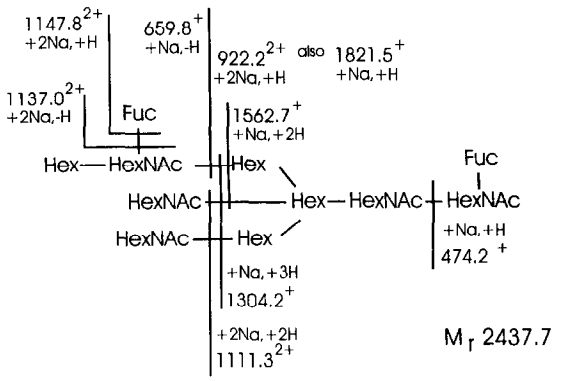


Fig. 3. (a) ESI-MS of the oligosaccharides released from a rat pineal gland glycopeptide and permethylated [33]. (b) CID mass spectrum of the $[M + 2Na]^{2+}$ ion of the minor component, m/z 1241.0. Product ion assignments are shown in Scheme 2 and Scheme 3.

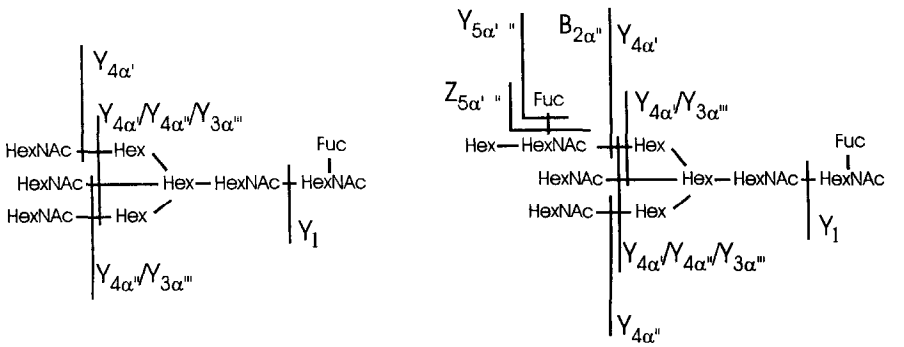


Scheme 2. Assignments of the fragment ions in the ESI-CID MS/MS spectrum of the permethylated minor glycoform released from the CPP glycopeptide (Fig. 3).

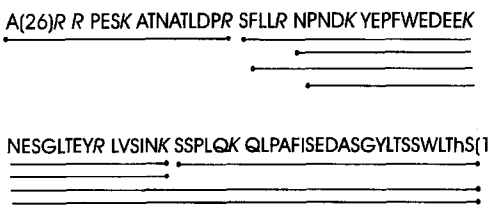
leased and permethylated oligosaccharide was analyzed by ESI-MS/MS with collision-induced dissociation (CID). Even for the minor oligosaccharide, the CID mass spectrum in Fig. 3(b) had excellent signal-to-noise ratio. The principal fragmentation pathway, glycosidic cleavage, yielded singly- and doubly-charged product ions that retained sodium (Scheme 2). The most significant fragments observed in the MS/MS spectra of the two glycoforms are indicated in Scheme 3, using the nomenclature system of Domon and Costello [35]; formation of *Y*-type ions involves cleavage of the glycosidic bond on the non-reducing-end side of the oxygen and includes hydrogen transfer to the observed fragment, whereas formation of *Z*-type fragments involves cleavage on the reducing-end side of the oxygen accompanied by hydrogen loss. *B*-type fragmentation does not involve hydrogen transfer. The spectra thus served to determine the residue sequences and branching patterns (but not the linkage sites). When adequate material is available to permit a second set of experiments, opening of the carbohydrate rings by periodate treatment can be used to provide linkage information [34,36], but the technique was not used in this case, because initial uncertainty about the amount and complexity of the sample suggested a conservative approach.

3.4. Cleavage of the activated thrombin receptor

The thrombin receptor (hTR) is a membrane-bound glycoprotein, with an exposed N-terminal region and seven transmembrane segments [37]. Binding of thrombin leads to cleavage of the R41–S42 bond and may lead to formation of an extracellular loop by implantation of the newly exposed terminus into the membrane [38,39]. While the receptor remains in this activated form, the membrane-bound G-protein exercises its signalling function and numerous physiological responses are stimulated, including blood coagulation, inflammation and mitosis. The process continues until the receptor is deactivated by an unknown mechanism. Because prolonged activation of the receptor can have deleterious effects, it would be advantageous to understand as completely as possible the mechanisms for activation and deactivation of the receptor. Our collaborator, A. Kuliopoulos at Tufts-New England Medical Center, is exploring the hypothesis that deactivation



Scheme 3. Carbohydrate sequence and branching patterns assigned on the basis of fragment ions observed in the ESI-CID MS/MS spectra of the major (left) and minor (right) oligosaccharide glycoforms released from the CPP glycopeptide from rat pineal gland and permethylated [33].



Scheme 4. Amino acid sequence of the thrombin receptor exodomain (expressed in *E. coli* and cleaved from the carrier by CNBr), showing potential α-thrombin cleavage sites in italics. Single-letter amino acid codes are given; hS = homoserine. Peptides observed by MALDI-TOF MS are underlined [40].

occurs upon further protease cleavage within the receptor's exodomain. As part of the investigation, the products of proteolysis with the various proteases known to be present, including thrombin itself, are being determined by MALDI-TOF MS of fractions separated by HPLC. The kinetics governing the appearance and disappearance of the various peptide products are being followed by HPLC. Because the sequence of the starting protein (corresponding to the external portion of the receptor, residues 26–103, a segment that has been expressed in *E. coli*) is known (Scheme 4), the molecular weights of potential products from cleavage with each protease can be calculated. In Scheme 4, spacings illustrate the peptides that would result from cleavage of the receptor exodomain at all the potential α-thrombin cleavage sites, the italicized *R* and *K* residues. MALDI-TOF MS analysis of the proteolysis products can identify the peptide(s) present in each of the HPLC fractions. This provides information on which of the potential cleavages occur under the experimental conditions and facilitates the kinetic studies. The underlines with circles at their termini in Scheme 4 indicate the peptides experimentally observed in α-thrombin digests, showing that not all the cleavage sites are actually utilized [40].

The MALDI-TOF mass spectra are simple, dominated by $[M_x + H]^+$ ions corresponding to each of the components in the HPLC fraction. Additional low abundance peaks can be assigned to doubly-charged species $[M_x + 2H]^{2+}$ and dimers $[2M_x + H]^+$, which may include mixed pairs of the various species $[M_x + M_y + H]^+$. As an example, Fig. 4 shows the MALDI-TOF mass spectrum of an HPLC fraction obtained after plasmin cleavage of the segment corresponding to the activated receptor exodomain (residues 42–103). Major components in these fractions are present at the pmol level; since the useable dynamic range is large, minor components may easily be detected at low fmol levels. The mass spectra thus provide sensitive and specific information on protease digestion products and permit the use of HPLC for the necessary kinetic analyses [40].

3.5. Pegylation of proteins

In addition to the requirement to determine naturally-occurring posttranslational modification of proteins, the analyst may also be faced with the problem of determining the extent of deliberate chemical modifications undertaken to change the circulation half-life of a protein or its biophysical properties. One such modification is the substitution of the polymer polyethylene glycol (PEG) onto the amino group of lysine residues. When the protein has use as a pharmaceutical, this modification may prolong its circulation, lower the dosage amount and frequency and facilitate the maintenance of stable therapeutic levels. The modification also changes solubility properties, and permits studies and utilization of enzymes in organic solvents. Nonaqueous media can be useful to probe biochemically significant enzyme intermediates. In both cases, for the sake of reproducibility and data interpretation, it is important to

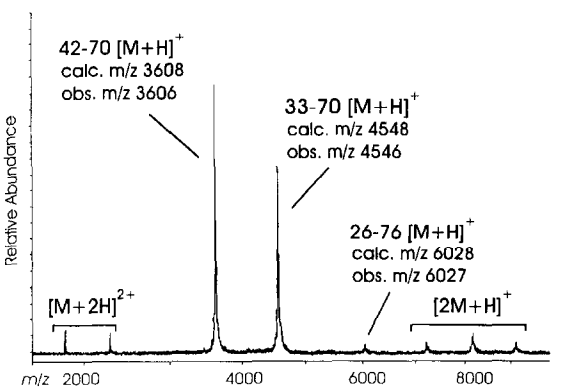


Fig. 4. MALDI-TOF mass spectrum of an HPLC fraction obtained after plasmin cleavage of the segment corresponding to the activated thrombin receptor exodomain (residues 42–103) [40].

ascertain the level of incorporation of the modifier. MALDI-TOF MS has been shown to be useful for this purpose [41,42].

Choice of measurement conditions (solvent, sample concentration, matrix, wavelength, laser power, analyzer) can influence results with this method; the selection becomes especially critical for the analysis of complex mixtures. Fig. 5(b) shows the ultraviolet (337 nm) MALDI-TOF mass spectrum obtained in the reflectron mode with 'super DHB' matrix (a 9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid [44]) for a sample of horseradish peroxidase modified under conditions selected to introduce only a few moles of PEG per mole of protein [41,43]. Peak width is limited by the half-width of the distribution of the PEG monomers. The distribution of the different species, spaced at intervals of 5000 daltons, the average molecular weight of the PEG, can easily be determined. This spectrum compares favorably with an earlier spectrum, Fig. 5(a), that we provided to this collaborator. The spectrum in Fig. 5(a) had been obtained for a similar sample using sinapinic acid matrix and a Vestec VT2000 linear MALDI-TOF MS, also operated at 337 nm [41].

3.6. Distribution of glycoprotein oligosaccharides

The extent of glycosylation and the glycoform distribution on a glycoprotein vary with the source of the glycoprotein and with the state of the organism or individual at the time the protein was processed—the patterns are affected, for example, by age, hormonal and immunological stresses, genetic mutations and specific tissue variations. By no means are the mechanisms nor all the effects of glycosylation changes yet fully understood. For structure–activity relationships to be established with a degree of confidence and pattern changes to be monitored, rapid and reliable methods for the determination of glycosylation patterns are required. In addition, since glycosylation can affect glycoprotein tissue distribution and half-life, the pattern expressed on recombinant proteins produced in batch processes for therapeutic uses should be determined initially for each protocol and monitored over the course of time. A simple approach for determination of carbohydrate moieties present on a glycoprotein (or other glycoconjugate) is the chemical or enzymatic release of the total oligosaccharide pool, followed by analysis of the released carbohydrates. Fig. 6 shows the ESI and MALDI mass spectra obtained for one such recombinant glycoprotein sample, after *N*-glycanase release of the oligosaccharides, followed by permethylation [45]. Once one includes a calculation that takes into account the distribution of components in the ESI spectrum over multiple charge states, the results of the two methods are in fact quite consistent. In these spectra, *N*-acetyl neuraminic acid (NA), known to be a labile group [46], was retained, protected as its fully methylated derivative.

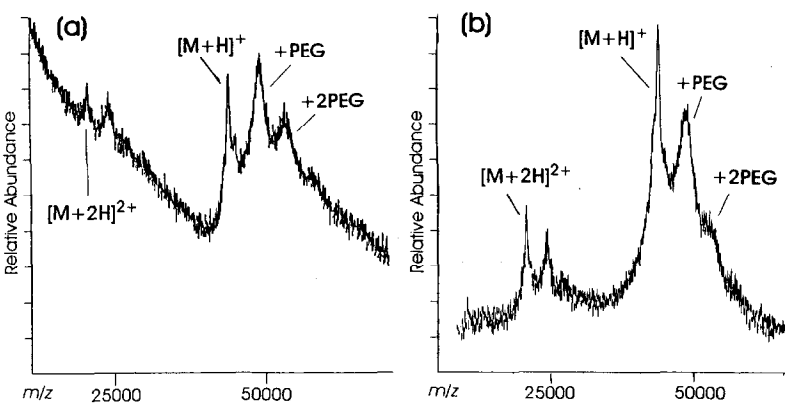


Fig. 5. MALDI-TOF mass spectra of different batches of horseradish peroxidase, $M_r = 43{,}400$, modified with polyethylene glycol, $M_w = 5000$, [41] dissolved in 0.5% trifluoroacetic acid recorded (a) for 3 pmol, in the linear mode with sinapinic acid matrix [41] and (b) for 200 fmol, in the reflector mode with 'super DHB' matrix [43].

In order to evaluate whether these methods give abundance results equivalent to the widely-used fluorophore-assisted carbohydrate electrophoresis method (FACE), we have analyzed a mixture of glucose oligomers ($n = 2 \rightarrow 10$) as their permethylated derivatives [45]. In Fig. 7, the relative abundances observed for the signals attributed to the several oligomers are plotted for each of the methods. The data indicate some variance for the smallest components that probably results from their volatility. At the high mass end of the distribution, the initial experiment (5 keV accelerating energy, 5 kV postacceleration before the detector) suggested discrimination in the MALDI-TOF mass spectrum. When the accelerating voltage was increased from 5 to 10 kV (or the postacceleration was raised), however, this phenomenon was eliminated, indicating that the overall energy of the ions impacting on the detector is important to maintain the sensitivity for larger species.

Compared to other types of mass spectral analysis, ESI and MALDI are less prone to competitive

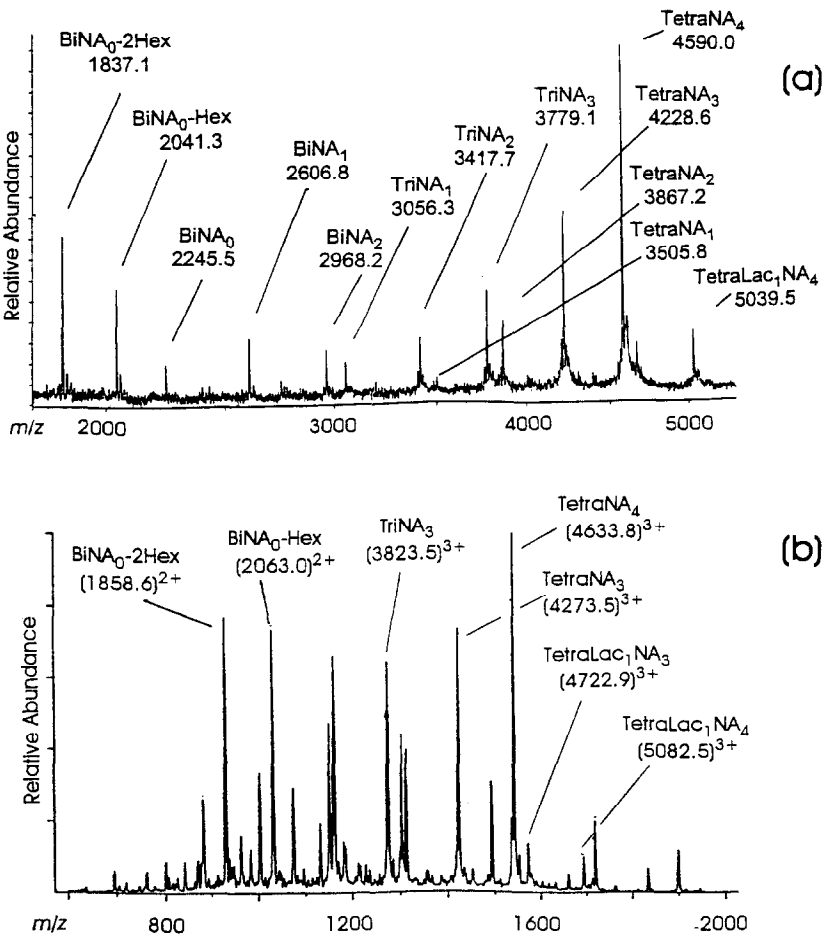


Fig. 6. Mass spectra of the oligosaccharide pool permethylated after *N*-glycanase release from a recombinant protein, LFA-3. Bi = biantennary; Tri = triantennary; Tetra = tetraantennary; NA = *N*-acetyl neuraminic acid; Lac = *N*-acetyllactosamine. Experimental values are given for mass assignments. (a) MALDI-TOF MS. All assignments correspond to $[M + Na]^+$. (b) ESI-MS. All assignments correspond to $[M + zNa]^{z+}$. Because of the space limitation within the figure, only the most abundant charge state has been marked for each component; undesignated peaks correspond to additional charge states of the various glycoforms. In order to obtain the relative abundance in the mixture for each component, the signals for all of its charge states (e.g., $z = 1$–4) must be summed.

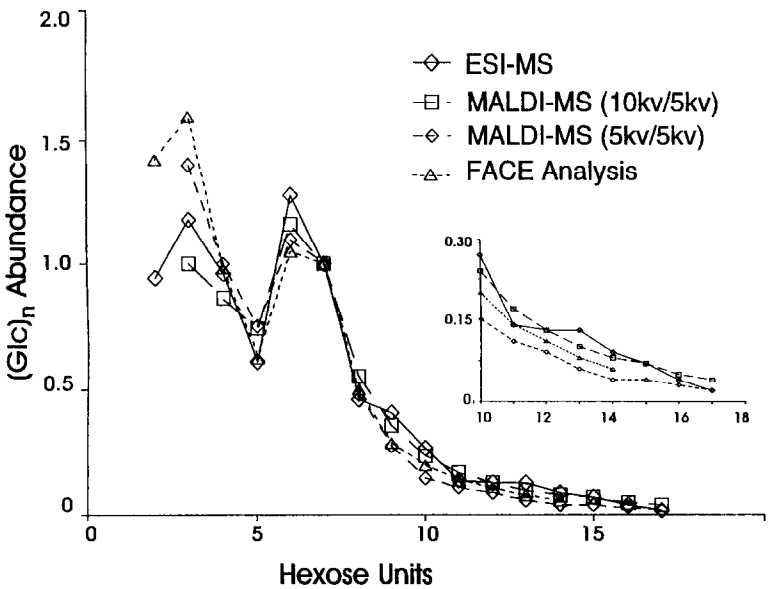
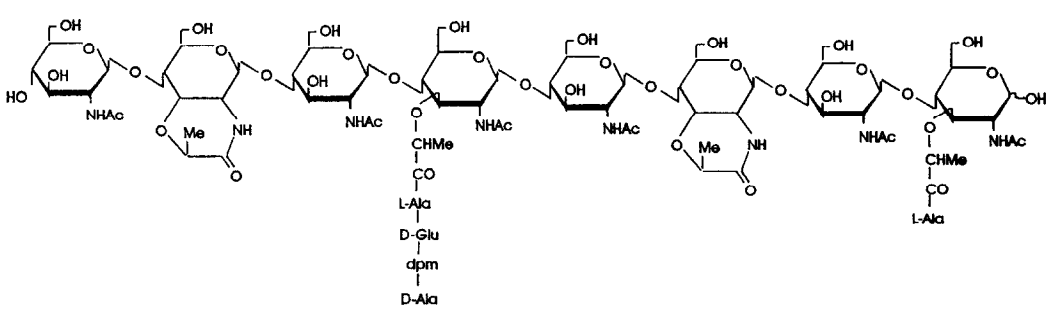


Fig. 7. Comparison of signal intensities observed for the fluorescence signal (FACE analysis) or $\Sigma_z[M + zNa]^{z+}$ of ions (MS) corresponding to components ($n = 2 \rightarrow 17$) for a permethylated (glucose)$_n$ polymer sample, each method's curve normalized against its value for $n = 7$.

effects in the ionization process, but some discrimination does still occur and its extent must be evaluated before quantitative measurements can be made. Carbohydrates bearing negatively-charged substituents (e.g., sulfate, phosphate) [47] or residues (e.g., *N*-acetyl neuraminic acid, glucuronic acid) [46] are more efficiently observed in the negative-ion mode.

For glycoproteins with multiple glycosylation sites, a more complete approach involves protease or chemical cleavage of the peptide backbone as a first step, because the glycosylation pattern is usually site-specific and information is lost by releasing the entire oligosaccharide pool at once. HPLC can provide separation of the individual glycosylation sites, and lectin affinity columns may be useful to separate glycotypes (e.g., biantennary vs. multiantennary, complex vs. high mannose). The molecular weight distributions of the eluted fractions may be determined for the glycopeptides [48], or for oligosaccharides released from them. Details of the structures may then be established using MS/MS [34], post-



Scheme 5. General structure for the muramyl peptides from bacterial spore peptidoglycan, proposed by Warth and Strominger [52].

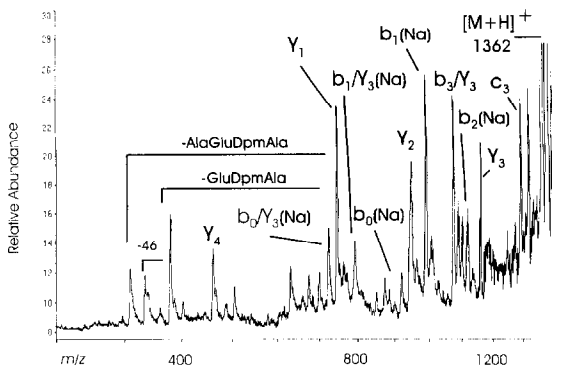


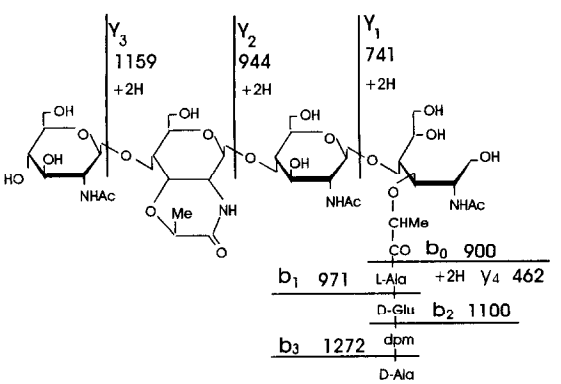
Fig. 8. MALDI-PSD-TOF mass spectrum of the $[M+H]^+$, m/z 1362, from a peptidoglycan tetrasaccharide tetrapeptide from a muramidase digest of wild-type *B. subtilis* spore cortex. Some contributions from the $[M+Na]^+$, m/z 1384, are also present. Peak assignments are shown in Scheme 6. Nitrogen laser, 337 nm. Accelerating voltage 10 kV.

source decay [7] and/or additional measurements to determine mass shifts after employing a series of exoglycosidases [49–51].

3.7. Muramyl peptides in wild and mutant bacterial spores

Gram-positive bacteria, e.g., *Clostridium*, *Bacillus*, form endospores as a means of shielding themselves from environmental challenges during a dormant state. The survival and long-term viability of the species may depend on the structure of the spore, and, in particular, on the peptidoglycan (PG) that makes up most of the bacterial cell's thick outer protective layer, the cortex [52,53]. Our collaborators, D. Popham and P. Setlow at the Univ. of Connecticut Health Center, have prepared cultures of wild and mutant bacteria in order to correlate variations in the endospore peptidoglycan with survival. The carbohydrate portion of the glycopeptide is composed of alternating muramic acid and *N*-acetyl glucosamine residues (Scheme 5); the muramic acid residues may be modified with amino acids or short peptides that can crosslink. Unsubstituted muramic acid residues can cyclize to lactams. In order to probe details of the cortex structure, it has been necessary to devise sensitive and specific methods to determine the nature and frequency of these crosslinks. Enzymatic digestion of the PG with muramidase, followed by HPLC separation of the released glycopeptides (after borohydride reduction to remove the anomeric heterogeneity) yields profiles of the PG components. MALDI-TOF MS analysis of the HPLC fractions provides molecular weight information that can be interpreted in terms of carbohydrate and amino acid composition, and reveals the presence of lactams and their reduction products. Because acidic residues (glutamic acid, diaminopimelic acid) occur frequently in these glycopeptides, they can efficiently generate stable carboxylate anions. Negative-ion MALDI-TOF analysis is therefore particularly useful for highly sensitive molecular weight determinations that are little affected by residual contaminants.

Previous reports had demonstrated that the structures of muramyl peptides may be determined at the nmol level by liquid secondary ionization [LSI] tandem mass spectrometry [54]. Our recent experiments have shown that MALDI-PSD-TOF MS yields similar information on sequence and cross-linking, but at the pmol level [55]. The PSD spectrum of a PG tetrasaccharide tetrapeptide from wild-type *Bacillus* spore cortex is given in Fig. 8. Because precursor ion selection was not employed in this case, fragments from both the $[M+H]^+$ ion, m/z 1362, and the $[M+Na]^+$ ion, m/z 1384 were observed. Most fragments originated from the less stable protonated species, as expected. Fragmentation in both the oligosaccharide and peptide portions of the structure permits the entire sequence to be deduced, and pres-



Scheme 6. Assignments of product ions in the MALDI-PSD-TOF mass spectrum (Fig. 9) of a tetrasaccharide tetrapeptide from peptidoglycan of wild-type *Bacillus* [55].

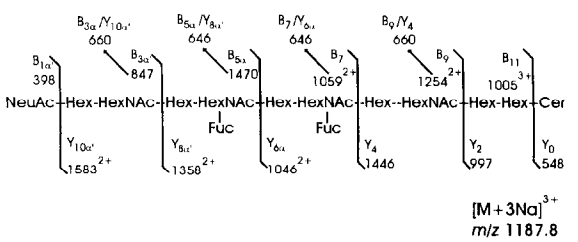
ence of the lactam to be observed, as indicated in Scheme 6. In this scheme, upper-case letters designate carbohydrate fragmentations [35] and lower-case letters designate peptide bond cleavages [56]. This sensitivity makes feasible determinations that would otherwise have required preparation and extraction of prohibitive amounts of the bacterial cultures.

These investigations have verified some of the assumptions made in previous work and provide tools for wider studies [53]. One recent phase of this continuing collaboration showed that a long-held assumption is incorrect, when we determined that cortex PG cross-linking is necessary but does not control the rate of dehydration in the dormant state, and is instead critically related to the success of reproduction following a dormant period [57].

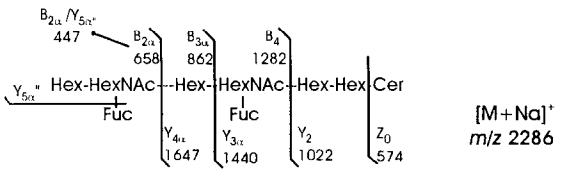
3.8. Studies of glycosphingolipids

Glycosphingolipids (GSLs) play many roles, from nervous system signalling to intercellular communications [58]. Heterogeneity is present in both the carbohydrate and lipid portions, and varies with tissue source, age, and immune status. Molecular weight profiles of such materials can be obtained by either MALDI-TOF MS [59,60] or ESI-MS [61]. Structural details must then be determined from fragments arising by CID or PSD.

Unusual structures are encountered in some instances, such as cell-surface GSLs on tumor tissues. For these, widely-used methods that base identifications on chromatographic behavior alone or in combination with exoglycosidase digestions are insufficient to specify unusual sequences and branching patterns. In a collaborative study carried out jointly with S.-I. Hakomori et al., at the Biomembrane



Scheme 7. Structure and CID product ion mass assignments for the permethylated derivative of a ganglioside from human myelogeneous leukemia HL60 cells that reacts with E-selectin [63].



Scheme 8. Structure and PSD product ion mass assignments for the permethylated derivative of a neutral glycosphingolipid from foetal calf brain [64].

Institute, Seattle, WA, human myelogeneous leukemia HL60 cells that bind to E-selectin and possibly P-selectin were shown to contain gangliosides with extended *N*-acetyllactosamine structures and varied fucosylations that were absent in normal leukocytes [62,63]. These represent a new type of epitope for E-selectin binding, different from the well-known sialylated Lewis^x ligand, Neu5NAcα2 $\rightarrow$ 3Galβ1 $\rightarrow$ 4(Fucα1 $\rightarrow$ 3)GlcNAcβ1 $\rightarrow$ 3R. The structures were elucidated based on ESI-MS/MS of major and minor GSLs, together with fast atom bombardment MS and NMR analysis of the more abundant components. The structure shown in Scheme 7 is representative of the set.

For an ongoing study that aims at delineating fucosyltransferase substrates in the developing brain, conducted in collaboration with R.H. McCluer et al. at the Shriver Center for Mental Retardation, Waltham, MA, it has been necessary to optimize methods for the analysis of GSLs at the low pmol level. To achieve sample cleanup and maximize sensitivity, the samples have been subjected to peracetylation, followed by permethylation. Molecular weight determinations have been made by LSI-MS on a four-sector tandem mass spectrometer and MALDI-TOF MS. Although LSI-MS/MS can be used for structural determinations of the more abundant model compounds, the sample amounts from foetal calf brain were too low for this method. Our model studies showed that MALDI-PSD-TOF MS can be useful for this class of compounds. PSD was performed on a sample corresponding to less than 700 fmol of one of the unknown calf brain GSLs and thereby enabled elucidation of its carbohydrate moiety (Scheme 8) [64]. PSD thus represents a valuable new tool for high sensitivity glycolipid structural determinations.

3.9. Emerging approaches for biopolymer analysis: Infrared-MALDI-TOF MS and FT-ICR-MS

MALDI-TOF mass spectra of ions desorbed with an infrared laser exhibit abundant multiply-charged species and multimers [65]. Decrease in metastable decomposition leads to improved peak shapes compared to UV-MALDI, especially noticeable for large molecules. As an example, Fig. 9 shows the 2.94 μm IR-MALDI mass spectrum of bovine serum albumin, recorded in the reflector mode, with succinic acid matrix. Infrared irradiation has been demonstrated to yield superior results for analysis of proteins and glycoproteins electroblotted onto transfer membranes, perhaps because of the deeper penetration of the infrared beam into the target surface [66].

The mass range for the TOF analyzer is theoretically unlimited, but is bounded by sample behavior and detector response. The practical upper mass limit for MALDI-TOF MS is around 1×10^6 daltons, a figure that has been achieved for IgM [67], but most bipolymer measurements to date have been in the range below 400 kDa. For industrial types of polymers, MALDI-TOF MS measurements of high mass compounds, even exceeding 1 MDa have been reported, although the dominant species for the very large ions bore two or more charges [68,69]. The mass range with other types of analyzers is more limited; for example, the highest singly-charged biopolymer ions observed by FT-ICR MS thus far are oligonucleotide 25-mers at m/z 7634 [70]. Multiple-charging in ESI-FT-ICR mass spectra extends the effective mass range, so that spectra of two different porcine serum albumins, M_r 66,736 and 66,886, have been obtained with isotopic resolution [71]. These instruments offer the possibility for multiple stages of mass analysis as well as the advantage of very high resolution and mass accuracy. Masses that are measured with the high accuracy provided by FT-ICR MS, sector instruments, or TOF MS with delayed extraction yield information on elemental composition in the lower mass ranges, and narrow the database search windows for larger compounds such as proteins [72].

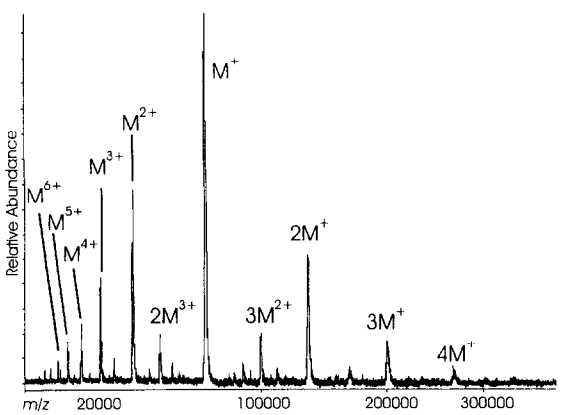


Fig. 9. MALDI-TOF mass spectrum of 1 pmol bovine serum albumin, M_r = 66,431, recorded with an infrared laser (2.94 μm) in the reflector mode, with succinic acid matrix.

4. Concluding remarks

Powerful new mass spectral ionization techniques, in particular electrospray and matrix-assisted laser desorption, together with microscale methods for sample preparation and improvements in mass analyzers, can have impact on wide areas of biological science. This brief survey of a few of the projects in which our own laboratory has recently been engaged is meant to provide a window on the types of 'real-world' problems that can now be approached. Other important areas include characterization of oligonucleotides and of non-covalent complexes. Developments are expected to continue at a rapid pace worldwide, increasing even further the reach of mass spectrometry into biology and medicine.

Acknowledgements

The author wishes to thank her colleagues J. Helin, X. Hronowski, P. Juhasz, L.C.N. Ngoka, H. Perreault, B.B. Reinhold, V.N. Reinhold and S. Ye, and the collaborators named in the text, for their contributions to the work described. The research was supported by the National Institutes of Health Grants P41 RR10888 and S10 RR10493 (to C.E. Costello), P41 RR00317 (to K. Biemann); R29 GM52926 (to A. Kuliopoulos) and R01 CA42505 (to S.-I. Hakomori); NATO Collaborative Grant 900605

(to C.E. Costello); National Science Foundation Grant No. MCB 9400633 (to V.N. Reinhold); the Vrie Univ., Amsterdam; the Swedish Medical Research Council (to R. Ekman) and Thermo BioAnalysis. J. Helin and H. Perreault received fellowships from the Academy of Finland and the National Research Council of Canada, respectively.

References

[1] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Proc. 78 (1987) 53.
[2] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.
[3] A. Varki, Glycobiology 3 (1993) 97.
[4] K.G. Standing, R.C. Beavis, G. Bollbach, W. Ens, F. Lafortune, D. Main, B. Schueler, X. Tang, J.B. Westmore, Anal. Instrumen. 16 (1987) 173.
[5] X. Tang, W. Ens, K.G. Standing, J.B. Westmore, Anal. Chem. 60 (1988) 1791.
[6] R. Kaufmann, B. Spengler, F. Lützenkirchen, Rapid Commun. Mass Spectrom. 7 (1993) 902.
[7] B. Spengler, D. Kirsch, R. Kaufmann, J. Lemoine, J. Mass Spectrom. 30 (1995) 782.
[8] W.C. Wiley, I.H. McLaren, Rev. Sci. Instr. 26 (1953) 1150.
[9] R.M. Whittal, L. Li, Anal. Chem. 67 (1995) 1950.
[10] R.S. Brown, J.J. Lennon, Anal. Chem. 67 (1995) 1998.
[11] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fostis, M. Mann, Nature 379 (1996) 466.
[12] K.J. Light-Wahl, B.L. Schwartz, R.D. Smith, J. Am. Chem. Soc. 116 (1994) 5271.
[13] N.L. Kelleher, C.A. Costello, T.P. Begley, F.W. McLafferty, J. Am. Soc. Mass Spectrom. 6 (1995) 981.
[14] H. Görisch, Anal. Biochem. 173 (1988) 393.
[15] J.C. Rouse, J.E. Vath, Anal. Biochem. 238 (1996) 82.
[16] E. Nordhoff, A. Ingendoh, R. Cramer, A. Overberg, B. Stahl, M. Karas, F. Hillenkamp, P.F. Crain, Rapid Commun. Mass Spectrom. 6 (1992) 771.
[17] I. Ciucanu, F. Kerek, Carbohydr. Res. 131 (1984) 209.
[18] K.B. Linsley, S.-Y. Chan, S. Chan, B.B. Reinhold, P.J. Lisi, V.N. Reinhold, Anal. Biochem. 219 (1994) 207.
[19] H. Perreault, C.E. Costello, Org. Mass Spectrom. 29 (1994) 720.
[20] D. Liu, C. Dass, G. Wood, D.M. Desiderio, J. Chromatogr. 500 (1990) 395.
[21] R. Ekman, Nord. Psykiatr. Tidskr. 39 (1985) 31.
[22] J.Z. Chou, M.J. Kreek, B.T. Chait, J. Am. Soc. Mass Spectrom. 5 (1994) 10.
[23] C.R. Jiménez, P.A. van Veelen, K.W. Li, W.C. Wildering, W.P.M. Geraerts, U.R. Tjaden, J. van der Greef, J. Neurochem. 62 (1994) 404.
[24] P. Adrén, R.M. Caprioli, J. Mass Spectrom. 30 (1995) 817.
[25] R. Ekman, P. Juhasz, M. Heili, H. Ågren, C.E. Costello, Peptides 17 (1996) 1107.
[26] K. Tatemoto, M. Carlquist, V. Mutt, Nature 296 (1982) 659.
[27] S.H.C. Hendry, in: W.F. Colmers, C. Wahlestedt (Eds.), The Biology of Neuropeptide Y. Humana Press, Totowa, NJ, 1993, p. 65.
[28] B.G. Stanley, in: W.F. Colmers, C. Wahlestedt (Eds.), The Biology of Neuropeptide Y. Humana Press, Totowa, NJ, 1993, p. 457.
[29] M. Heilig, G.F. Koob, R. Ekman, K.T. Britton, TINS 17 (1994) 80.
[30] E. Widerlöv, L.H. Lindström, C. Wahlestedt, R. Ekman, J. Psychiat. Res. 22 (1988) 69.
[31] W.H. Berrettini, A.R. Doran, J. Kelsoe, A. Roy, D. Pickar, Neuropsychopharmacology 1 (1987) 81.
[32] L. Grundemar, S.P. Sheikh, C. Wahlestedt, in: W.F. Colmers, C. Wahlestedt (Eds.), The Biology of Neuropeptide Y. Humana Press, Totowa, NJ, 1993, p. 197.
[33] C.R. Jiménez, K.W. Li, K. Dreisewerd, H.D. Mansvelder, A.B. Brussaard, B.B. Reinhold, R. Kingston, R.C. van der Schors, M. Karas, F. Hillenkamp, J.P.H. Burbach, C.E. Costello, P.M. Cevaerts, Proc. Natl. Acad. Sci, in press.
[34] V.N. Reinhold, B.B. Reinhold, C.E. Costello, Anal. Chem. 67 (1995) 1772.
[35] B. Domon, C.E. Costello, Glycoconj. J. 5 (1988) 397.
[36] A.S. Angel, B. Nilsson, Biomed. Environ. Mass Spectrom. 19 (1990) 721.
[37] T.-K. Vu, D.T. Hung, V.I. Wheaton, S.R. Coughlin, Cell 64 (1994) 1057.
[38] R.E. Gerszten, J. Chen, M. Ishii, K. Ishii, L. Wang, T. Nanevicz, C.W. Turk, T.-K.H. Vu, S.R. Coughlin, Nature 368 (1994) 648.
[39] T. Nanevicz, M. Ishii, L. Wang, M. Chen, C.W. Turk, F.E. Cohen, S.R. Coughlin, J. Biol. Chem. 270 (1995) 21619.
[40] A. Kuliopulos, P.J. Sheridan, S.K. Seeley, J. Helin, C.E. Costello, Abstr. Am. Soc. Hematology, November, 1996.
[41] P.A. Mabrouk, J. Am. Chem. Soc. 117 (1995) 2141.
[42] S.K. Chowdhury, M. Doleman, D. Johnston, J. Am. Soc. Mass Spectrom. 6 (1995) 478.
[43] C.E. Costello, J. Helin, L.C.M. Ngoka, Progr. Biomed. Opt. 2680 (1996) 406.
[44] M. Karas, H. Ehring, E. Nordhoff, B. Stahl, K. Strupat, F. Hillenkamp, M. Grehl, B. Krebs, Org. Mass Spectrom. 28 (1993) 1476.
[45] B.B. Reinhold, S. Ye, J. Helin, V.N. Reinhold, C.E. Costello, Proc. ASMS Conf. on Mass Spectrom. Allied Topics 44 (1996) 222.
[46] M.C. Huberty, J.E. Vath, W. Yu, S.A. Martin, Anal. Chem. 65 (1993) 2791.
[47] G.R. Hayes, A. Williams, C.E. Costello, C.A. Enns, J.J. Lucas, Glycobiology 5 (1994) 227.
[48] M.J. Treuheit, C.E. Costello, H.B. Halsall, Biochem. J. 283 (1992) 105.
[49] B. Stahl, T. Klabunde, H. Witzel, B. Krebs, M. Steup, M. Karas, F. Hillenkamp, Eur. J. Biochem. 220 (1994) 321.
[50] E. Mørtz, T. Sareneva, I. Julkunen, P. Roepstorff, J. Mass Spectrom. 31 (1996) 1109.
[51] R. Küster, T.J.P. Naven, D.J. Harvey, J. Mass Spectrom. (1996) 1131.

[52] A.D. Warth, J.L. Strominger, Biochemistry 11 (1972) 1389.
[53] D.L. Popham, J. Helin, C.E. Costello, P. Setlow, J. Bacteriol. 178 (1996) 6451.
[54] S.A. Martin, R.S. Rosenthal, K. Biemann, J. Biol. Chem. 262 (1987) 7514.
[55] J. Helin, C.E. Costello, D. Popham, P. Setlow, Proc. ASMS Conf. on Mass Spectrom. Allied Topics 44 (1996) 1000.
[56] K. Biemann, in: J.A. McCloskey (Ed.), Mass Spectrometry (Methods in Enzymology, Vol. 193) Academic Press, San Diego, 1990, p. 886.
[57] D.L. Popham, J. Helin, C.E. Costello, P. Setlow, Proc. Natl. Acad. Sci. USA 93 (1996) 15405.
[58] S.-I. Hakomori, J. Biol. Chem. 265 (1990) 18713.
[59] H. Egge, J. Peter-Katalinic, M. Karas, B. Stahl, Pure Appl. Chem. 69 (1991) 491.
[60] P. Juhasz, C.E. Costello, J. Am. Soc. Mass Spectrom. 3 (1992) 785.
[61] B.B. Reinhold, S.-Y. Chan, S. Chan, V.N. Reinhold, Org. Mass Spectrom. 29 (1994) 736.
[62] M.R. Stroud, K. Handa, M.E.K. Salyan, K. Ito, S.B. Levery, S.-I. Hakomori, B.B. Reinhold, V.N. Reinhold, Biochemistry 35 (1996) 758.
[63] M.R. Stroud, K. Handa, M.E.K. Salyan, K. Ito, S.B. Levery, S.-I. Hakomori, B.B. Reinhold, V.N. Reinhold, Biochemistry 35 (1996) 770.
[64] H. Perreault, O. Koul, J. Street, R.H. McCluer, C.E. Costello, Proc. ASMS Conf. on Mass Spectrom. Allied Topics 43 (1995) 1179.
[65] A. Overberg, M. Karas, U. Bahr, R. Kaufmann, F. Hillenkamp, Rapid. Commun. Mass Spectrom. 4 (1990) 293.
[66] K. Strupat, C. Eckerskorn, M. Karas, F. Hillenkamp, in: A.L. Burlingame, S.A. Carr (Ed.), Mass Spectrometry in the Biological Sciences, Humana Press, Clifton, NJ, 1995, p. 365.
[67] R.W. Nelson, D. Dogruel, P. Williams, Rapid. Commun. Mass Spectrom. 9 (1995) 625.
[68] P.O. Danis, D.E. Carr, Org. Mass Spectrom. 28 (1993) 923.
[69] D.C. Schriemer, L. Li, Anal. Chem. 68 (1996) 2721.
[70] Y. Li, K. Tang, D.P. Little, H. Köster, R.L. Hunter, R.T. McIver Jr., Anal. Chem. 68 (1996) 2090.
[71] F.W. McLafferty, J. Mass Spectrom. 30 (1995) 39.
[72] O.N. Jensen, A. Podtelejnikov, M. Mann, Rapid Commun. Mass Spectrom. 10 (1996) 1371.