





This is to certify that the

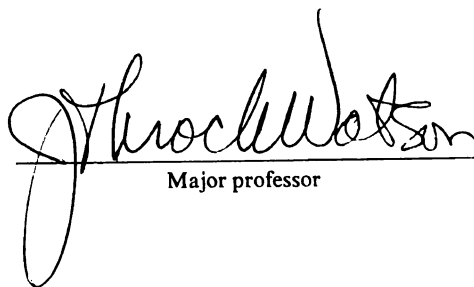
dissertation entitled

Structural Characterization of Glycoconjugates and
Proteins by Matrix-assisted Laser Desorption
Ionization Mass Spectrometry
presented by

Naxing Xu

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Chemistry


Major professor

Date 1 July 1997

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
 TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

MSU Is An Affirmative Action/Equal Opportunity Institution

c:\cric\data\due.pm3-p.1

STRUCTURAL CHARACTERIZATION OF GLYCOCONJUGATES AND PROTEINS
BY MATRIX-ASSISTED LASER DESORPTION IONIZATION MASS
SPECTROMETRY

by

Naxing Xu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1997

ABSTRACT

STRUCTURAL CHARACTERIZATION OF GLYCOCONJUGATES AND PROTEINS BY MATRIX-ASSISTED LASER DESORPTION IONIZATION MASS SPECTROMETRY

by Naxing Xu

Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) has quickly emerged as a powerful analytical tool for sensitive and accurate molecular weight determination of biomolecules. In this research, 2-mercaptobenzothiazole (MBT) and its analogs have been discovered as a class of new matrices for MALDI-MS. These compounds are structurally distinct from the conventional matrices. The information gained from the matrix studies has been applied to the structural characterization of peptidoglycans, a type of complex glycoconjugates consisting bacterial cell wall with important immunological significance. 5-Chloro-2-mercaptobenzothiazole (CMBT) was employed for the routine MALDI-MS analysis of mucopeptides derived from peptidoglycans and structural information of mucopeptide was obtained by further post-source decay (PSD) analysis.

Glycoprotein characterization has been a challenging topic in bioanalytical field because their complexity and microheterogeneity. To minimize sample handling between the stages of lectin affinity binding and MALDI-MS, a micro-batch lectin binding process was also developed for direct MALDI-MS analysis for isolated glycopeptides. This approach is capable of identifying glycopeptides derived from ribonuclease B and avidin, proteins having a single glycosylation site.

11
12
13
14
15
16
17
18
19
20

A new glycosylamine derivative of oligosaccharides has been developed for facile sequence analysis in MALDI-MS. 3-Aminoquinoline was used as the derivatizing reagent in making glycosylamine derivative at the reducing terminus. This type of glycosylamine derivatives produces primarily protonated molecules in MALDI-MS. Reaction cleanup is not necessary because excess reagent can be used as MALDI matrix without interfering MALDI-MS analysis. PSD analysis of the protonated molecule yields a spectrum dominated by reducing terminal fragments, allowing easy determination of the oligosaccharide sequence, which is required for the complete characterization of glycoconjugates and glycoproteins.

[illegible]

ACKNOWLEDGMENT

I would like to acknowledge Dr. Zhi-Heng Huang and Dr. Douglas A. Gage for their expert assistance in organic chemistry and biochemistry and enormous encouragement they have given me during my five years at Michigan State. A special thanks to the members of the staff of the MSU-NIH Mass Spectrometry Facility: Mel Micke for fixing the computer, Melinda Berning for fixing the paper work, and Bev Chamberlin for fixing JEOL 505, morning-coffee and delicious cakes. A thank-you to Mike Davenport and Tony Sills for their help with the ion-trap and Finnigan 4600 instruments, to Joe Hiller, Pao-chi Liao and Tun-li Shen for sharing their experience on MALDI-MS. Thanks to Jian Wang, Huanqin Wu, and Kuruppu Dharmasiri for the useful discussion and making ASMS hotel reservations. I would also like to thank other Watson group members, Jiang Wu, Ken Roth, Jennifer Johnson, Ben Gardner, Ying Yang, and Qing Li, who made my journey through through graduate school a lot easier. And finally, I would like to thank Dr. Watson for his academic advice, research support, the freedom he allowed me in the lab, and for those great salmon steaks.

1175

1176

1177

1178

1179

1180

1181

1182

1183

1184

1185

1186

1187

1188

1189

1190

1191

1192

1193

1194

1195

1196

TABLE OF CONTENTS

List of Figures	vii
List of Tables	xiii
Key to Abbreviations	xiv
CHAPTER 1. INTRODUCTION	1
1. Recent Advances of Mass Spectrometry of Biological Molecules	1
2. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry	2
2.1 Operating Principles	3
2.2 Sample Preparation	4
2.3 Post-source Decay (PSD)	8
2.4 Delayed Extraction (DE)	11
3. Glycoconjugate and Protein Analysis by MALDI-MS	14
3.1 Status of Glycoconjugate Analysis	14
3.2 Status of Protein Analysis	15
References	18
CHAPTER 2. Mercaptobenzothiazoles: A New Class of Matrices for Laser Desorption Ionization Mass Spectrometry	22
1. Introduction	22
2. Experimental	27
3. Results and Discussion	29
3.1 Peptides and Proteins	29
3.2 Carbohydrates	37
3.3 Synthetic Polymers	47
4. Conclusions	49
References	50
CHAPTER 3. Structural Characterization of Peptidoglycan Muropeptides by MALDI-MS and Post-Source Decay Analysis	53
1. Introduction	53
2. Experimental	56
3. Results and Discussion	58
3.1 Molecular Weight Determination	59
3.2 Analysis by MALDI-PSD	65
3.3 Lysostaphin Digestion and Mass Mapping Analysis	70
4. Conclusions	77
References	78
CHAPTER 4. Structural Characterization of Glycoproteins by MALDI-MS	81
1. Introduction	81
1.1 Glycoproteins	81

1.2 Glycoproteins Analysis	86
1.3 Fractionation and Structural Assessment of Glycoconjugates by Lectin Immobilized Column Chromatography	88
2. Experimental	93
3. Results and Discussion	96
3.1 Structural Characterization of Porcine Oocyte Zona Pellucida Glycoprotein 3 α (ZP3 α)	96
3.2 A Micro-batch Lectin Affinity Binding Process for Direct MALDI-MS Analysis	109
4. Conclusions	120
References	121
CHAPTER 5. A New Glycosylamine Derivative for Sequence Analysis of Oligosaccharides by MALDI-PSD-MS and ESI-MS	
1. Introduction	125
2. Experimental	129
3. Results and Discussion	132
4. Conclusions	146
References	147
CHAPTER 6. Summary	
References	154

List of Figures

Figure 1.1	Voyage-Elite MALDI time-of-flight mass spectrometer	5
Figure 1.2	Differentiation of PSD fragment ion, M_1^+ , from the intact ionized molecule MH^+ by using the reflector. M_2 is the neutral dissociated from MH^+ . With no potential applied to the reflector (top panel), M_1^+ and M_2 maintain the initial velocity of MH^+ and reach the linear detector, L, simultaneously. With potential applied to the reflector (low panel), M_1^+ is reflected to the reflectron detector, R, and recorded while M_2 still hit the linear detector.	10
Figure 1.2	Comparison of MALDI mass spectra of ACTH (18-39) obtained in linear mode without DE (Linear), linear mode with DE (Linear DE), and reflectron mode with DE (reflectron DE) showing the effect of DE on mass resolution of TOF-MS	10
Figure 2.1	Light micrographs (magnification: $\times 14$) obtained from samples of bovine pancreatic insulin in different matrices (molar ratio of insulin to matrix: 1:1000). top left panel: MBT; top right panel: CMBT; (c) bottom left panel: DHB; bottom right panel: α CHCA.	30
Figure 2.2	Positive MALDI mass spectra of three proteins with MBT as matrix. (a) human angiotensin I (MW 1,296.5 Da); (b) porcine pancreastatin (MW 5,103 Da); (c) bovine transferrin (MW 78000 Da). The data were acquired by summing 30 to 100 laser shots. The amount of proteins loaded on the sample plate was about 1 to 5 picomoles.	32
Figure 2.3	Negative ion MALDI spectrum of bovine serum albumin (MW 66,430 Da) with MBT as matrix; 200 laser shots summed. Total protein loaded: 5 pmol	34
Figure 2.4	Positive MALDI spectrum of horse skeletal muscle myoglobin (10 pmol); 100 laser shots summed. The sample was dissolved in acetonitrile/0.1% (v/v) aqueous trifluoroacetic acid (1:1) solution. (a) no sodium dodecylsulfate (SDS) added, matrix: MBT; (b) with the addition of 0.1% SDS, matrix: MBT; (c) with the addition of 10% SDS, matrix: MBT	35
Figure 2.5	Positive MALDI spectrum of bovine pancreatic insulin (MW 5,733.5 Da) using CMBT as matrix; 30 laser shots summed. The region of the spectrum shown corresponds to a mass range of 5,500-6,150. Total protein loaded: 2 pmol	38

Figure 2.6	Positive MALDI spectrum of oligomannose-type N-linked oligosaccharide [Man] ₅ [GlcNAc] ₂ (MW 1,235.1 Da); 40 laser shots summed. (a) sample loaded: 2 pmol, matrix: CMBT; (b) sample loaded: 2 pmol, matrix: DHB; (c) sample loaded: 100 fmol, matrix: CMBT	40
Figure 2.7	Positive MALDI spectrum of dextran 5,000 with CMBT as matrix; 50 laser shots summed. Adjacent oligomer peaks in the spectrum are spaced by 162 mass units	42
Figure 2.8	Negative MALDI spectrum of underivatized G _{T1b} with CMBT as matrix; 40 laser shots summed. M ₁ and M ₂ are two molecular species in the sample differing by two methylene units in the long-chain base. [M ₁ + K - 2H] ⁺ and [M ₂ + K - 2H] ⁺ of two molecular species are represented by dominant peaks in the spectrum	44
Figure 2.9	Structure of trimeric mucopeptide derived from the peptidoglycan of <i>S. aureus</i> strain COL. The following abbreviations were used: G, <i>N</i> -acetylglucosamine; M, <i>N</i> -acetylmuramic acid	45
Figure 2.10	Positive MALDI spectrum of trimeric mucopeptide derived from peptidoglycan of <i>S. aureus</i> strain COL with (a) CMBT (b) α CHCA as matrix. 40 laser shots summed. Total sample loaded was approximately 3 pmol	46
Figure 2.11	Positive MALDI spectrum of PEG 3,350 with AMBT as matrix. 30 laser shots summed. No alkaline salts were added to the polymer solution	48
Figure 2.12	Positive MALDI spectrum of PEG 10,000 with AMBT as matrix. 100 laser shots summed. NaCl was added to the polymer solution	48
Figure 3.1	(a) The <i>Staphylococcus aureus</i> bacterial cell wall peptidoglycan. (b) The repeating unit of peptidoglycan is a GlcNAc-MurNAc disaccharide whose lactyl side chain forms an amide bond with a peptide	54
Figure 3.2	Muropeptides derived from peptidoglycan of the highly methicillin-resistant <i>Staphylococcus aureus</i> strain COL. (a) unsubstituted monomer; (b) pentaglycine-substituted monomer (c) muropeptide oligomers (n = 1, 2, 3, ...). The following abbreviations were used: GlcNAc, <i>N</i> -acetylglucosamine; MurNAc, <i>N</i> -acetylmuramic acid; Ala, alanine; iGln, isoglutamine;	

	Lys, lysine; Gly, glycine	60
Figure 3.3	Linear MALDI mass spectra of an unsubstituted monomer obtained in (a) positive and (b) negative ion mode with CMBT as matrix. The amount of analyte loaded on the sample plate was <i>ca.</i> 2 picomoles	61
Figure 3.4	Linear MALDI mass spectra of pentaglycine-substituted monomer obtained in (a) positive and (b) negative ion mode with CMBT as matrix. The amount of analyte loaded on the sample plate was approximately 2 picomoles. Peaks labeled with an asterisk are separated by 57 mass units representing mucopeptide impurities in the sample that are tetra-, tri-, di- glycine substituted	63
Figure 3.5	Composite positive ion linear MALDI mass spectra of mucopeptides ranging from monomer to pentamer with CMBT as matrix. The amount of analyte loaded on the sample plate was between 1 to 5 picomoles	64
Figure 3.6	Positive ion MALDI-PSD mass spectrum of an unsubstituted monomer. Matrix: CMBT. Sample loading: 2 picomoles	66
Figure 3.7	Positive ion MALDI-PSD mass spectrum of a pentaglycine-substituted monomer. Matrix: CMBT; Sample loading: 2 picomoles	68
Figure 3.8	Negative ion MALDI-PSD mass spectrum of a pentaglycine-substituted monomer. Matrix: CMBT; Sample loading: 2 picomoles	69
Figure 3.9	(a) Negative ion MALDI mass spectrum of the lysostaphin digest of (b) mucopeptide dimer I. Matrix: CMBT. The digestion mixture was used directly for MALDI-MS analysis without further purification. The arrows in Fig. 3.9b indicate the favorable sites of hydrolysis, based on the intensities of peaks in the MALDI mass spectrum of the lysostaphin digest. Thicker arrows indicate more favorable hydrolysis sites	72
Figure 3.10	Negative ion MALDI mass spectrum of the lysostaphin digest of (b) mucopeptide dimer II. Matrix: CMBT	75
Figure 4.1	<i>N</i> -linked oligosaccharides. (a) All <i>N</i> -glycosidic protein attachments occur through a β - <i>N</i> -acetylglucosamino-Asn bond in which the Asn occurs in the sequence Asn-X-Ser/Thr. (b) <i>N</i> -linked oligosaccharides usually have the branched (Man) ₃ (GlcNAc) ₂ core structure	82

Figure 4.2	Structures of major types of N-linked oligosaccharides. The box area encloses the core structure common to all N-linked oligosaccharides	84
Figure 4.3	Most common <i>O</i> -glycosidic attachments of oligosaccharides to glycoproteins	85
Figure 4.4	Amino acid sequence of porcine oocyte zona pellucida Glycoprotein 3 α (ZP3 α) glycoproteins	98
Figure 4.5	(a) HPLC of tryptic digest of ZP3 α glycoprotein. (b) HPLC of the glycopeptide fractions from ZP3 α glycoprotein after lectin (jacalin) affinity chromatography	100
Figure 4.6	<i>N</i> -linked glycopeptides at Asn 220 of ZP3 α glycoprotein. The MALDI-MS spectrum was obtained from a collected HPLC fraction	101
Figure 4.7	<i>O</i> -linked glycopeptides of ZP3 α glycoprotein after purification by lectin (jacalin) affinity chromatography and HPLC	102
Figure 4.8	<i>N</i> -linked Glycopeptides at Asn 220 of ZP3 α Glycoprotein. The MALDI-MS spectrum was obtained on (a) after β -galactosidase treatment. (b) after β -galactosidase & <i>N</i> -acetylhexoaminidase treatment	104
Figure 4.9	<i>N</i> -linked Glycopeptides at Asn 220 of ZP3 α Glycoprotein. Observed <i>m/z</i> and calculated <i>m/z</i> were assigned to the [M+H] ⁺ ions of the glycopeptides	106
Figure 4.10	<i>N</i> -linked Glycopeptides at Asn 333 of ZP3 α Glycoprotein. (a) MALDI-MS spectrum was obtained on original HPLC fraction. (b) MALDI-MS spectrum was obtained after β -galactosidase & <i>N</i> -acetylhexoaminidase treatment	107
Figure 4.11	<i>N</i> -linked glycopeptides at Asn 333 of ZP3 α glycoprotein. Observed <i>m/z</i> and calculated <i>m/z</i> were assigned to the [M+Na] ⁺ ions of the glycopeptides	108
Figure 4.12	<i>N</i> -linked glycopeptides at Asn 203 of ZP3 α glycoprotein. (a) MALDI-MS spectrum obtained on original HPLC fraction. (b) MALDI-MS spectrum obtained after β -galactosidase & <i>N</i> -acetylhexoaminidase treatment	110
Figure 4.13	<i>N</i> -linked glycopeptides (unfucosylated structures) at Asn 203	

	of ZP3a glycoprotein. Observed m/z and calculated m/z were assigned to the $[M+Na]^+$ ions of the glycopeptides	111
Figure 4.14	<i>N</i> -linked glycopeptides (fucosylated structures) at Asn 203 of ZP3a glycoprotein. Observed m/z and calculated m/z were assigned to the $[M+Na]^+$ ions of the glycopeptides	112
Figure 4.15	A micro-batch lectin affinity binding process for subsequently direct MALDI-MS analysis	113
Figure 4.16	Direct MALDI-MS analysis of tryptic digest of ribonuclease B	115
Figure 4.17	MALDI-MS Analysis of Glycopeptides from the Tryptic Digest of ribonuclease B after Micro-batch Lectin (ConA) Affinity Binding	117
Figure 4.18	Direct MALDI-MS analysis of tryptic digest of avidin	118
Figure 4.19	MALDI-MS Analysis of Glycopeptides from the Tryptic Digest of avidin after Micro-batch Lectin (ConA) Affinity Binding	119
Figure 5.1	Structures of monosaccharides commonly found in eukaryotic glycoproteins and glycolipids	126
Figure 5.2	Types of oligosaccharide fragmentation	127
Figure 5.3	Linear positive MALDI-MS spectrum of 3-aminoquinoline (3AQ) derivatized maltohexaose (G6). The reaction mixture was used for direct MALDI-MS analysis	134
Figure 5.4	3-Aminoquinoline glycosylamine derivative formation using maltohexaose as the model compound	135
Figure 5.5	Positive ion MALDI-PSD mass spectrum of 3AQ-G6. The $[M+H]^+$ ion at m/z 1117.4 was selected as the precursor ion. The Oligosaccharide fragment ions in all figures were labeled according to the nomenclature of Domon and Costello	137
Figure 5.6	(a) Linear positive MALDI-MS spectrum of 3AQ-derivatized LS-tetrasaccharide c (LSTc). The small peak at m/z 1021.3 represented the $[M+Na]^+$ ion of underivatized LSTc. (b) Positive ion MALDI-PSD mass spectrum of 3AQ-LSTc with $[M+H]^+$ ion at m/z 1125.4 selected as the precursor ion. For symbols, see Figure 5.5	138

Figure 5.7	Positive ion MALDI-PSD mass spectrum of 3AQ-derivatized (a) LNFP-1 (b) LNFP-2. The $[M+H]^+$ ion of either derivatized isomer had same m/z value, 980.4, and were selected as the precursor ions for both analyses. For symbols, see Figure 5.5	140
Figure 5.8	Positive ion MALDI-PSD mass spectrum of a 3AQ-derivatized triantennary oligosaccharide which was derived from Cyclamen seed xyloglucan. The $[M+H]^+$ ion at m/z 1513.5 was selected as the precursor ion. For symbols, see Figure 5.5	141
Figure 5.9	HPLC separation of 3-aminoquinoline (3AQ)-derivatized maltohexaose (G6). on a Phase Sep Spherisorb S5 amino column. The system was run from 20% to 60% water-acetonitrile in 1 ml/min with a 30 minute linear gradient program	143
Figure 5.10	Positive ion ESI-MS spectrum of HPLC separated 3AQ-G6. Cone voltage (capillary-skimmer voltage): 100 V. The spectra shown represented an average over 1 min recording time. For symbols, see Figure 5.5	145



List of Tables

Table 2.1	Physical properties of MBT and its analogs	26
Table 3.1	MALDI Mass Mapping Analysis of Lysostaphin Digest of Muropeptide Dimer I	74
Table 3.2	MALDI Mass Mapping Analysis of Lysostaphin Digest of Muropeptide Dimer II	77



KEY TO ABBREVIATIONS

MALDI	matrix-assisted laser desorption/ionization
ESI	electrospray ionization
MS	mass spectrometry
TOF	time-of-flight
PSD	post-source decay
MBT	2-mercaptobenzothiazole
CMBT	5-chloro-2-mercaptobenzothiazole
AMBT	6-amino-2-mercaptobenzothiazole
DE	delayed extraction
FAB	fast atom bombardment
KE	kinetic energy
MS/MS	tandem mass spectrometry
HPLC	high performance liquid chromatography
UV	ultraviolet
α CHCA	α -cyano-4-hydroxycinnamic acid
DHB	2,5-dihydroxybenzoic acid
HABA	2-[(4-hydroxyphenyl)azo]benzoic acid
3AQ	3-aminoquinoline
THF	tetrahydrofuran
SDS	sodium dodecylsulfate

SDS-PAGE	sodium dodecylsulfate-polyacrylamide gel electrophoresis
DTT	dithiothreitol
CAD	collisionally activated dissociation
GlcNAc	<i>N</i> -acetylglucosamine
GalNAc	<i>N</i> -acetylgalactosamine
NeuAc	<i>N</i> -acetylneuraminic acid
MurNAc	<i>N</i> -acetylmuramic acid
Asn	asparagine
Ser	serine
Thr	threonine
Man	mannose
Gal	galactose
Glc	glucose
Fuc	fucose
Xyl	xylose
Con A	concanavalin A

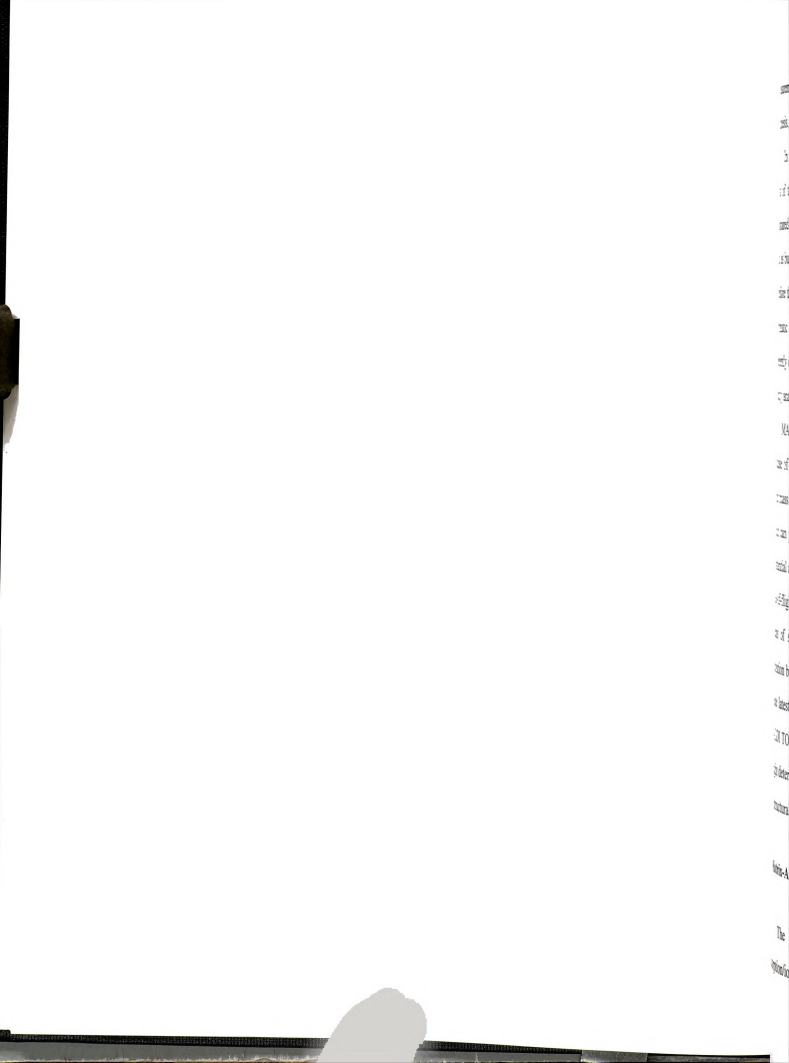
[illegible]

Chapter 1. Introduction

1. Recent Advances of Mass Spectrometry of Biological Molecules

Mass spectrometry is based on producing, differentiating, and detecting ions in the gas phase. The transfer of small molecules into the gas phase has traditionally been accomplished by thermal vaporization; however, thermal vaporization for biopolymers and other nonvolatile or thermally unstable molecules has little use. Fast atom bombardment mass spectrometry (FAB-MS), discovered in the early 80s (1-2), opened the possibility for the analytical chemists to obtain mass spectral signals directly from larger, thermally labile molecules like peptides and oligosaccharides. But FAB-MS analysis is limited to nanomole sensitivity range and usually not amenable to biomolecules with molecular mass higher than 5,000 daltons.

It was the developments of two powerful ionization techniques, matrix-assisted laser desorption ionization (MALDI) (3-4) and electrospray ionization (ESI) (5-6), that dramatically expanded the field of biological mass spectrometry. Both techniques offer picomole to femtomole sensitivity in determination of molecular mass up to a few hundred thousand daltons. Since introduced at almost the same time in 1988, MALDI mass spectrometry (MALDI-MS), and ESI mass spectrometry (ESI-MS) have shown considerable promise in characterization of a wide range of biopolymers including proteins, peptides (7-10), carbohydrates (11-12), and oligonucleotides (13-14). In combination with biochemical techniques, numerous applications in biomedical research have been achieved. The development of MALDI-MS and ESI-MS coincides with an increasing demand for more accurate and sensitive techniques in fields such as biochemistry, molecular biology, genetics and biotechnology (15). Their utility for mass



measurement also facilitates routine characterization in small molecule synthesis, protein synthesis, and compounds obtained directly from biological matrices.

In many respects, MALDI and ESI are complementary. MALDI-MS, used for most of this thesis work, is unique in its capacity to analyze a complex mixture. Compared to ESI-MS or FAB-MS, MALDI-MS is relatively tolerant to contaminants, such as buffers, salts, and denaturants which are common in biological samples. It has therefore the capacity to analyze unseparated, heterogeneous complex mixtures such as enzymatic or chemical digest mixtures of proteins. ESI-MS, on the other hand, is inherently compatible for analyzing mixtures separated on-line by HPLC; yet its ability to directly analyze very heterogeneous samples is limited.

MALDI is typically used in conjunction with time-of-flight (TOF) mass analyzers because of the pulsed nature of laser desorption in MALDI. In addition, the time-of-flight mass analyzer has virtually no upper mass range and is compatible with MALDI, which can produce ions at very high mass-to-charge ratio (m/z). In the last few years substantial advancement also has been made in the two areas of the instrumentation of the time-of-flight mass analyzer: (a) enhanced capability of structural determination by means of so called post-source decay (PSD) analysis (16-17); (b) increased mass resolution by incorporating delayed extraction (DE) technique in the ion source (18-21). These latest developments, as described in the remainder of this chapter, have made MALDI TOF-MS an extremely useful technique not only for more accurate molecular weight determination (up to resolution in excess of 10,000 at mass up to 5 kDa), but also for structural elucidation of biomolecules.

2. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS)

The first attempt to generate ions of organic molecules by direct laser desorption/ionization dated back to early 1970s (22). However, the size of the analytes

[illegible]

which can be desorbed and ionized was limited to about 1,000 daltons. In early applications (23-24) of laser desorption to volatilize large biopolymers, the neat analyte (no matrix) was irradiated directly with intense pulses of laser light for short durations. However, mass spectra for large molecular weight compounds were frequently dominated by fragment ions with few if any intact ionized molecules. This technique (direct laser desorption) had limited applications to biological samples and has been quickly supplanted by MALD-MS..

2.1 Operating Principles

The key to the MALDI process is the incorporation of the analyte into a matrix that absorbs radiation from the ultraviolet (or infrared) laser. The matrix compound is a small organic molecule and selected for its properties of laser energy absorption, solubility characteristics, and other criteria as described in detail in Chapter 2. The sample is typically mixed into a solution of the matrix which is in large excess with a matrix:analyte molar ratio in the order of 1000:1. Once the solvent is removed, the sample and the matrix cocrystallize after they have been deposited on a metal surface. The physicochemical events leading to the transfer of biopolymers to the gas phase and their ionization in MALDI have not yet been fully elucidated (25). The matrix is believed to serve to minimize sample fragmentation from the laser beam by absorbing the incident laser energy, resulting in the sample and matrix molecules being ejected into the gas phase. One model (15, 26) for the mechanism assumes that the uppermost layers of matrix are induced to undergo a phase transition from the solid to the gas phase. The subsequent expansion of these matrix molecules into the vacuum drags the matrix-isolated biopolymer molecules into the gas phase. During the transfer from the solid phase to the gas phase, the biopolymer undergoes ionization through proton transfer or cation attachment with the matrix by reaction processes that remain to be explained

[illegible]

(27-30). Once ions are formed in the gas phase, they can be accelerated in an electric field into to a mass analyzer.

The operating principle and the configuration of the MALDI instrument used for the thesis work are illustrated in Figure 1.1. This instrument, a PerSeptive Biosystems Voyager-Elite time-of-flight (TOF) mass spectrometer, is equipped with a nitrogen laser emitting at 337 nm, and a single-stage reflector (or reflectron). Dual microchannel plates were employed for ion detection. The sample plate holder used in Voyager Elite system can accommodate 100 sample wells and permits a high sample throughput for analysis. The laser radiation, controlled by an attenuator, is focused on the sample plate well through appropriate optics. A fraction of light split from the laser pulse initiates the timing circuitry for measuring the time of flight for the ions produced by each laser pulse. In continuous extraction mode, a static electric field is imposed upon ions generated from the sample by application of a two-stage high voltage (20-30 kV) to the sample plate with regard to a closely spaced accelerating electrode, which corresponds to the grounded aperture in Figure 1.1. As all of the ions are given the same energy in the ion acceleration region, The lighter ions formed during each pulse travel faster in a field-free region, the flight tube as shown in Figure 1, and reach the detector earlier. The time required for ions to traverse the flight tube (t.o.f.) is dependent on their masses and described by the relationship: $t.o.f. = L/v = L(m/2zeV)^{1/2}$, where L is the length of flight tube, v is the ion velocity, m is the mass of the ion, and V is the acceleration potential. Thus the ions are separated into a series of spatially discrete individual ion packets, each traveling with a velocity characteristic of its m/z ratio. A detector positioned at the end of the field-free region produces a signal as each ion packet strikes it. The transient (duration of 300 microseconds typically) mass

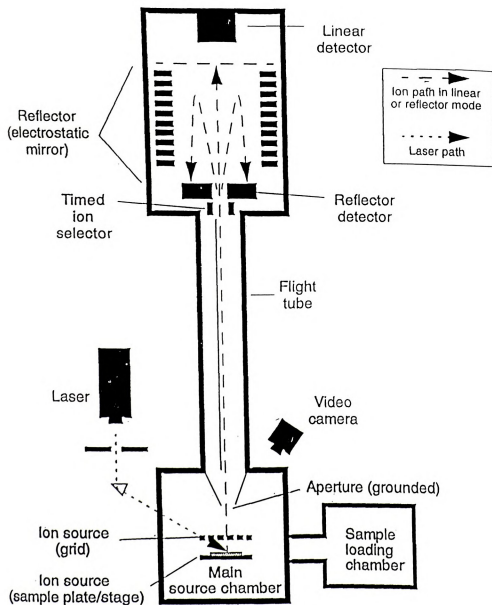


Figure 1.1: Voyager-Elite MALDI time-of-flight mass spectrometer.

simple
Prop
Marin
at a
The
a misc
amacet
Give a f
a 2 µl o
either
the dry

spectrum is recorded by a transient recorder at the detector. Commonly 10 to 250 of these transient mass spectra are summed up to produce a usable mass spectrum. To correct the initial kinetic energy spread of ions generated by MALDI, an electrostatic mirror (reflector) can be used to improve mass resolution for some applications.

MALDI is typically considered as a soft ionization technique that yields little fragmentation. The most attractive feature of MALDI spectrum is its simplicity. Unlike ESI spectra which are dominated by multiply charged species, the predominant analyte signal in MALDI corresponds to the singly charged species, though the doubly-charged and triply-charged molecules are also observed to varying degrees. Also commonly observed in MALDI mass spectra are ions for gas phase aggregates such as dimers, trimers, etc. Other often observed characteristics of MALDI-MS spectra are adducts arising from analyte-matrix interactions or analyte-alkali metal ion interactions depending on the analyte, choice of matrix, and laser power.

2.2 Sample Preparation

Proper sample preparation is critical for successful analysis by MALDI-MS (31-32). Matrix solutions are typically prepared in water, water-acetonitrile or water-ethanol mixtures at a concentration of 5-100 mg/ml depending on the solubility properties of the matrix. The analyte is prepared at a concentration of about 0.01-0.1 mg/ml in a solvent that is miscible with the matrix solution (for peptides and proteins, aqueous 0.1% trifluoroacetic acid (TFA) is frequently used). The matrix and analyte solutions are mixed to achieve a final matrix:analyte molar ratio in the range of 100:1 to 10,000:1. An aliquot of 1 to 2 μ l of the mixture is applied to a MALDI-MS sample plate well, and allowed to dry by either ambient evaporation, heating with a stream of warm air, or under vacuum. During the drying process, the matrix codeposits from solution with the analytes.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

To date, the nature of the matrix-analyte interaction has not yet been defined, but it may involve analyte molecules that have become embedded into the matrix crystal lattice or on the surface of rapidly forming matrix crystals. Studies based on optical microscopy have shown that matrix crystal formation and size vary, depending on the matrix and solvent. In general, it has been observed that a decrease in the "quality" of matrix crystal morphology (where the matrix-analyte deposit appears as a film or non-crystalline glassy surface) is associated with a lower quality MALDI mass spectrum (33). These conditions can result from the presence of high levels of contaminating salts, buffers, or surfactants. Some means to desalt the sample may be required for successful analysis by MALDI (34). Alternatively, dilution of a contaminant which can interfere with analyte signal acquisition, may also lead to improved spectra. In general, sample preparation by different matrices, matrix additives, solvents, molar ratio between the analyte and the matrix affects the sensitivity and resolution of MALDI. Optimal results require parallel analyses under different conditions.

Other than a metal surface, inert polymeric supports such as transfer membranes have also been used for picomole-level sample handling in protein mass spectrometry (35-37). Proteins are often immobilized by adsorption onto synthetic supports at some stage as is performed routinely in automated Edman sequence analysis. There is also considerable interest now in the direct analysis by MALDI-MS of biological samples that have been adsorbed onto a transfer membrane because of implications in coupling gel electrophoresis with MALDI-MS. Nylon-66 and positive charge-modified nylon (ZETABIND) membranes were reported to be suitable for analyzing picomole levels of a protein that had been immobilized by adsorption onto the nylon membrane, washed to remove MALDI contaminants, and digested enzymatically and/or chemically prior to adding matrix (35). Another approach, affinity mass spectrometry employs a modified target surface. In one example, a target composed of porous agarose immobilized with single-stranded DNA has been developed as an affinity-capture medium in the assay of the 80-kDa glycoprotein lactoferrin from biological samples (38).

These findings suggest that the use of a single, standardized, and validated instrument to assess the impact of the intervention on the target population is a critical component of the evaluation process. The use of a single instrument allows for the collection of data that can be used to compare the results of the intervention to the results of the control group. The use of a validated instrument ensures that the data collected are reliable and valid. The use of a standardized instrument ensures that the data collected are comparable to the results of other studies. The use of a validated and standardized instrument is a critical component of the evaluation process.

2.3 Post-source Decay (PSD)

MALDI TOF-MS has been initially recognized as "molecular weight machine" by the biochemists because of its superior capability in molecular mass determination of a wide range of biopolymers such as proteins, peptides, and carbohydrates. However, it lacked the capacity for structural analysis because relatively little or no fragmentation was observed in linear MALDI TOF-MS instrument. Metastable fragments formed in the flight tube could not be detected. This deficiency has been overcome since the introduction of the so called post-source decay (PSD) technique (16-17) which can separate and resolve metastable fragments on a reflector-equipped TOF instrument in 1993.

In this technique, the excess energy deposited into the intact ionized molecules was found to result in the decomposition of a portion of these species in the flight tube after they leave the ion source. If the ions dissociate within the source, the fragment ions and the intact ionized molecules would have obtained the same energy during acceleration, and because they have different masses, would travel with different velocities and have different time of flight values. This type of fragmentation, named as "prompt fragmentation" (39-40), does occur in MALDI TOF-MS, but was only observed for some very special analytes such as some carbohydrates and mucopeptides derived from peptidoglycans as described in Chapter 2 and 3. However, for those ions that undergo PSD in the flight tube, the resulting fragment ions (or product ions) continue toward the detector with the same velocity as the intact ionized molecules (precursors), and thus are detected as a single peak in a conventional linear TOF-MS. But these product ions have kinetic energies (KE) that are different from each other (if they have different masses) and less than that of the precursor ion, and it is this characteristic that enables them to be separated by mass in the reflector. Consider a molecule M, which was

21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 54

ionized in the MALDI source to form a protonated molecule, MH^+ , at m/z 1,000 and accelerated by a electric field of 20 kV. MH^+ dissociates into a fragment ion M_1^+ at m/z 500 and a neutral M_2 during its flight from the source to the detector.

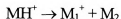


Figure 1.2 shows the projected flights of these ions within the reflector portion of the mass spectrometer. When M_1^+ ions were formed in flight, they retain the same velocity as MH^+ ions which remained intact, but the kinetic energy of M_1^+ is half that of MH^+ . In the top panel, when no potential is applied to the reflector, both ions MH^+ and M_1^+ as well as neutral M_2 arrive at the linear detector at approximately the same time, but with some variability, resulting in an increase of the peak width representing the intact MH^+ . However, the ions M_1^+ can be distinguished very well from MH^+ when a retarding potential, 12 kV, is applied to the reflector as shown in the low panel of Figure 1.2. The ions MH^+ entering the reflector with a KE of 20 keV will traverse the mirror with a final kinetic energy of 8 keV and hit the linear detector. The neutrals, M_2 , would be unaffected and hit the linear detector as before. However, the fragment ions, M_1^+ , entering the reflector with a KE of 10 KeV, will be slowed down in the retarding potential first and then turned around by the reflector. Because of the dispersion process by the inherent radial velocity, M_1^+ ions will reach the reflectron detector and be recorded. By lowering the reflector potential in several intervals, the entire mass range of PSD fragment ions for a particular precursor can be detected in the same manner.

Like other tandem mass spectrometric (MS/MS) techniques, PSD can provide full or partial sequence information of molecules less than 2,500 daltons. Precursor ion selection for PSD was typically performed by an electrostatically switched ion gate, marked as timed ion selector Figure 1.1. Since its commercialization about three years

100

100

Figure 1

Microle

600 no

2000

potential

factor

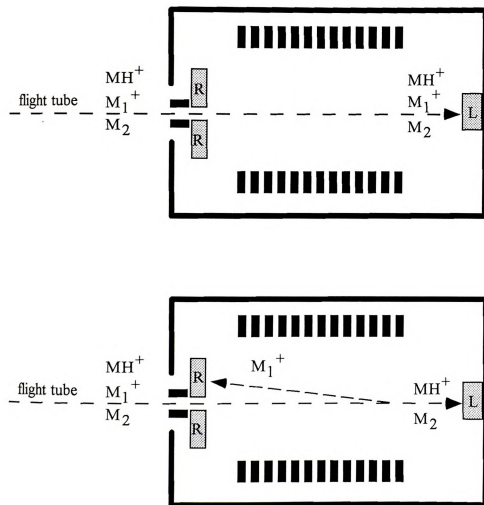
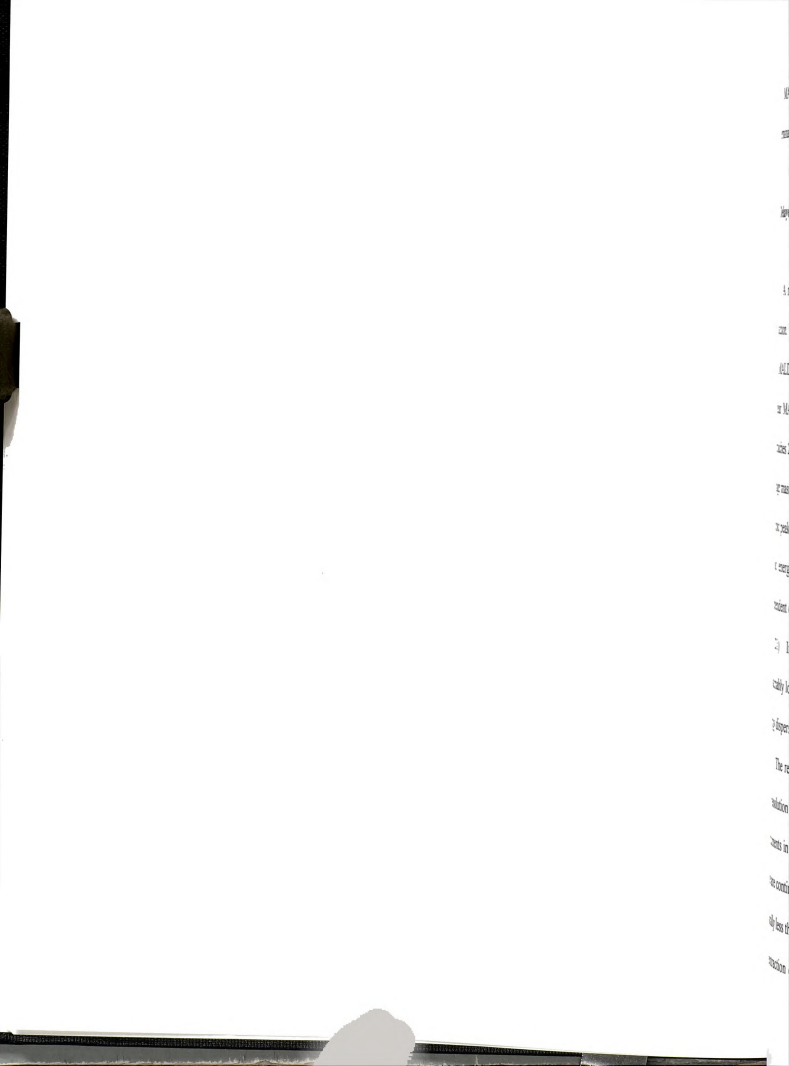


Figure 1.2: Differentiation of PSD fragment ion, M_1^+ , from the intact ionized molecule MH^+ by using the reflector. M_2 is the neutral dissociated from MH^+ . With no potential applied to the reflector (top panel), M_1^+ and M_2 maintain the initial velocity of MH^+ and reach the linear detector, L, simultaneously. With potential applied to the reflector (low panel), M_1^+ is reflected to the reflectron detector, R, and recorded while M_1^+ and M_2 still hit the linear detector.



ago, MALDI-PSD has quickly evolved into a powerful technique for sequence determination of biopolymers like peptides at the low picomole scale (41).

2.4 Delayed Extraction (DE)

A major limitation of TOF mass analyzers has often been relatively poor mass resolution. The use of a reflector has overcome this limitation for some applications, but with MALDI in particular, resolution is often inadequate. Typically the resolving power of a linear MALDI TOF-MS is well below 1,000. With a reflector, the resolving power approaches 2,000 or sometimes even greater. Because of the poor resolving power and also the high mass of the ions analyzed by MALDI TOF-MS, it is not possible to discern individual isotopic peaks. The main problem is that the ions produced by MALDI exhibit a rather broad energy distribution. The initial velocity of desorbed analyte ions is nearly independent of mass and the initial kinetic energy is proportional to the mass of the analyte (18, 21). In addition, when desorption occurs in a strong electric field, energy is presumably lost by collisions within the neutral plume, resulting in further mass-dependent energy dispersion (21).

The recently introduced technique of delayed extraction (DE) dramatically improves the resolution achievable in MALDI TOF-MS. In contrast to conventional MALDI instruments in which the ions generated by the laser beam near the surface of the sample plate are continuously extracted by a dc potential, in delayed extraction, a short time delay (typically less than 300 nanoseconds) is inserted between the laser desorption/ionization and ion extraction events. The basic idea is to allow those ions having the greater initial kinetic



energy to travel farther into the ion source than ions having a lesser kinetic energy, so that when the extraction field is applied, the ion having the lesser initial kinetic energy will traverse a greater distance through the electric field, thereby picking up more kinetic energy in a compensatory effort to give all ions of the same mass the same kinetic energy. Three important parameters need to be adjusted to optimize the performance. These include the field magnitude and direction during and immediately following ion production, the magnitude of the extraction field, and the time delay between the laser pulse and application of the extraction field. In principle, delayed extraction, causes ions of the same mass to arrive at the detector at nearly the same time giving rise to a sharp peak or greater resolution. Some improvement in resolution with delayed extraction may be due to dissipation of the very high local pressure (tens of atmospheres) in the laser plume following (for picoseconds) the laser pulse (18, 21).

The benefit of delayed extraction in MALDI-TOF-MS is illustrated in Figure 1.3 which compares the MALDI mass spectrum of human adrenocorticotrophic hormone fragment (ACTH) 18-39 obtained by conventional use of a linear TOF-MS (top panel) with those obtained by delayed extraction into a linear (middle panel) and a reflector (low panel) TOF-MS. Delayed extraction improves resolution, mass accuracy, and the quality of MALDI mass spectra by suppressing matrix background, reducing chemical noise, and minimizing the effect of laser intensity on performance. The potential of delayed extraction MALDI has been demonstrated for proteins and peptides as well as for oligonucleotides (42). The delayed extraction technique was added to our Voyager Elite MALDI-MS instrument in late 1996 and was used for the experiments described in Chapter 4 and 5.

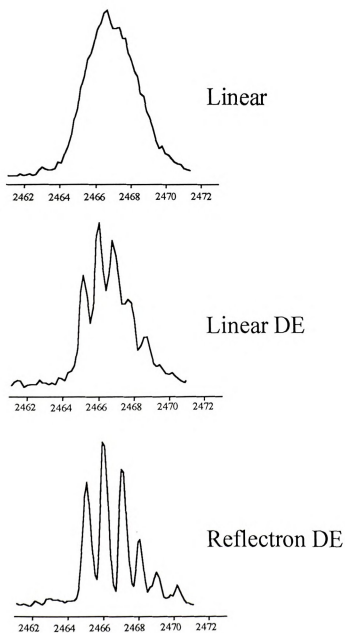
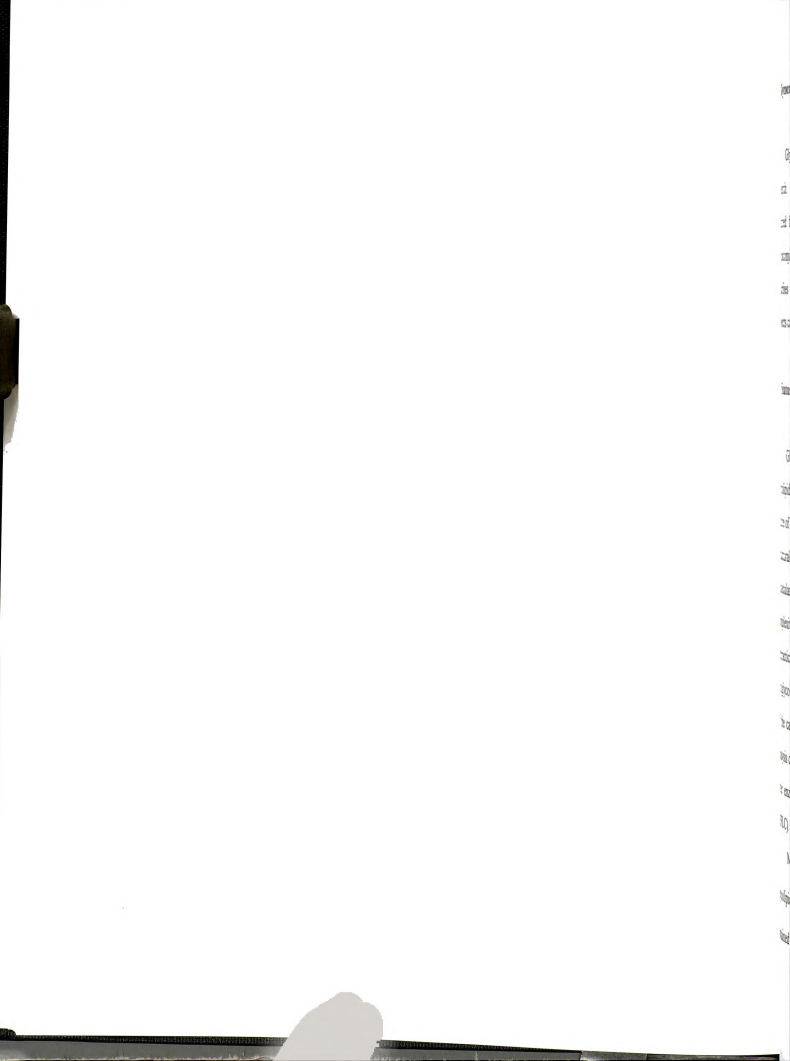


Figure 1.3: Comparison of MALDI mass spectra of ACTH (18-39) obtained in linear mode without DE (Linear), linear mode with DE (Linear DE), and reflectron mode with DE (reflectron DE) showing the effect of DE on mass resolution of TOF-MS.



3. Glycoconjugate and Protein Analysis by MALDI-MS

Glycoconjugates and proteins are two major categories of molecules in biomedical research. Methods for the characterization of carbohydrates and proteins are both required for the analysis of glycoproteins. Like other applications of MALDI-MS, glycoconjugate and protein analysis is still a rapidly developing field. This section describes the latest developments in this field as well as major goals of the research projects carried out in this thesis work.

3.1 Status of Glycoconjugate Analysis

Glycoconjugates include glycoproteins, glycopeptides, peptidoglycans, and glycolipids. Glycoconjugate analysis is usually complicated because of the heterogeneous nature of its carbohydrate moiety. The development of MALDI-MS has greatly facilitated structural analysis of glycoproteins. To date, the most reliable method of determining the molecular mass of a heavily glycosylated protein is MALDI TOF-MS (43-44) because the complexity derived from the multiple charge states present in ESI-MS generally renders it impractical for this purpose. New approaches incorporating endoglycosidase and exoglycosidase have been used to identify the site of glycosylation and obtain the sequence of the carbohydrate moiety (45-46). It was also demonstrated (12) that MALDI-PSD analysis could provide structural information for glycopeptides derived from glycoproteins after enzymatic digestion and separation by high-performance liquid chromatography (HPLC).

MALDI TOF-MS has been shown to be a viable technique for analysis of glycolipids including gangliosides (47-48). Although satisfactory spectra have been obtained from underivatized molecules, permethylation has been shown to increase the ion



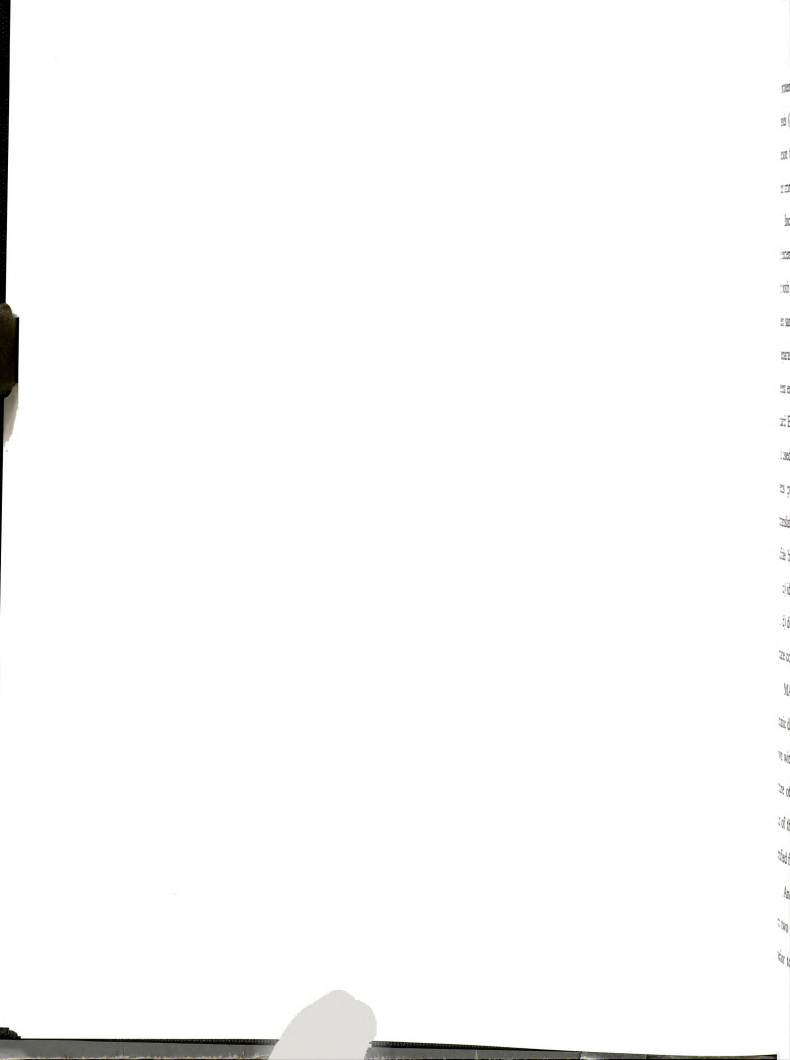
yield considerably. In general, negative ion analysis has been preferred to work in the positive ion mode, an observation that also applies to the FAB-MS.

Biemann et al. (49) recently found that the molecular weight of highly sulfated glycoconjugates could be measured by MALDI-MS at the picomole level by mixing them with a basic peptide of known mass. These highly acidic compounds, derived from heparin and other polysaccharides, cannot be ionized in any other way with such sensitivity. Such methodology using MALDI-MS leads to a way of sequencing heparin, the detailed structure of which is not known, although the compound has been used clinically as an anticoagulant for many decades.

The complete characterization of glycoconjugates usually requires structural analysis of their respective oligosaccharides. Analysis of oligosaccharide structure has been dominated by nuclear magnetic resonance (NMR) in the past. However, NMR lacks the sensitivity to address many biological problems. MALDI-MS not only has become a sensitive and accurate means in determining the molecular weight of oligosaccharides purified from various biological sources (50), but also provides the possibility of sequence and branching pattern determination of oligosaccharides at the picomole level (51). In addition, MALDI-MS has been used for the direct analysis of carbohydrate mixtures such as oligosaccharides from human milk (11).

3.1 Status of Protein Analysis

Sodium dodecyl sulfate-polycrylamide gel electrophoresis (SDS-PAGE) has been used by the biochemists as a universal technique in protein molecular mass determination for a long period. Over the past few years, MALDI-MS has quickly gained its popularity



for protein molecular mass determination with much better accuracy relative to other methods (e.g., SDS-PAGE). Because of the unlimited mass range of this technique, detection for very large proteins such as 500-kDa gramicidin S-synthase and 939-kDa human immunoglobulin IgM (52) was achieved.

Incorporation with other biochemical tools like peptide mapping, MALDI-MS has been extensively used for the determination of the primary structure of proteins derived from both natural and recombinant sources. Peptide mapping is a technique whereby a protein sample is digested either enzymatically or chemically, and the resulting peptides are separated and analyzed. Overlapping sets of peptide fragments are generated using different enzymes or chemical agents and the peptides are separated and sequenced using standard Edman degradation reactions. Nowadays, peptide mapping based on MALDI-MS is used in the following applications: (a) confirmation of a protein sequence, especially of proteins produced through recombinant DNA techniques (53); (b) characterization of posttranslational modifications including covalently modified protein N- and C- termini (54), disulfide bond pairing (55-56), phosphorylation (57), glycosylation (58), and lipidation (59); (c) identification of variants in homologous proteins from different biological sources (60); (d) determination of the antigenic site or epitope, the specific region of the protein in intimate contact with the antibody (61).

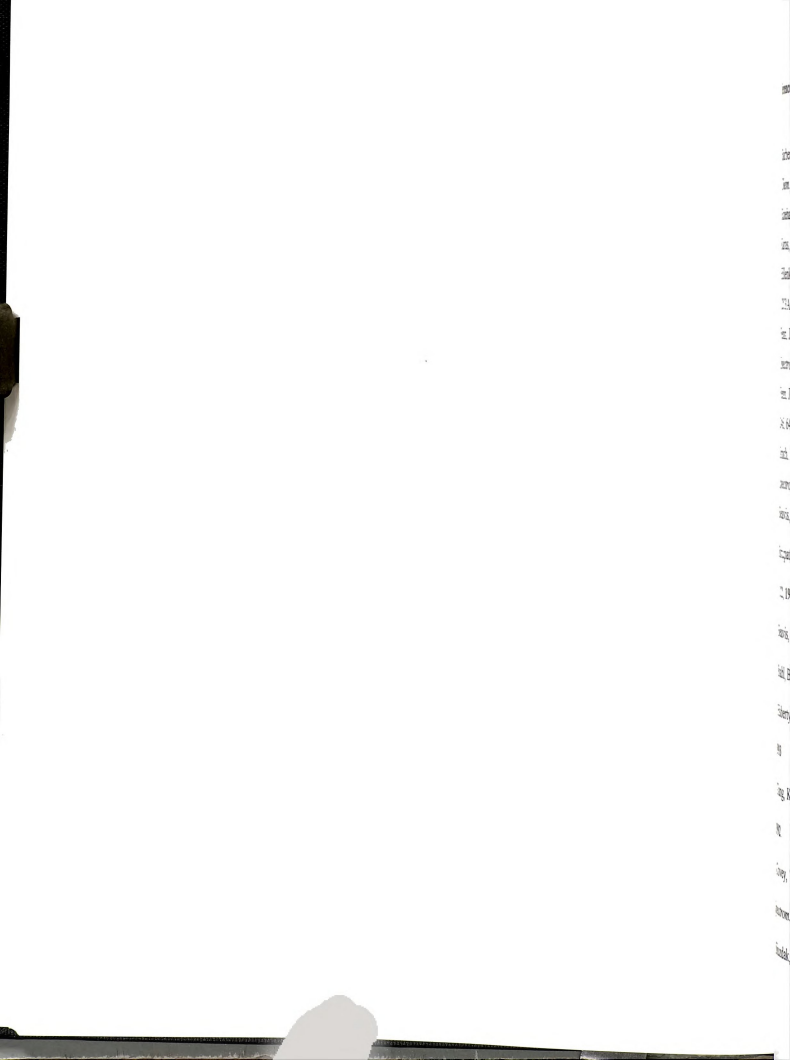
MALDI-MS provides routine means for direct analysis of a tryptic digest (62). A dramatic demonstration of the capability of MALDI in analyzing heterogeneous samples is shown with "protein ladder sequencing" which involves the simultaneous analysis of a mixture of peptide/proteins that have undergone a stepwise Edman degradation (63). Each of the fragments differs from the next by one amino acid residue, which can be identified from the mass difference between successive peaks.

Another highly promising area is the unambiguous identification of protein spots from two dimensional gel electrophoresis by MALDI-MS (64-66). This method is superior to Edman microsequencing for protein spot identification as the latter often

suffers from problems such as insufficient protein quantities in 2D gel spots, widespread occurrence of N-terminally blocked proteins, poor protein recoveries from gels, and other unknown factors.

3.3 Major Goals of this Research

The three major goals of the research work described in this thesis are: (a) the investigation of a group of compounds, mercaptobenzothiazoles, as a new class of matrices for MALDI-MS and their application in structural characterization of peptidoglycan muropeptides; (b) the development of a new type of derivative for oligosaccharide sequence determination by mass spectrometry; (c) the study of lectin affinity binding in structural analysis of glycoproteins.



References:

1. Barber, M., Bordoli, R. S., Elliot, G. J., Sedgwick, R. D., and Tyler, A. N., *Anal. Chem.*, 54, 645A-657A, **1982**.
2. Rinehart, K. L., *Science*, 218, 254-260, **1982**.
3. Karas, M., and Hillenkamp, F., *Anal. Chem.*, 60, 2299-2301, **1988**.
4. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T., *Anal. Chem.*, 63, 1193A-1203A, **1991**.
5. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M., *Mass Spectrom. Rev.*, 9, 37-70, **1990**.
6. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M., *Science*, 246, 64-71, **1989**.
7. Smith, R. D., Loo, J. A., Ogorzalek Loo, R. R., Busman, M., and Udseth, H. R., *Mass Spectrom. Rev.*, 10, 359-451, **1991**.
8. Beavis, R. C.; Chait, B. T. *Anal. Chem.*, 62, 1836-1840, **1990**.
9. Strupat, K.; Karas, M.; Hillenkamp, F., *Int. J. Mass Spectrom. Ion Processes*, 111, 89-102, **1991**.
10. Beavis, R. C.; Chait, B. T., *Proc. Natl. Acad. Sci. U.S.A.*, 87, 6873-6878, **1990**.
11. Stahl, B.; Steup, M.; Karas, M.; Hillenkamp, F., *Anal. Chem.*, 63, 1463-1466, **1991**.
12. Huberty, M. C.; Vath, J. E.; Yu, W.; Martin, S. A., *Anal. Chem.*, 65, 2791-2796, **1993**.
13. Tang, K.; Allman, S. L.; Chen, C. H., *Rapid Commun. Mass Spectrom.*, 6, 365-368, **1992**.
14. Covey, T. R., Bonner, R. F., Shushan, B. I., and Henion, J., *Rapid Commun. Mass Spectrom.*, 2, 249-256, **1988**.
15. Siuzdak, G., *Proc. Natl. Acad. Sci. U.S.A.*, 91, 11290-11297, **1994**.

[illegible]

16. Kaufmann, R.; Kirsch, D.; and Spengler, B., *Int. J. Mass Spectrom. Ion Proc.* 131, 355-385, 1994.
17. Kaufmann, R.; Spengler, B.; and Lutzenkirchen., *Rapid Commun. Mass Spectrom.*, 7, 902-910, 1993.
18. Brown, R. S., and Lennon, J. J., *Anal. Chem.*, 67, 1998-2003, 1995.
19. Colby, S. M., King, T. B., and Reilly, J. P., *Rapid Commun. Mass Spectrom.*, 8, 865-868, 1994.
20. Whittall, R. M., and Li, L., *Anal. Chem.*, 67, 1950-1954, 1995.
21. Vestal, M. L., Juhasz, P., and Martin, S. A., *Rapid Commun. Mass Spectrom.*, 9, 1044-1050, 1995.
22. Posthumus, M. A., Kistemaker, P. G., Meuzelaar, H. L. C., and Brauw, M. C., *Anal. Chem.*, 50, 985-991, 1978.
23. Cotter, R., *Anal. Chem. Acta.*, 195, 45-59, 1987.
24. Nuwaysir, L., and Wilkins, C., *Anal. Chem.*, 60, 279-282, 1988.
25. Chait, B. T.; and Kent, S. B. H., *Science*, 257, 1885-1894, 1992.
26. Busch, K. L., *J. Mass Spectrom.*, 30, 233-240, 1995.
27. Karas, M., Bahr, U., Ingendoh, A., Nordhoff, E., Stahl, B., and Hillenkamp, F., *Anal. Chem. Acta.*, 241, 175-185, 1990.
28. Ehring, H., Karas, M., and Hillenkamp, F., *Org. Mass Spectrom.*, 27, 472-480, 1992.
29. Gimon, M. E., Preston, L. M., Solouki, T., White, M. A., and Russell, D. H., *Org. Mass Spectrom.*, 27, 827-830, 1992.
30. Solovki, T. and Russell, D., *Appl. Spectros.*, 47, 211-217, 1993.
31. Cohen, S. L. and Chait, B. T., *Anal. Chem.*, 68, 31-37, 1996.
32. Vorm, O., Roepstorff, P., and Mann, M., *Anal. Chem.*, 66, 3281-3287, 1994.
33. Doktycz, S. J., Savickas, P. J., and Krueger, D. A., *Rapid Commun. Mass Spectrom.*, 5, 145-148, 1991.

196
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
2639
2640
2641
2642
2643
2644
2645
2646
2647
264

34. Zhang, H., Andren, P. E., and Caprioli, R. M., *J. Mass Spectrom.*, 30, 1768-1771, 1995.
35. Zaluzec, E. J., Gage, D. A., Allison, J., and Watson, J. T., *J. Am. Soc. Mass Spectrom.*, 5, 230-237, 1994.
36. Blackledge, J. A. and Alexander, A. J., *Anal. Chem.*, 67, 843-848, 1995.
37. Mock, K. K., Sutton, C. W., and Cottrell, J. S., *Rapid Commun. Mass Spectrom.*, 6, 233-238, 1992.
38. Hutchens, T. W. and Yip, T. T., *Rapid Commun. Mass Spectrom.*, 7, 576-580, 1993.
39. Xu N., Huang Z. H., Watson J. T., and Gage D. A., *J. Am. Soc. Mass Spectrom.* 8, 116-124, 1997.
40. Scott, P. D. and Viswanatham, K., *Anal. Chem.* 66, 3727-3733, 1994.
41. Rouse, J. C., Yu, W., and Martin, S., *J. Am. Soc. Mass Spectrom.*, 6, 822-835, 1995.
42. Juhasz, P., Roskey, M. T., Smirnov, I. P., Haff, L. A., Vestal, M. L., and Martin, S. A., *Anal. Chem.*, 68, 941-946, 1996.
43. Wada, Y., Gu, J., Okamoto, N., and Inui, K., *Biol. Mass Spectrom.*, 23, 108-110, 1994.
44. Lapolla, A., Baldo, L., Aronica, R., Fedele, D., and Traldi, P., *Biol. Mass Spectrom.*, 23, 241-245, 1994.
45. Sutton, C. W., O'Neil, J. A., and Cottrell, J. S., *Anal. Biochem.*, 218, 34-36, 1994.
46. Treuheit, M. J., Costello, C. E., and Kirley, T. L., *J. Biol. Chem.*, 268, 13914-13918, 1993.
47. Juhasz, P. and Costello, C. E., *J. Am. Soc. Mass Spectrom.*, 3, 785-796, 1992.
48. Harvey, D., *J. Mass Spectrom.*, 30, 1311-1324, 1995.
49. Juhasz, P. and Biemann, K., *Proc. Natl. Acad. Sci. U.S.A.*, 91, 4333-4337, 1994.
50. Harvey, D., Rudd, P. M., Bateman, R. H., Bordoli, R. S., Howes, K., Hoyes, J. B., and Vickers, R. G., *Org. Mass Spectrom.*, 29, 753-758, 1994.
51. Spengler, B., Kirsch, D., Kaufmann, R. J., and Lemoine, J., *J. Mass Spectrom.*, 30, 782-787, 1995.

10

2

10

1

256

1

310

17

Clayton

58

29.9

During

Floor

Задача

CONCEPT

19, 1

20, Y

知

日

versc

time

3,2

See,

24

52. Nelson, R. W., Dogruel, D., and Williams, P., *Rapid Commun. Mass Spectrom.*, 8, 627-631, 1994.
53. Burlingame, A. L., Carr, S. A., Eds., *Mass Spectrometry in the biological sciences*, Humana Press, Totowa, NJ, 1996.
54. Specht, B., Oudenampsen-Kruger, E., Ingendoh, A., Hillenkamp, F., Leaius, A. G., and Spener, F., *J. Biotechnol.*, 33, 259-269, 1994.
55. Robertson, J. G., Adams, G. W., Medzihradszky, K. F., Burlingame, A. L., and Villafranca, J. J., *Biochemistry*, 33, 11563-11575, 1994.
56. Chang, J. Y., Schindler, P., and Chatrenet, B., *J. Biol. Chem.*, 270, 11992-11997, 1995.
57. Liao, P-C., Leykam, J., Andrew, P. C., Gage, D. A., and Allison, J., *Anal. Biochem.*, 219, 9-20, 1994.
58. Burlingame, A. L., *Curr. Opin. Biotechnol.*, 7, 4-10, 1996.
59. Wilcox, M. D., Schey, K. L., Busman, M., and Hildebrandt, J. D., *Biochem. Biophys. Res. Commun.*, 212, 367-374, 1995.
60. Kerner, J., Zaluzec, E. J., Gage, D. A., and Bieber, L. L., *J. Biol. Chem.*, 269, 8209-8219, 1993.
61. Zhao, Y., and Chait, B. T., *Anal. Chem.*, 66, 3723-3726, 1994.
62. Billeci, T. M., and Stults, J. T., *Anal. Chem.*, 65, 1709-1716, 1993.
63. Chait, B. T., Wang, R., Beavis, R. C., and Kent, S. B., *Science*, 262, 89-92, 1993.
64. Patterson, S. D., and Aebersold, R., *Electrophoresis*, 16, 1791-1814, 1995.
65. Erdjument, B. H., Lui, M., Sabatini, D. M., Snyder, D. H., and Tempst, P., *Protein Sci.*, 3, 2435-2446, 1994.
66. Clauser, K. R., Hall, S. C., Smith, D. M., Webb, J. W., Andrews, L. E., Tran, H. M., Epstein, L. B., and Burlingame, A. L., *Proc. Natl. Acad. Sci. U.S.A.*, 92, 5072-5076, 1995.

Chapter 2. Mercaptobenzothiazoles: A New Class of Matrices for Laser Desorption Ionization Mass Spectrometry

1. Introduction

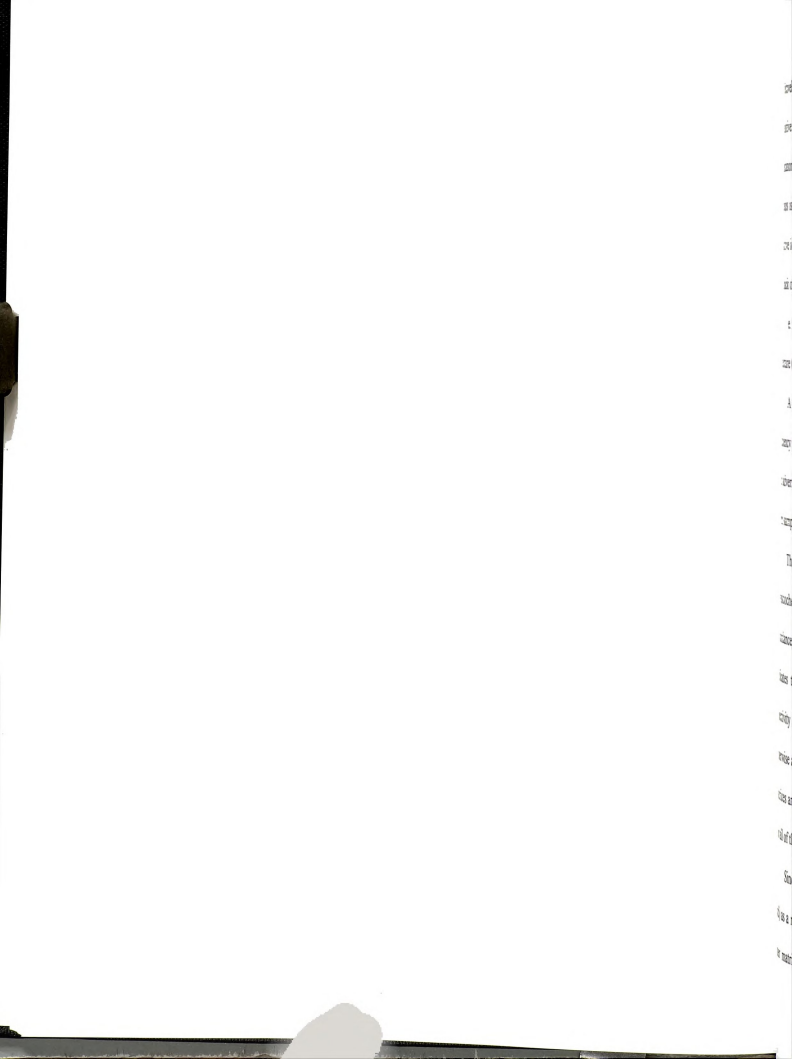
An essential feature of the MALDI process is the use of a low molecular weight organic compound as a matrix to absorb the laser energy. The matrix probably plays multiple roles in the ion formation process (1-4). The step of absorbing the laser radiation leads to the breakup of a microvolume of the solid phase. Reactive precursors are formed and released from the sample surface as protonated and deprotonated matrix molecules, radical ions, and electronically and/or vibrationally excited neutral matrix molecules. Efficient matrices generate a high yield of ions from the analyte relative to the abundance of the matrix ions. To foster this highly complex process of ionization, we assume that a good matrix substance should have the following properties:

a. A reasonably high molar absorptivity coefficient at the wavelength used ($\epsilon > 10^4$ L cm⁻¹ mol⁻¹ for ultraviolet (UV)-MALDI). A higher absorptivity coefficient at the given wavelength does not necessarily make a matrix work better.

b. Miscibility with the analyte in the solid phase. The matrix should be soluble in the solvent of the analyte. For application to proteins, the usual solvent is water or a mixture of water and an alcohol or acetonitrile.

c. Good vacuum stability.

d. "Proper chemical composition". Most matrices reported to date that promote protonation of the analyte contain OH and/or NH groups. It is plausible that protons are



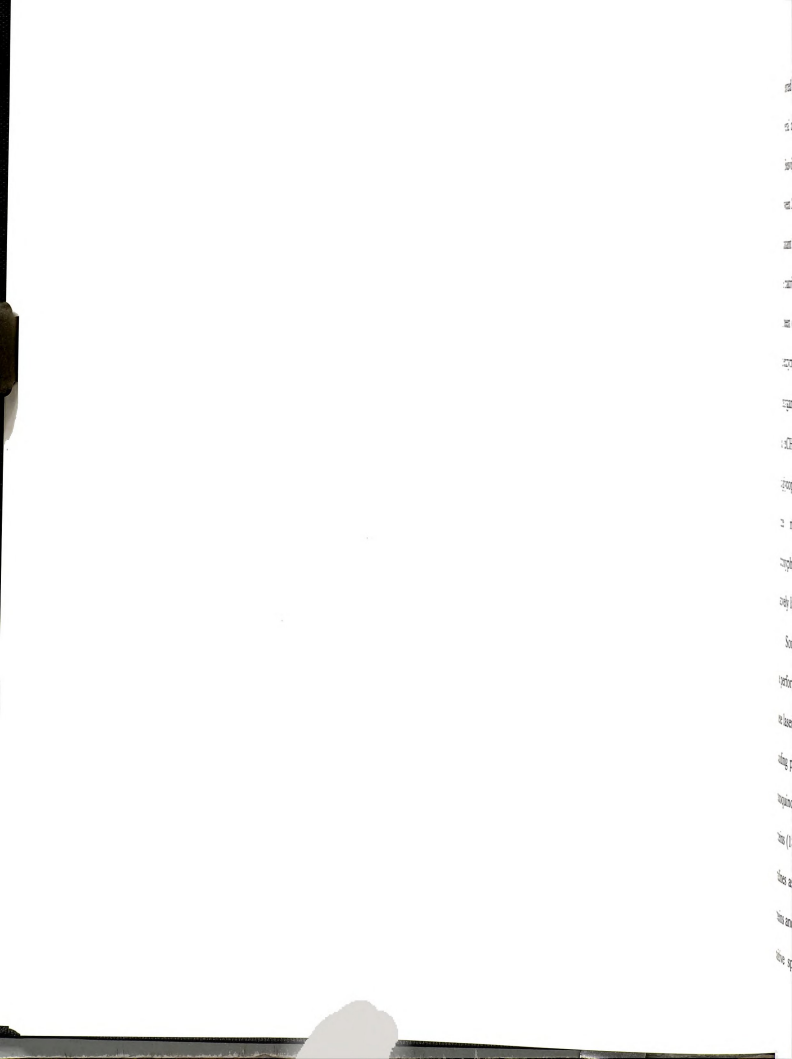
effectively transferred from these groups during ionization. However, many good "positive ion mode" matrices are also applicable in the negative ion mode. If the major ionization pathway corresponds to cation or anion attachment, the presence of OH or NH groups is not necessary. Most matrix development studies have been restricted to the positive ion mode; no generalization can yet be made for the negative ion mode because of the lack of sufficient experimental data.

e. Other physical properties, such as the heat of sublimation (5) and the lattice structure (6), may affect the efficiency of a matrix.

A number of additional features are desirable, but do not necessarily influence the efficiency of the ionization process. For instance, matrix-analyte photoadduct formation can adversely affect mass assignment accuracy when the adduct peaks are not resolved from sample peaks (7-9).

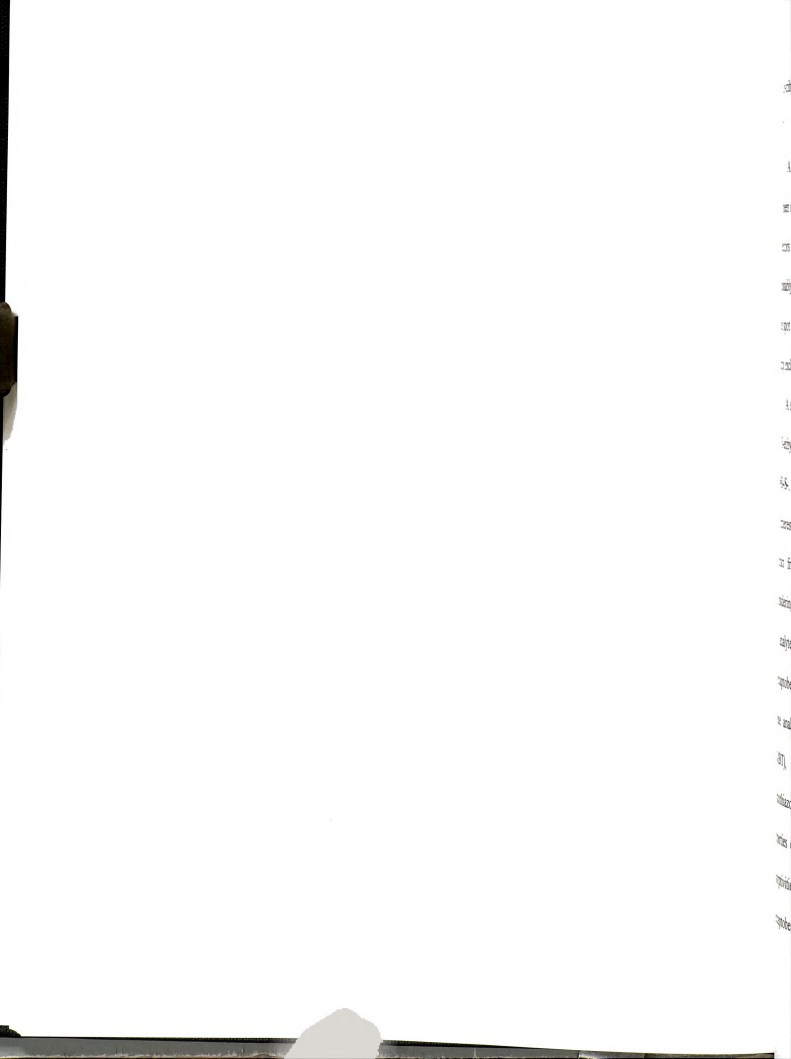
The fulfillment of some of these criteria can be assessed readily because physicochemical data on organic solids are available. Further, comparison of the relative abundance of analyte ions with the relative concentration of the analyte in the sample indicates that the ionization process is highly selective with efficient matrices. This selectivity requires that matrix-derived precursors ionize the analyte preferentially, otherwise a high matrix background and abundant metastable ions can be expected. Some matrices are selected for optimal performance with particular classes of analytes (11-14). For all of these reasons, new matrices are still usually found by trial and error.

Since Karas and Hillenkamp first reported nicotinic acid (3-pyridine carboxylic acid) as a matrix to desorb large proteins with 266-nm laser radiation (15), a number of other matrices of substantial utility has been discovered. Most of these compounds



reported to date are substituted aromatic compounds containing a carboxylic acid group. Several cinnamic acid derivatives, including ferulic, caffeic, and sinapinic acid, evaluated by Beavis and Chait (5), have made analysis by MALDI possible at laser wavelengths between 266 nm and 355 nm. These matrices, in particular sinapinic acid, form much less abundant photochemically-generated adducts with the analyte than does nicotinic acid. The matrix, 2,5-dihydroxybenzoic acid (DHB), introduced by Karas and co-workers (9), has been extremely useful for the analysis of proteins, carbohydrates, synthetic polymers and enzymatic digests at 337 and 355 nm, and it has a high tolerance to contaminants such as inorganic salts, buffers and detergents. Another matrix, α -cyano-4-hydroxycinnamic acid (α CHCA), is now widely accepted as the choice for the analysis of low-mass peptides and glycopeptides (2). The matrix 3-hydroxy-picolinic acid (16) was shown to desorb and ionize nucleotides effectively in the negative ion mode, and 2-[(4-hydroxyphenyl)azo]benzoic acid (HABA) was reported to be effective for the analysis of relatively large proteins (11).

Some neutral and basic matrices also have been reported in the literature (17-19). The performance of coumarin laser dyes was investigated by Perera and co-workers (17). These laser dyes were shown to be useful in the analysis of a variety of macromolecules including peptides, proteins, deoxynucleotides, and polymers at 337 nm. The matrix 3-aminoquinoline was reported to be effective for the analysis of polysaccharides and proteins (18). Fitzgerald et. al screened 37 substituted pyrimidines, anilines, and amino-pyridines as potential matrices (19). Although limited to the analysis of relatively small proteins and oligonucleotides, these basic matrices extended the utility of MALDI to acid-sensitive species. Other neutral matrices, such as 2,3,4-trihydroxyacetophenone and



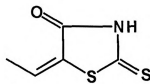
2,4,6-trihydroxyacetophenone were also reported to be good matrices for DNA detection (20).

A new liquid matrix which is a binary mixture of α CHCA and 3-aminoquinoline has been reported by Orlando's group for magnetic sector mass spectrometers with point detectors (21). This matrix was demonstrated to provide a very long-lasting and reasonably constant ion signal with each laser pulse, without the need to reposition the laser spot or adjust tuning of the mass spectrometer. The successful applications of this matrix included the analysis of oligosaccharides and polymers.

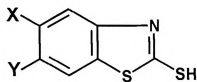
A survey of the literature indicated that many compounds such as dithioacetate (**I**) and 5-ethylidene-rhodanine (**II**) which contain the common chromophoric functionality - C(=S)-S-, have strong UV-absorption in the region of 330-350 nm (22). This initiated our interest in these compounds as part of a search for new matrices which are structurally distinct from the conventional hydroxybenzoic acids and hydroxycinnamic acids. Considering other physico-chemical characteristics, including compatible solubility with the analytes, vacuum stability and solid state morphology, we chose to evaluate 2-mercaptobenzothiazole (MBT) and its analogs (**III**) as potential matrices in MALDI-MS. These analogs having the core structure **III** are: 5-chloro-2-mercaptobenzothiazole (CMBT), 6-amino-2-mercaptobenzothiazole (AMBT), 2-mercapto-5 methoxybenzothiazole (MMBT), 6-ethoxy-2-mercaptobenzothiazole (EMBT). Pertinent properties of these mercaptobenzothiazoles are summarized in Table 2.1. The molar absorptivities at 337 nm (ϵ_{337}) were determined by measuring the absorbance of mercaptobenzothiazoles in a mixture of 1:1:1 EtOH/THF/water.



I



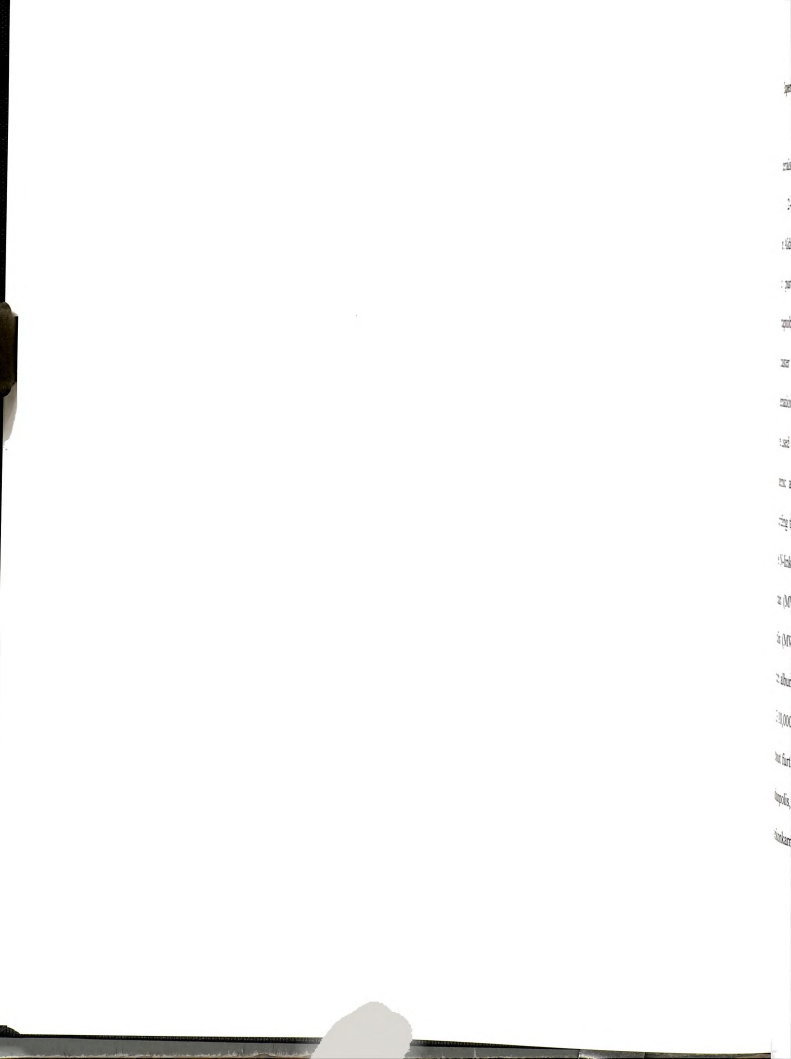
II



III

Table 1. Physical properties of MBT and its analogs.

Matrix	Structure	$\lambda_{\max}(\text{nm})$	ϵ_{337} ($\text{l mol}^{-1}\text{cm}^{-1}$)	MW
MBT	X = H, Y = H	326.8	1.4×10^3	167.25
CMBT	X = Cl, Y = H	331.7	1.4×10^4	201.70
AMBT	X = H, Y = NH_2	339.0	1.0×10^4	182.27
MMBT	X = OMe, Y = H	337.0	4.4×10^3	197.29
EMBT	X = H, Y = OEt	330.0	1.1×10^4	211.32



2. Experimental

Materials

2-Mercaptobenzothiazole and 5-chloro-2-mercaptobenzothiazole were purchased from Aldrich Chemical Co. (Milwaukee, WI, USA). 6-Amino-2-mercaptobenzothiazole was purchased from TCI America (Portland, OR, USA). 5-Methoxy-2-mercaptobenzothiazole and 6-ethoxy-2-mercaptobenzothiazole were purchased from Lancaster Synthesis Inc. (Windham, NH, USA). As purification by recrystallization, sublimation and ion exchange did not improve performance, all these matrix materials were used as received without further cleanup. Conventional matrices (including α CHCA, sinapinic acid, and DHB; Aldrich Chemical Co.) have been recrystallized and used according to literature protocols [5,6,10,18] to assure valid comparisons. Oligomannose-type N-linked oligosaccharide [Man]₅[GlcNAc]₂ (MW 1235.1 Da), Angiotensin I from human (MW 1,296.5 Da), porcine pancreastatin (MW 5,103 Da), bovine pancreatic insulin (MW 5,733.5 Da), horse skeletal muscle myoglobin (MW 16,951 Da), bovine serum albumin (MW 66,430 Da), bovine transferrin (MW 78,000 Da), PEG 3,350, and PEG 10,000 were purchased from Sigma Chemical Co. (St. Louis, MO, USA) and used without further purification. Ganglioside G_{T1b} was purchased from Boehringer Mannheim (Indianapolis, IN, USA). Dextran 5,000 was purchased from Fluka Chemical Corp. (Ronkonkoma, New York, USA).

200

M

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

198

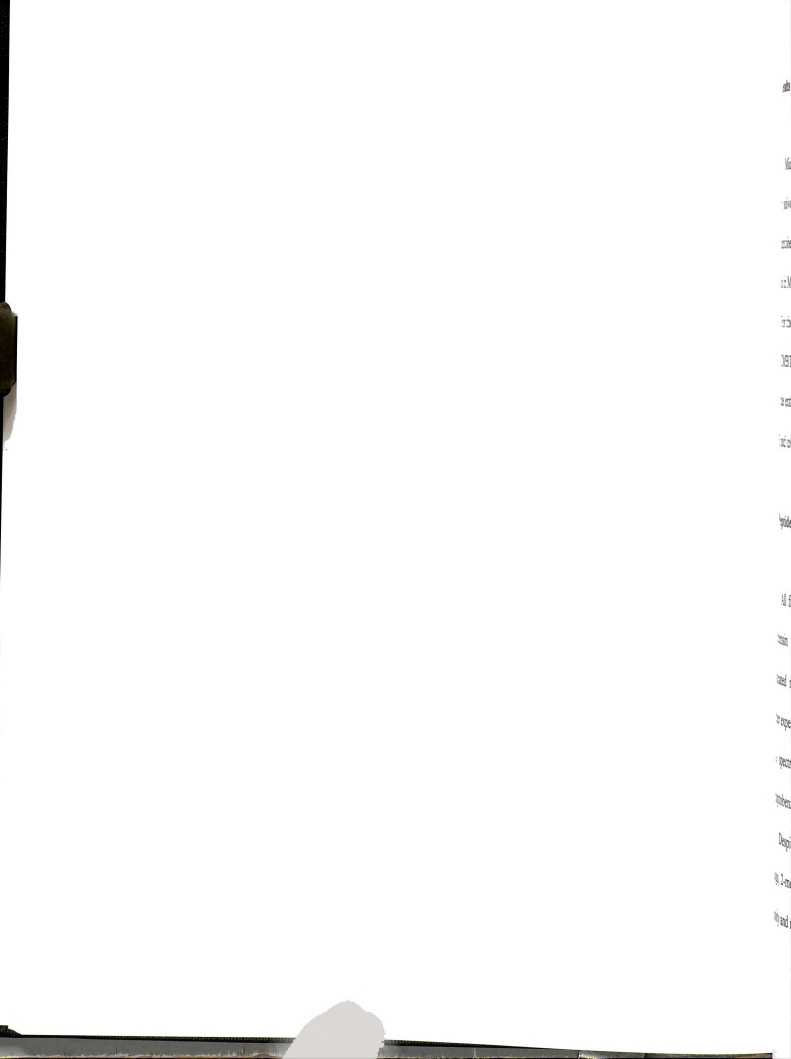
198

Instrument

MALDI-MS experiments were carried out on the PerSeptive Biosystems Voyager-Elite time-of-flight mass spectrometer equipped with a Model VSL-337ND nitrogen laser (Laser Science, Newton, MA, USA) (337 nm; 3-nsec pulse length) and a dual microchannel plate detector (Galileo, Sturbridge, MA). The accelerating voltage in the ion source was 30 KV. All MALDI-MS results were obtained in the linear mode without DE. The most important operational parameter was the careful regulation of the laser intensity. The laser irradiance per shot was kept as low as possible, i.e., close to the threshold of analyte ion production.

Sample Preparation

2-Mercaptobenzothiazole and analogs were dissolved in a mixture of ethanol (EtOH)/tetrahydrofuran (THF)/water (1:1:1) to make a solution of 10g/l. This matrix solution concentration was found to be the most generally applicable for a variety of analytes. Peptide and protein samples were dissolved in a solution of 1:1 acetonitrile/0.1% (v/v) aqueous trifluoroacetic acid or 1:1:1 EtOH/THF/water. Carbohydrate samples were dissolved in a solution of 1:1 EtOH/water or 1:1:1 EtOH/THF/water. Polymer samples were dissolved in a solution of 1:1:1 EtOH/THF/water. The samples for mass spectrometric analysis were typically prepared by depositing about 1 μ l of the matrix solution and an equal volume of the analyte solution on a sample plate well, which were then mixed by means of the pipette tip. The solvents were then removed by drying the samples in air at room temperature.



3. Results and Discussion

Microscopic inspection of the matrix/analyte mixture on the sample plate surface after solvent evaporation revealed homogeneous crystalline phases for all five mercaptobenzothiazoles. Light micrographs obtained from samples of bovine pancreatic insulin in MBT and CMBT (molar ratio of analyte to matrix: 1;1000) are shown in Figure 2.1. For the samples with MBT, evenly spread large crystals were observed. The samples with CMBT, on the other hand, formed a uniform layer of very fine crystals distributed over the entire sample plate surface. Light micrographs obtained from the same sample in DHB and α CHCA are also included in Figure 2.1 for comparison.

3.1. Peptides and Proteins

All five selected mercaptobenzothiazoles were capable of desorbing and ionizing angiotensin and bovine pancreatic insulin with very strong, stable signals for the protonated molecules, comparable to those obtained with α CHCA or sinapinic acid. Further experiments with higher molecular weight proteins showed that the quality of the mass spectra obtained and the mass range applicable varied with different mercaptobenzothiazoles.

Despite a lower molar absorptivity at 337 nm relative to that of some of its analogs, 2-mercaptobenzothiazole (MBT) produced MALDI spectra with the best signal intensity and resolution for the peptides and proteins with molecular weights up to at least

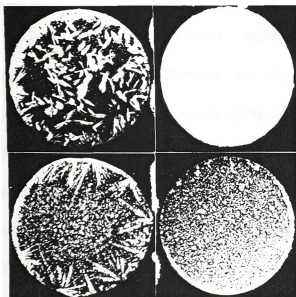


Figure 2.1: Light micrographs (magnification: $\times 14$) obtained from samples of bovine pancreatic insulin in different matrices (molar ratio of insulin to matrix: 1:1000). top left panel: MBT; top right panel: CMBT; (c) bottom left panel: DHB; bottom right panel: α CHCA.

[illegible]

100,000 daltons. Fig. 2 presents the positive MALDI spectra of human angiotensin I (Fig. 2a), porcine pancreastatin (Fig. 2b), and bovine transferrin (Fig. 2c) using MBT as the matrix. The amount of protein loaded on the sample plate cell was about 1 to 5 picomoles. In all these spectra, the major peak observed corresponds to the singly charged (protonated) analyte denoted by $[M+H]^+$. Signals corresponding to the dimer of bovine transferrin (Fig. 2c) were also observed. For this, and other proteins examined (data not shown), the high-mass cluster ions are usually of significantly lower abundance. Multiply-charged analyte signals were also detected for all the analytes except angiotensin. The highest charge state observed was +4 in the spectrum of bovine transferrin (Fig. 2c). No obvious fragmentation was observed. The matrix adduct ion peaks were weak, but well resolved for small proteins such as porcine pancreastatin (MW 5,103 Da, Fig. 2b). Like α CHCA and sinapinic acid, MBT forms matrix-generated adducts of much lower abundance than that of the protonated molecules. In these spectra, the measured mass resolution ($M/\Delta M$, FWHM) for the peak representing the protonated molecule, was 766 for angiotensin, 540 for porcine pancreastatin, and 63 for bovine transferrin, respectively. Additionally, MBT is a suitable matrix for the analysis of small peptides and proteins because of the absence of matrix ions above m/z 500. These and other results (data not shown) from examining over 30 different peptides and proteins demonstrate that MBT provides similar performance to that obtained with α CHCA or sinapinic acid in terms of sensitivity and mass resolution.

12.
13.
14.
15.
16.

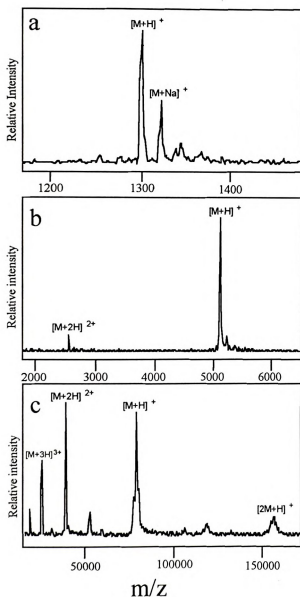


Figure 2.2: Positive MALDI mass spectra of three proteins with MBT as matrix. (a) human angiotensin I (MW 1,296.5 Da); (b) porcine pancreastatin (MW 5,103 Da); (c) bovine transferrin (MW 78000 Da). The data were acquired by summing 30 to 100 laser shots. The amount of proteins loaded on the sample plate was about 1 to 5 picomoles.



MBT also has been used successfully for protein analysis in the negative ion mode.

A typical negative ion mass spectrum of bovine serum albumin (BSA) with MBT as matrix is shown in Figure 2.3. In addition to the deprotonated molecule $[M-H]^-$, doubly- $[M-2H]^{2-}$ and triply- $[M-3H]^{3-}$ charged (deprotonated) molecular ion signals were observed in this spectrum.

One of the strengths of MALDI-MS is its relative tolerance to contaminants (such as buffers, salts) in the sample. However, it is generally agreed that contaminants like detergents which are sometimes used for protein sample preparation and stability degrade the quality of spectra with respect to their signal-to-noise ratio and mass resolution, making the analysis of protein samples in the presence of surfactant agents by MALDI-MS very challenging (9-10, 25). The most deleterious spectral effects have been assigned to sodium dodecylsulfate (SDS), an ionic detergent (10, 25). With MBT as matrix, inorganic salts, and buffers in the analyte solution up to concentrations of at least 200 mM, as well as non-ionic detergents such as urea up to concentrations of at least 1M, do not lead to a significant decrease in detection limits or quality of spectra. Furthermore, MBT was also found to have excellent tolerance to ionic detergents such as SDS. Figure 2.4 (a) presents the positive MALDI spectrum of myoglobin (10^{-5} M in acetonitrile/0.1% (v/v) aqueous trifluoroacetic acid (1:1) solution) with MBT as matrix. The tolerance of MBT to SDS is illustrated in Figure 2.4(b) and Figure 2.4(c) where the myoglobin sample was applied in a 1 g/l SDS (0.1%, w/v) and 100 g/l (10%, w/v) solution, respectively. SDS caused the background signal to be elevated in the low-mass region. The signal-to-background ratio for myoglobin has deteriorated compared to that in Fig. 4(a), but singly-charged and doubly-charged protein signals are clearly present in both Figure 2.4(b) and Figure 2.4(c).

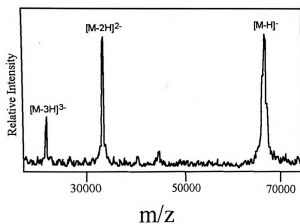


Figure 2.3: Negative ion MALDI spectrum of bovine serum albumin (MW 66,430 Da) with MBT as matrix; 200 laser shots summed. Total protein loaded: 5 pmol.

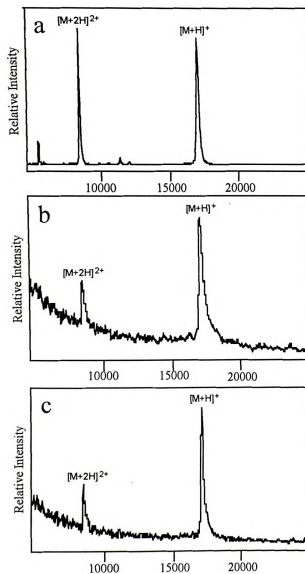
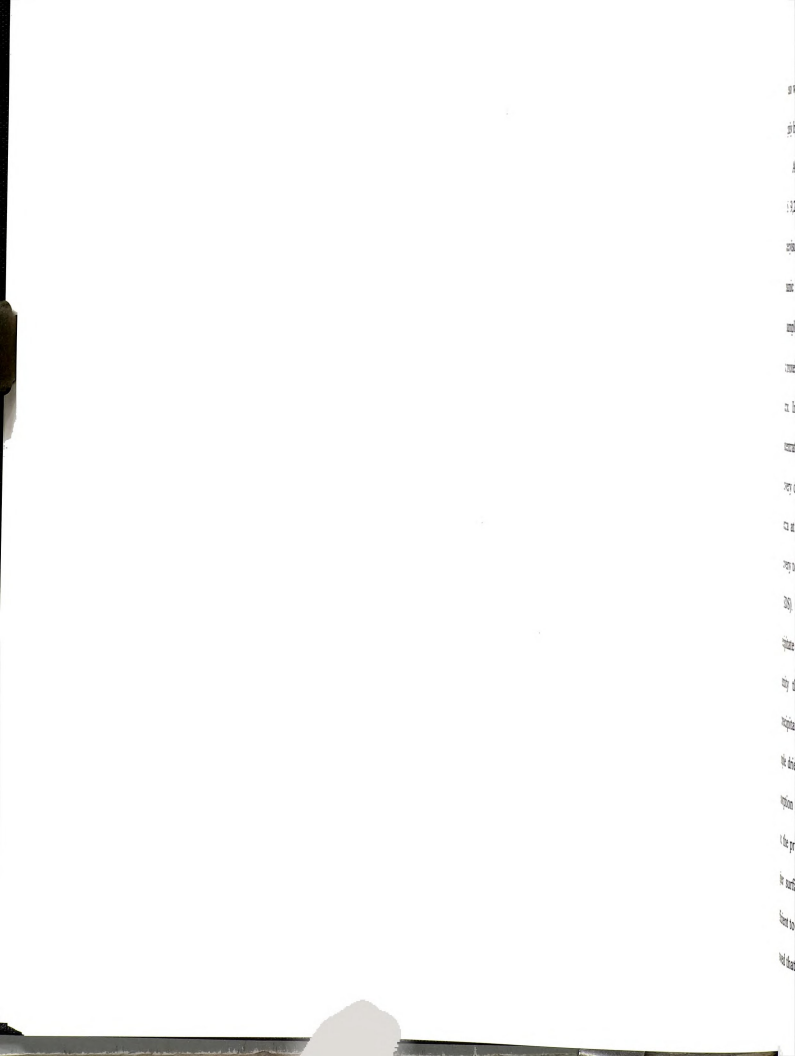


Figure 2.4: Positive MALDI spectrum of horse skeletal muscle myoglobin (10 pmol); 100 laser shots summed. The sample was dissolved in acetonitrile/0.1% (v/v) aqueous trifluoroacetic acid (1:1) solution. (a) no sodium dodecylsulfate (SDS) added, matrix: MBT; (b) with the addition of 0.1% SDS, matrix: MBT; (c) with the addition of 10% SDS, matrix: MBT.



It also was noted that the signal-to-noise ratio and mass resolution in Figure 2.4(c) are slightly better than that in Figure 2.4(b) in spite of more SDS added.

Among the matrices reported in the literature, only DHB shows some tolerance to SDS (9,25). It has been reported before that even minimal contamination with sodium dodecylsulfate (SDS) completely suppresses the protein signals when matrices such as cinnamic acid analogs including α -cyano-4-hydroxycinnamic acid (α CHCA) are used in the sample preparation (10, 25). However, a recent article by Amado et al. (26) argued that protein samples containing SDS can be analyzed by MALDI-MS with α CHCA as the matrix. In their results, a decrease in the protein signal also was observed for increasing concentrations of SDS up to 0.1% (w/v), at which no signal was observed. But a recovery of the signals was observed at concentrations of SDS above 0.3%, with useful spectra at concentrations of SDS as high as 10%. They attribute the sample signal recovery to the formation of micelles above the critical micelle concentration (cmc, 0.23% for SDS). In their theory, at subcritical concentrations, the protein molecule may partially precipitate by the formation of ion pairs with the surfactant monomers, diminishing the quantity that cocrystallizes with the matrix, while the excess of surfactant may coprecipitate with the matrix, a part of the surfactant remains in the liquid phase. As the sample dries, the surfactant effectively coats the crystals, which may partially prevent desorption of the protein. Also in their theory, when SDS concentrations are above the cmc, the protein is stabilized and solubilized by surfactant micelles, which may lead to a higher surface concentration of the protein resulting in an increasing ion formation sufficient to compete with surfactant ion formation. Our results using MBT as the matrix showed that the spectra quality with 10% SDS addition was better than that with 0.1%

166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

SDS addition, which somewhat agreed with what Amado et al. have observed. However, contrary to what they observed, the protein signal was detectable with 0.1% SDS addition using MBT as the matrix as shown in Figure 2.4(b).

Other analogs of 2-mercaptobenzothiazole were found to be effective matrices for the analysis of peptides and low-mass proteins in the range of 800 to 6000 daltons. Figure 2.5 shows a mass spectrum of bovine pancreatic insulin (sample loading: 3 pmol) using CMBT as the matrix. The mass resolution (FWHM) for the major peak at m/z 5734.5 is about 650. Peaks for the protonated molecule $[M+H]^+$, sodium adduct $[M+Na]^+$, potassium adduct $[M+K]^+$, and matrix adduct $[M+202]^+$ are well resolved. When these analogs are used for the analysis of higher-mass proteins, however, peak broadening becomes more obvious, resulting in lower signal intensity and mass resolution, which may partly be attributed to the formation of unresolved matrix adduct ion peaks.

3.2. Carbohydrates

Although sodium and potassium adduct ions are almost always present in the MALDI spectra of peptides and proteins, the most abundant signals observed are from protonation of analyte molecules. However, when MALDI is applied to carbohydrates, cation attachment usually predominates, so that protonated molecules are rarely detected. The cinnamic acid-type matrices have been used successfully for the analysis of peptides and proteins, but these matrices are not useful for the analysis of carbohydrates. However, DHB, which is a good cationizing matrix, was found effective for carbohydrate analysis.

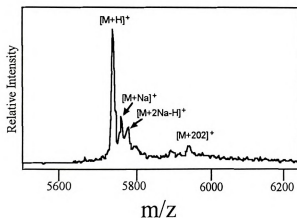


Figure 2.5: Positive MALDI spectrum of bovine pancreatic insulin (MW 5,733.5 Da) using CMBT as matrix; 30 laser shots summed. The region of the spectrum shown corresponds to a mass range of 5,500–6,150. Total protein loaded: 2 pmol.

As a result of the above, the following conclusions can be drawn:

Among the five selected mercaptobenzothiazoles, only CMBT was found to be a successful cationizing matrix for the analysis of carbohydrates. The performance of CMBT as a cationizing matrix was compared with that of DHB in the analysis of an oligomannose-type N-linked oligosaccharide, $[\text{Man}]_5[\text{GlcNAc}]_2$ (MW 1,235.1 Da), as illustrated in Figure 2.6. The amount of carbohydrate loaded on the sample plate cell was 2 pmol. With either CMBT or DHB as matrix, the major peak in the mass spectrum represents the sodium-cationized molecule, $[\text{M}+\text{Na}]^+$. A potassium adduct ion, $[\text{M}+\text{K}]^+$, is also observed with much less abundance. However, CMBT (Figure 2.6(a)) offers a much higher signal-to-background ratio for this oligosaccharide sample than does DHB (Figure 2.6(b)). When sample loading was decreased to 100 fmol, no signal could be obtained with DHB as matrix. Even with binary matrices such as super-DHB (a mixture of DHB and 5-methoxysalicylic acid) (23) and DHB/HIC (a mixture of DHB and 1-hydroxyisoquinoline) (24), which have been reported to improve the performance of DHB, no signal could be obtained at the 100-fmol level (data not shown). However, with CMBT the carbohydrate signal is still detectable as shown in Figure 2.6(c), indicating the better sensitivity offered by this new matrix. Furthermore, cation doping experiments were performed to examine the effect of adventitious salts on performance. NaCl was added to the sample solution (at final concentrations of 50 mM) when CMBT or conventional matrices (DHB and binary DHB mixtures) were used in the analysis of $[\text{Man}]_5[\text{GlcNAc}]_2$. No significant difference in terms of the performance of the individual matrices was observed.

As reported by other research groups (25), the solid state morphology of the matrix/analyte mixture plays an important role in MALDI ion yields. A practical

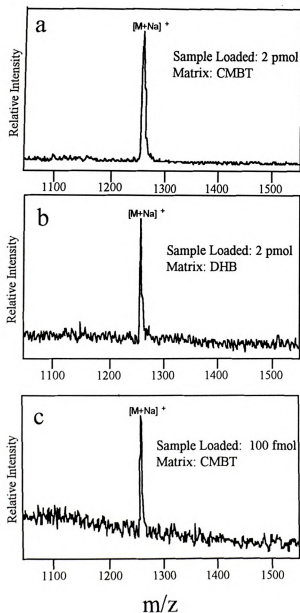


Figure 2.6: Positive MALDI spectrum of oligomannose-type N-linked oligosaccharide

$[\text{Man}]_5[\text{GlcNAc}]_2$ (MW 1,235.1 Da); 40 laser shots summed. (a) sample loaded: 2 pmol, matrix: CMBT; (b) sample loaded: 2 pmol, matrix: DHB; (c) sample loaded: 100 fmol, matrix: CMBT.

1884

1885

1886

1887

1888

1889

1890

1891

1892

1893

1894

1895

1896

1897

1898

1899

1900

1901

1902

1903

1904

1905

1906

1907

advantage of using CMBT as a matrix is the capacity to generate ions uniformly from sites across the target owing to its homogeneous crystallization with the analyte over the entire sample surface area. Almost any portion of the sample surface can be used to obtain the resulting spectra shown above. Good quality spectra could be obtained with DHB, but the results were not as reproducible as those obtained by CMBT because of the irregular crystallization process for DHB during sample preparation. DHB tends to form large crystals on the rim of the target (although DHB/HIC is more homogenous), leaving a microcrystalline region in the center, which yields the strongest signals for carbohydrates (9,12). In this study, different regions of the target were irradiated to generate the best signal possible.

The potential of CMBT as a matrix for oligosaccharide analysis is demonstrated in Figure 2.7 in which dextran 5,000 was selected as a representative example of oligosaccharides with a high molecular mass range. Dextran 5000 consists of a mixture of polysaccharides containing a backbone of D-glucose units produced by bacteria growing on a sucrose substrate. The MALDI spectrum by CMBT shows oligomers up to 5,500 daltons with good resolution and signal-to-background ratio.

Gangliosides are a class of glycosphingolipids that consist of a hydrophobic ceramide and a hydrophilic carbohydrate chain containing one or more sialic acid residues at various linkage sites. These compounds can be analyzed by MALDI-MS with a number of matrices such as DHB, HABA, and 6-aza-2-thiothymine (26). CMBT also was found to be an efficient matrix for these glycolipids, in both the positive and negative ion mode. Negative ion mass spectra of the underivatized gangliosides were always of better quality

101

102

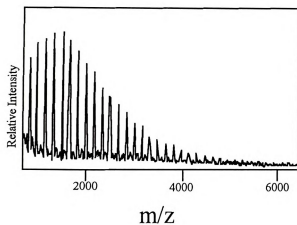


Figure 2.7: Positive MALDI spectrum of dextran 5,000 with CMBT as matrix; 50 laser shots summed. Adjacent oligomer peaks in the spectrum are spaced by 162 mass units.

1. The first step is to identify the problem. This involves understanding the current situation and the goals that need to be achieved.

than the positive ion spectra, exhibiting better sensitivity, resolution, and less fragmentation from loss of sialic acid residues. Figure 2.8 shows the mass spectrum of underivatized G_{T1b} , containing three sialic acid residues, obtained in the negative ion mode by using CMBT as matrix. No fragments were observed. Two prominent molecular species were detected in the sample, differing from each other by two methylene units. The most abundant peaks corresponded to $[M-2H+K]^-$ ions of each species, while $[M-H]^-$ peaks were less intense.

CMBT also was found to have superior performance over conventional matrices in analyzing complex carbohydrates such as the muropeptides (27) derived from peptidoglycan, a structural component of the cell walls of bacteria. Figure 2.9 shows the structure of a trimeric muropeptide derived from peptidoglycan that was isolated from a highly methicillin-resistant *Staphylococcus aureus* Strain COL (28-29). Figure 2.10(a) displays the mass spectrum of this trimeric muropeptide using CMBT as matrix. Sodiated analyte signals are the dominant peaks in the spectra. A multiple sodium adduct ion, $[M+2Na-H]^+$, and a prompt fragmentation peak $[M+Na-GlcNAc]^+$ also are present in the mass spectra. Most of the other peaks present in the spectra are suspected to be impurities that coeluted in the HPLC fraction, or decomposition products. The mass spectrum presented in Figure 2.10(a) was recorded from a total sample load of 3 pmol indicating high sensitivity analysis can be achieved by MALDI-MS using CMBT as the matrix. In contrast, only weak signals were observed for this sample with conventional matrices such as α CHCA, DHB, or sinapinic acid as shown in Figure 2.10(b) where α CHCA was used as an example.

1. 20

2.8

races.

10

10

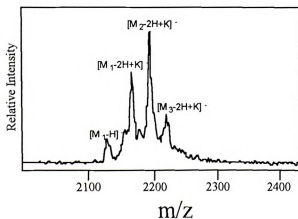
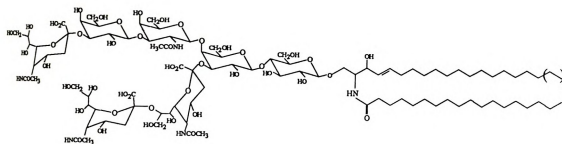


Figure 2.8: Negative MALDI spectrum of underivatized G_{T1b} with CMBT as matrix; 40 laser shots summed. M_1 and M_2 are two molecular species in the sample differing by two methylene units in the long-chain base. $[M_1 + K - 2H]^-$ and $[M_2 + K - 2H]^-$ of two molecular species are represented by dominant peaks in the spectrum.

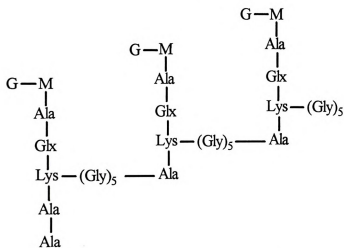


Figure 2.9: Structure of trimeric mucopeptide derived from the peptidoglycan of *S. aureus* strain COL. The following abbreviations were used: G, *N*-acetylglucosamine; M, *N*-acetylmuramic acid.

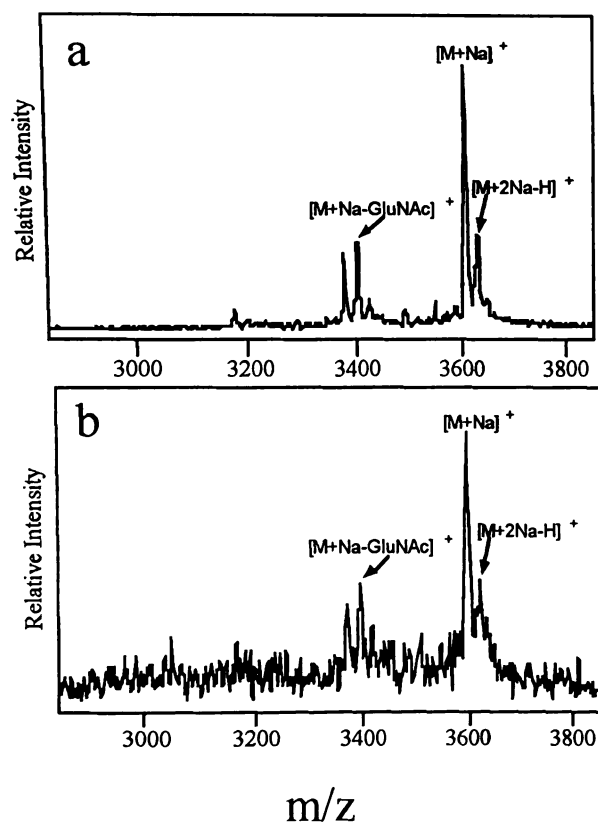


Figure 2.10: Positive MALDI spectrum of trimeric mucopeptide derived from peptidoglycan of *S. aureus* strain COL with (a) CMBT (b) α CHCA as matrix. 40 laser shots summed. Total sample loaded was approximately 3 pmol.



3.3. Synthetic Polymers

Synthetic polymers are another class of compounds that is typically more easily ionized by cation attachment (30-32). All five selected mercaptobenzothiazoles proved to be useful for analyzing synthetic polymers. The positive ion mass spectrum obtained from PEG 3,350 with AMBT as matrix is illustrated in Figure 2.11. No alkaline salts were added to this polymer sample prior to analysis. The spectrum clearly shows well resolved adjacent oligomer signals with masses ranging from 2,440 daltons to 4,600 daltons in Figure 2.11, reflecting the molecular weight distribution of PEG 3,350. Two series of peaks were observed originating from Na^+ and K^+ adducts. The peaks corresponding to the $[\text{M}+\text{Na}]^+$ ions of the oligomers were centered at a value corresponding to a M_w of 3,362 daltons, close to the weight-averaged molecular weight provided by the manufacturer (M_w : 3,350 Da). The mass difference between adjacent $[\text{M}+\text{Na}]^+$ ion signals was found to be approximately 44.2 daltons.

The potential of using mercaptobenzothiazoles for the analysis of higher molecular weight synthetic polymers such as PEG 10,000 is demonstrated in Figure 2.12. Again, AMBT is used as the matrix, although similar results can be obtained with other mercaptobenzothiazoles, including MBT and CMBT. NaCl was doped into the PEG 10,000 sample to assist cationization. The shape of the molecular weight distribution and mass of the repeat unit were well defined, although the oligomer ion peaks of PEG 10,000 could not be resolved completely. The oligomer distribution ranged from m/z 10,300 to ca. 14,600. The weight-averaged molecular weight obtained from this MALDI

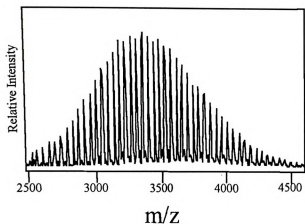


Figure 2.11: Positive MALDI spectrum of PEG 3,350 with AMBT as matrix. 30 laser shots summed. No alkaline salts were added to the polymer solution.

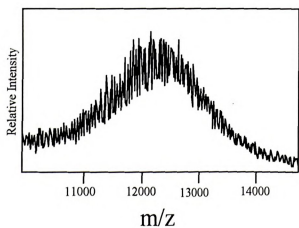


Figure 2.12: Positive MALDI spectrum of PEG 10,000 with AMBT as matrix. 100 laser shots summed. NaCl was added to the polymer solution.

1. The first step is to identify the problem. This involves understanding the current situation and the goals that need to be achieved.

spectrum is ca. 12,600 which shows about 10% positive discrepancy from the GPC data provided by the manufacturer ($M_w = 11,500$ daltons). This observation agrees well with the results reported by two other research groups (33-34), where MALDI-MS showed about 10% higher weight-averaged molecular weight for PEG 8,000, relative to GPC. We also have applied mercaptobenzothiazoles in the analysis of nonpolar synthetic polymer molecules, such as poly(methyl methacrylate) 15,000. Although molecular weight distribution information could be determined from the spectrum (data not shown), the signal intensity is not as strong as that obtained from PEG samples.

4. Conclusions

The performance of five selected mercaptobenzothiazoles has been evaluated as a matrix for MALDI-MS at 337 nm. It was demonstrated that 2-mercaptobenzothiazole is applicable to the analysis of peptides and proteins of molecular weight up to at least 100,000 daltons with comparable sensitivity and resolution to those achieved with conventional matrices such as α CHCA and sinapinic acid. 5-Chloro-2-mercaptobenzothiazole (CMBT) was found not only to be effective as a matrix for the analysis of peptides, low-mass proteins, and glycolipids, but also superior to conventional matrices for the analysis of oligosaccharides and mucopeptides. Synthetic polymer analysis also is possible by MALDI-MS with these new matrices. In view of the fact that most matrices presently employed are essentially acidic in nature, introduction of a family of neutral materials (e.g., pKa for MBT is 7.5) will complement the limited choice of matrices available to deal with acid-sensitive biomolecules.

References:

1. Ehring, H., Karas, M., and Hillenkamp, F., 1992, 27, 472-480.
2. Beavis, R. C. and Chait, B. T., Org. Mass Spectrom., 1992, 27, 156-158.
3. Gimon, M. E., Preston, L. M., Solouki, T., White, M. A., and Russell, D. H., Org. Mass Spectrom., 1992, 27, 827-830.
4. Juhasz, P., Wang, B. H., and Biemann, K., *Proceedings of the 40th Conference on Mass Spectrometry and Allied Topics*, Washington, DC, May 31-June 5, 1992, p 372-373.
5. Beavis, R. C. and Chait, B. T., Rapid Commun. Mass Spectrom., 1989, 3, 233-237.
6. Beavis, R. C., *Proceedings of the 40th Conference on Mass Spectrometry and Allied Topics*, Washington, DC, May 31-June 5, 1992, p 5.
7. Hillenkamp, F.; Karas, M.; Beavis, R. C.; and Chait, B. T., Anal. Chem., 1991, 63,1193A-1198A.
8. Beavis, R. C. and Chait, B. T., Anal. Chem., 1990, 62, 1836-1840.
9. Strupat, K., Karas, M., and Hillenkamp, F., Int. J. Mass Spectrom. Ion Processes, 1991, 111, 89-102.
10. Beavis, R. C. and Chait, B. T., Proc. Natl. Acad. Sci. 1990, 87, 6873-6878.
11. Juhasz, P., Costello, C. E., and Biemann, K., J. Am. Soc. Mass Spectrom., 1993, 4, 399-409.
12. Stahl, B., Steup, M., Karas, M., and Hillenkamp, F., Anal. Chem., 1991, 63, 1463-1466.

13. Huberty, M. C., Vath, J. E., Yu, W., and Martin, S. A., *Anal.Chem.*, **1993**, 65, 2791-2796.
14. Tang, K., Allman, S. L., Chen, C. H., *Rapid Commun. Mass Spectrom.*, **1992**, 6, 365-368.
15. Karas, M. and Hillenkamp, F., *Anal Chem.*, **1988**, 60, 2299-2303.
16. Wu, K. J., Steding, A., Becker, C. H., *Rapid Commun. Mass Spectrom.*, **1993**, 7, 142-146.
17. Perera, I. K., Kantartzoglou, S., Dyer, P. E., *Int. J. Mass Spectrom. Ion Processes*, **1994**, 137, 151-156.
18. Metzger, J. O., Woisch, R., Tuszyński, W., Angermann, R. *Fresenius J. Anal. Chem.*, **1994**, 349, 473-475.
19. Fitzgerald, M., Parr, G. R., Smith, L., M. *Anal. Chem.*, **1993**, 65, 3204-3211.
20. Zhu, Y. F., Chung, C. N., Taranenko, N. I., Allman, S. L., Martin, S. A., Chen, C. H., Haff, L., *Rapid Commun. Mass Spectrom.*, **1996**, 10, 383-388.
21. Kolli, V. S. K. and Orlando, R., *Anal. Chem.*, **1997**, 69, 327-332.
22. *Organic Electronic Spectral Data*, Vol. 1, **1946-52**.
23. Rosinke, B., Strupat, K., Hillenkamp, F., Rosenbusch, J., Dencher, N., Krüger, U.; Galla, H. J., *Mass Spectrom.*, **1995**, 30, 1462-1468.
24. Karas, M., Ehring, H., Nordhoff, E., Stahl, B., Strupat, K., Hillenkamp, F., Grehl, M., Krebs, B., *Org. Mass Spectrom.*, **1993**, 28, 1476-1481.
25. Mohr, M.D., Börnsen, K.O., Widmer, H.M., *Rapid Commun. Mass Spectrom.*, **1995**, 9, 809-814.

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

26. Amado, F. M. L., Santana-Marques, M. G., Ferrer-Correia, J., and Tomer, K. B., *Anal. Chem.*, **1997**, 69, 1102-1106.
27. Allmaier, G. and Schmid, E. R., in *Bacterial Growth and Lysis*; de Pedro M. A., Höltje J.- V., Löffelhardt, W., ED.; Plenum Press, New York, **1993**; FEMS Symposium No.65, p 23.
28. de Jonge, B. L. M., Chang, Y. -S., Gage D. A., Tomasz, A., *J. Biol. Chem.*, **1992**, 267, 11248-11253.
29. de Jonge, B. L. M., Chang, Y. -S., Gage D. A., Tomasz, A., *J. Biol. Chem.*, **1992**, 267, 11255-11259.
30. Bahr, V., Deppe, A., Karas, M., Hillenkamp, F., *Anal. Chem.*, **1992**, 64, 2866-2870.
31. Cotterl, J. S., Koerner, M., Gerhards, R., *Rapid Commun. Mass Spectrom.*, **1995**, 9, 1562-1564.
32. Larsen, B. S., Simonsick, W. J. Jr., McEwen, C. N., *J. Am. Soc. Mass Spectrom.*, **1996**, 7, 287-292.
33. Tang, X., Dreifuss, P. A., Vertes, A., *Rapid Commun. Mass Spectrom.*, **1995**, 9, 1141-1147.
34. Dey, M., Castoro, J. A., Wilkins, C. L., *Anal. Chem.*, **1995**, 67, 1575-1580.

Imp

for

type

every

action

the

every

every

every

every

the

comes

is well

the, exp

and be

He

Mon

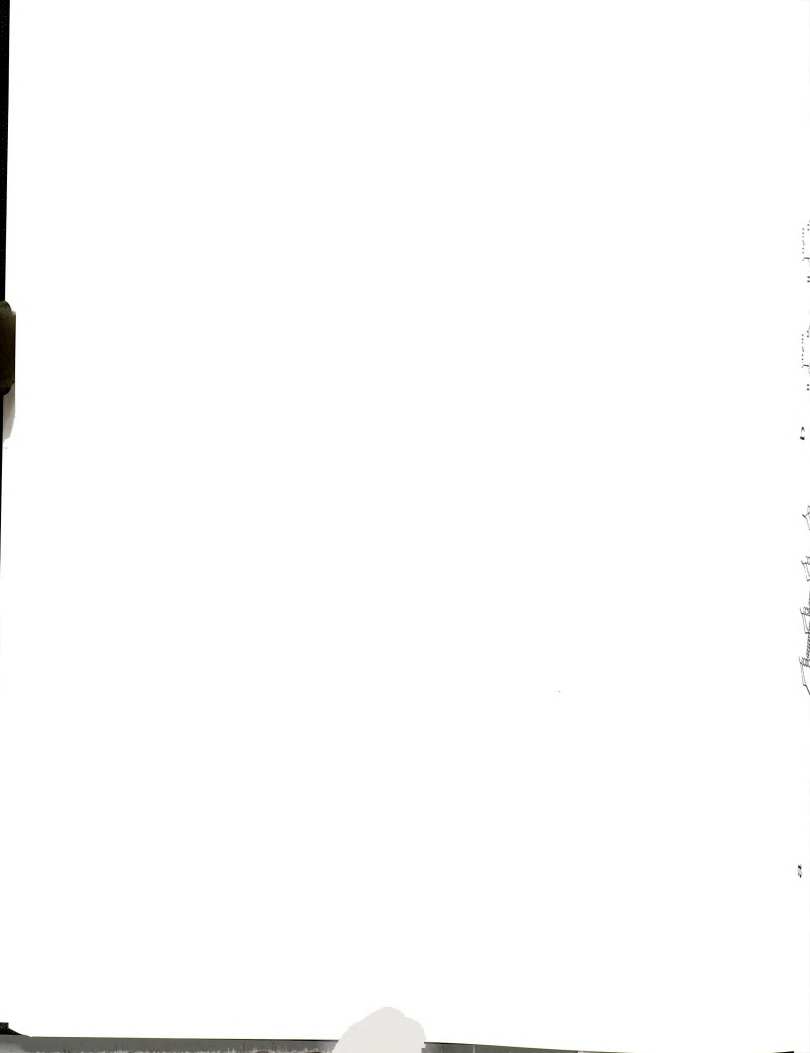
abstain

Chapter 3. Structural Characterization of Peptidoglycan Muropeptides by MALDI-MS and Post-Source Decay Analysis

1. Introduction

Peptidoglycan, a key component of bacterial cell walls, is a heteropolymer with unique chemical composition (1). It is composed of sugar chains consisting of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues that are interlinked *via* the muramic acid residues through short peptides, which include D-amino acids as illustrated in Figure 3.1 with *Staphylococcus aureus* bacterial cell wall peptidoglycan used as an example. The network that this polymer forms surrounds the cytoplasmic membrane and protects the cell from osmotic rupture (2). Because of this property, its ubiquity among bacteria and absence in humans, antibiotics that interfere with the peptidoglycan metabolism have been successfully used to combat bacterial infections. Such antibiotics including penicillin and methicillin specially bind to and inactivate enzymes that function to cross-link the peptidoglycan strands of bacterial cell walls. Since cell wall expansion also requires the action of enzymes that degrade and resynthesize cell walls, exposure of growing bacteria to such antibiotics results in their lysis, that is, the normal balance between cell wall biosynthesis and degradation is disrupted.

However, recently, growing resistance has emerged against this class of drugs (3-5). Mounting numbers of bacterial species have been found to be multiply resistant to antibacterial drugs. In particular, infections caused by methicillin-resistant



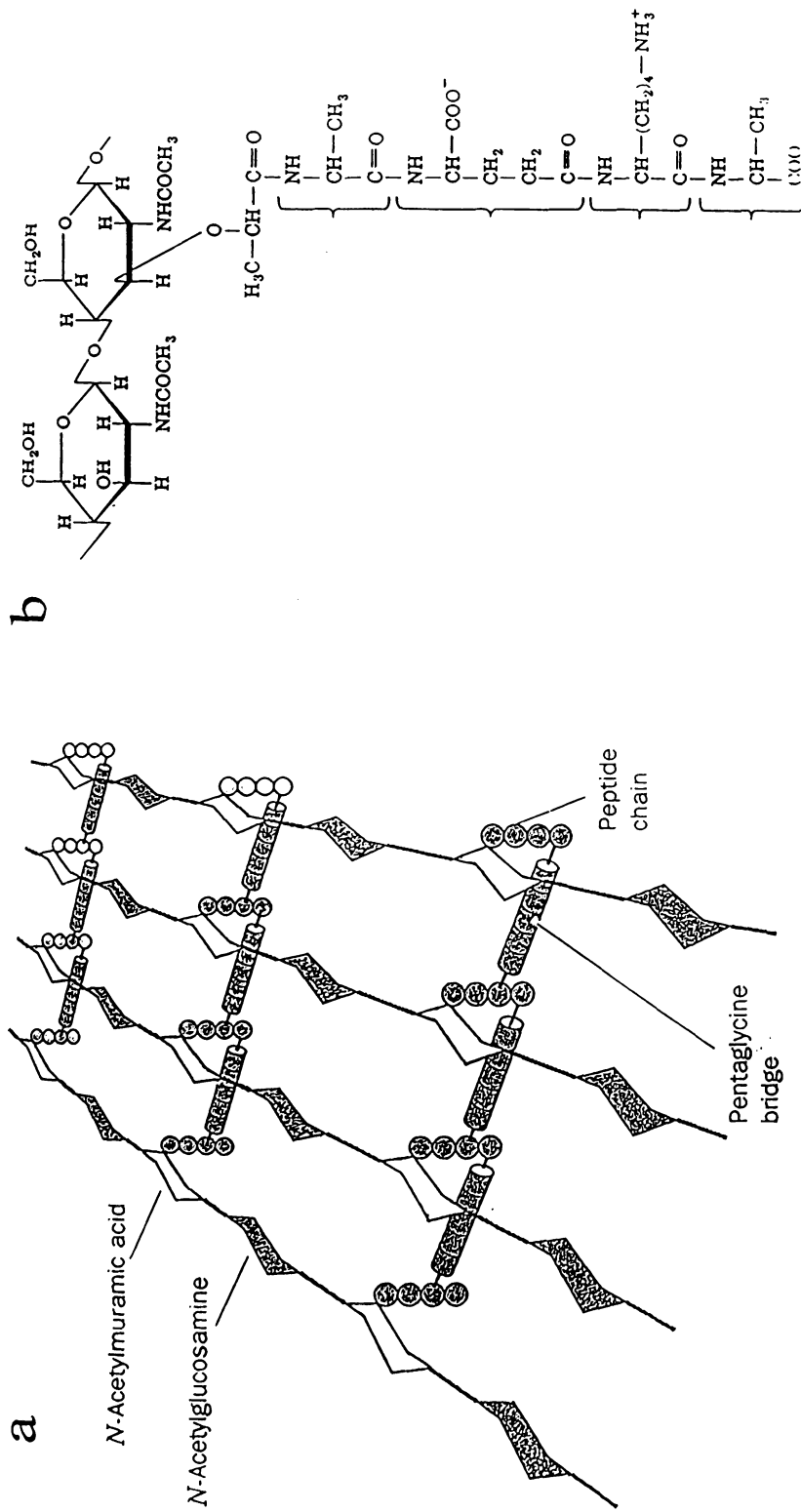


Figure 3.1: (a) The *Staphylococcus aureus* bacterial cell wall peptidoglycan. (b) The repeating unit of peptidoglycan is a *N*-acetylglucosamine (GlcNAc) - *N*-acetylmuramic acid (MurNAc) disaccharide whose lactyl side chain forms an amide bond with a peptide.

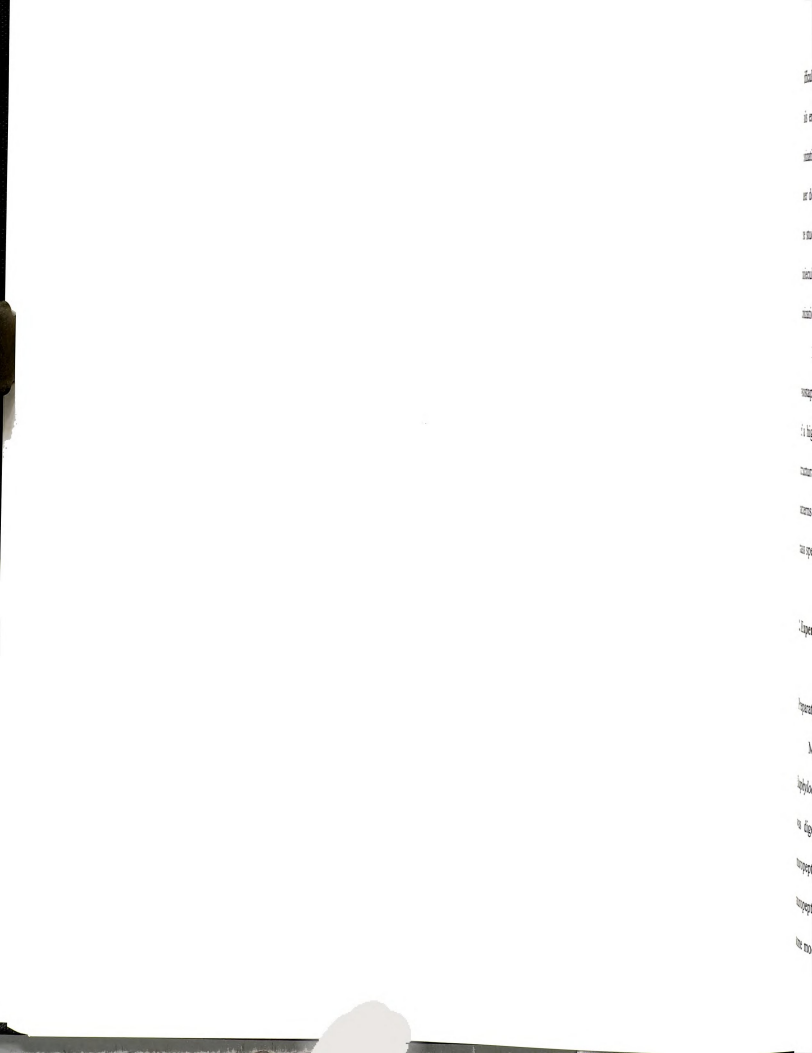


Staphylococcus aureus and vancomycin-resistant enterococci have now become difficult to treat, presenting a potential world-wide health crisis. The Centers for Disease Control & Prevention in Atlanta reported that the death rate from infectious diseases crept up 58% between 1980 and 1992, from 41 deaths per 100,000 population to 65. With the realization that the last of the drug weapons against bacterial infections could be blunted, medicinal chemists have intensified their effort to restock the arsenal.

To assess the kind of modifications that are needed in antimicrobial therapy, understanding of the resistance mechanisms is required. One approach to determine whether resistance is caused by changes in the peptidoglycan metabolism is a comparison of the peptidoglycan compositions of the resistant and sensitive strains. We became interested in this area in the course of a collaboration with a group at The Rockefeller University who had isolated peptidoglycans from methicillin-resistant *Staphylococcus aureus* strains.

HPLC, amino acid analysis, fast atom bombardment mass spectrometry (FAB-MS) and collisionally activated dissociation tandem mass spectrometry (FAB-CAD-MS/MS) (6-10) have been primarily used to characterize small disaccharide fragments (called muropeptides) produced from peptidoglycan by degradation with a muramidase, an enzyme that specifically hydrolyzes the β 1 \rightarrow 4 linkages between N-acetylglucosamine and N-acetylmuramic acid. These muropeptides can exist as monomers (the disaccharide unit plus short peptide chain) or can be cross-linked *via* their peptide chains to varying degrees to form dimers and higher oligomers.

While FAB-MS and FAB-CAD-MS/MS have been successfully employed to structurally characterize muropeptide monomers and dimers at the nanomole level, it is



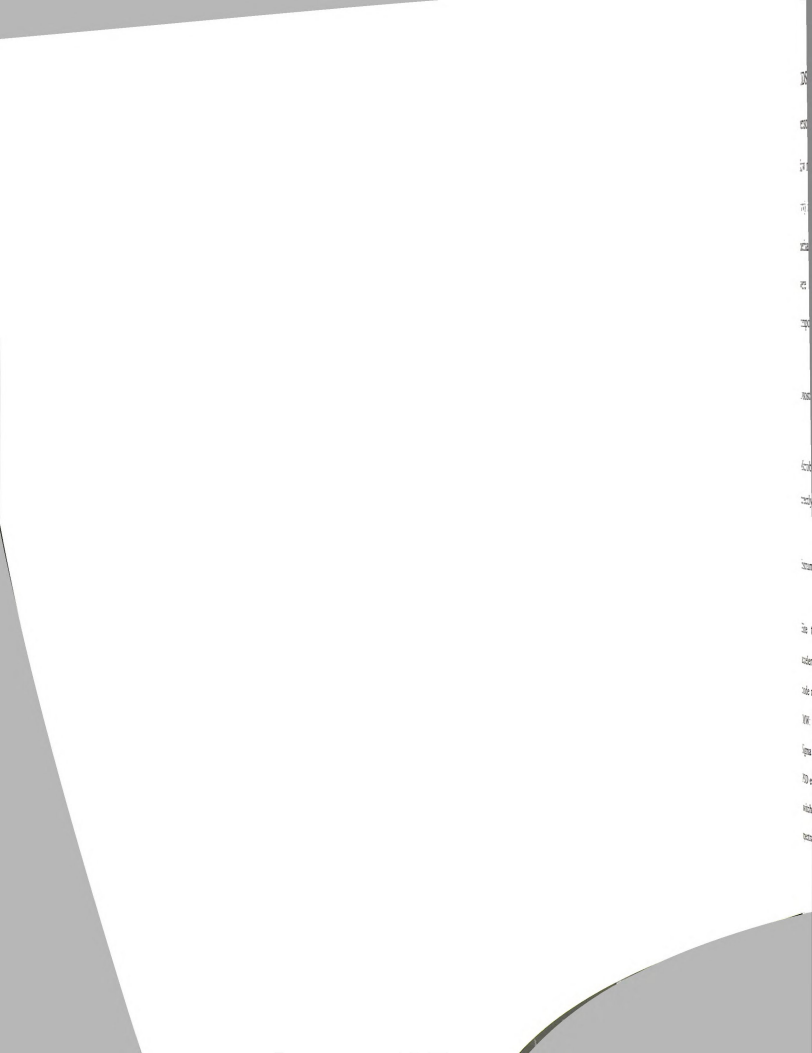
difficult to obtain even molecular weight information from trimers, and FAB-MS usually fails entirely for higher oligomers. Within the past decade, the development of new ionization techniques such as electrospray ionization (ESI) MS (11) and matrix-assisted laser desorption ionization (MALDI) MS (12-15) have made significant contributions to the study of biological molecules. An early study by Allmaier et al. (16) reported the molecular weight determination of mucopeptides at the picomole level using these new ionization techniques.

In this chapter, MALDI-MS, PSD, and specific enzymatic degradation with lysostaphin are demonstrated as an approach to characterize the mucopeptide composition of a highly methicillin-resistant *Staphylococcus aureus* strain. With these techniques, structural information from mucopeptides including peptide sequence and cross-linking patterns can be obtained with greatly improved sensitivity, relative to that of conventional mass spectrometric methods.

2. Experimental

Preparation of Mucopeptides from Peptidoglycan

Mucopeptides were obtained from a highly homogeneous methicillin-resistant *Staphylococcus aureus* strain COL as described by de Jonge *et al.* (9). Peptidoglycan was digested with muramidase (Sigma Chemical Co., St. Louis, MO) and the mucopeptides were reduced with sodium borohydride. Separation of the reduced mucopeptides by HPLC was performed essentially by the method of Glauner (22), with some modifications. Samples were applied to a 250 × 4.6-mm reversed-phase column



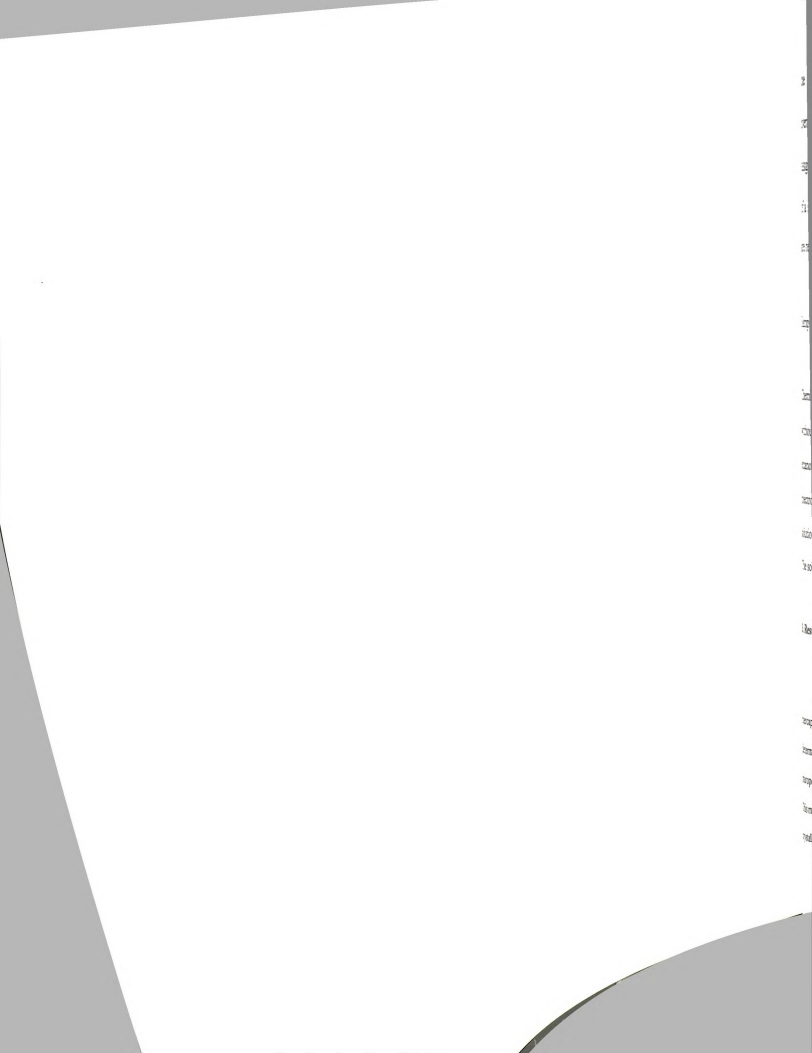
(ODS-Hypersil, 3 μ m; Keystone Scientific, Bellefonte, PA), guarded by a (70 \times 4.6-mm) perisorb RP-18 (EM Science, Elmsford, NY) precolumn. The column was eluted at a flow rate of 0.5 ml/min with a linear gradient starting immediately after injection of 5% (v/v) methanol in 100 mM NaH₂PO₄ (pH 2.5), 0.00025% azide per liter to 30% (v/v) methanol in 100 mM NaH₂PO₄ (pH 2.8) in 150 min. Collected mucopeptide fractions were desalted with acetonitrile according to Dougherty (23) and the amino acid composition was determined.

Lysostaphin Digestion of Mucopeptides for MALDI Mass Mapping

Desalted mucopeptide samples were digested with lysostaphin (5 μ g/ml, Applied Microbiology, NY) in 12.5 mM sodium phosphate buffer (pH 5.5) for 16 h and analyzed directly without further purification.

Instrumentation

MALDI-MS experiments were carried out on the PerSeptive Biosystems Voyager-Elite time-of-flight (TOF) mass spectrometer and continuous extraction with an accelerating voltage of 25 kV was used in all experiments. Linear mode and reflector mode mass spectra were calibrated with a model mucopeptide GlcNAc-MurNAc-Ala-Gln (MW: 694.7, Sigma Chemical Co., St. Louis, MO) and human gastrin I (MW: 2098.2, Sigma Chemical Co., St. Louis, MO) as standards in both positive and negative modes. In PSD experiments, precursor ion selection was performed by means of an electrostatically switched ion gate, which allows a mass selectivity of about 90 M/dM. PSD fragment ion spectra were sequentially recorded over 10-12 mass windows by incremental reduction of



the reflectron voltages. In each spectral window 30-100 single-shot spectra were averaged depending on the sample conditions. Accuracy and precision of PSD mass assignments reach ± 0.3 - 0.4 u in the mass range of m/z 300-1300. Compilation (stitching) of a set of PSD mass windows was performed by the data system. Average mass values are reported throughout this chapter.

Sample Preparation

5-Chloro-2-Mercaptobenzothiazole (CMBT) was purchased from Aldrich Chemical Co. (Milwaukee, WI) and used as the MALDI matrix in mucopeptide analysis without further purification (24). CMBT was dissolved in a mixture of ethanol/tetrahydrofuran/water (1:2:1) to make a solution of 10g/l. The samples for mass spectrometric analysis were typically prepared by depositing about 1 μ l of the matrix solution and an equal volume of the analyte solution on a sample plate well and mixed. The solvents were removed by air-drying.

3. Results and Discussion

Among many MALDI matrices employed to date, 5-chloro-2-mercaptobenzothiazole (CMBT) has been found to be most suitable for mucopeptide mass determinations (24). For peptidoglycan analysis, CMBT offers sensitive detection of mucopeptides at low picomole to femtomole range in both positive and negative modes. This matrix also exhibits excellent experimental reproducibility owing to the homogeneous crystallization of the analyte/matrix mixture over the entire sample surface area. In

contrast to the high sensitivity
samples are typically required

3. Molecular Weight Determination

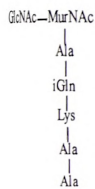
Micropeptides derived from
are shown in Figure 3.2. The
MALDI-TOF-MS with the ion
in positive and negative ion
manner. Because of the
glycosylated, the positive
[M+Na]⁺ ion signals. Proteo-
is observed with much less
signal in the negative ion
mode in the negative ion
hydrated molecule [M-H]⁻
manner. The simplicity of
molecular weight determination
In the positive and
manner shown in Figure
dominant. Some ion signals
the ion source) (25), denoted

contrast to the high sensitivity achieved by MALDI-MS, nanomole quantities of the sample are typically required in analysis by FAB-MS.

3.1. Molecular Weight Determination

Muropeptides derived from peptidoglycan of *Staphylococcus aureus* strain COL are shown in Figure 3.2. The molecular masses of these compounds were determined by MALDI-TOF-MS with the instrument operated in the linear mode. Figure 3.3 presents the positive and negative ion MALDI mass spectra of the unsubstituted muropeptide monomer. Because of the high affinity of alkali ions for the sugar moiety of the peptidoglycans, the positive ion MALDI mass spectrum (Figure 3.3(a)) is dominated by $[M+Na]^+$ ion signals. Protonated molecules $[M+H]^+$ and potassium adducts $[M+K]^+$ are also observed with much less abundance. Alternatively, muropeptides can be easily analyzed in the negative ion mode without any special sample preparation. The major peak in the negative ion MALDI mass spectrum spectrum (Figure 3.3(b)) is the deprotonated molecule $[M-H]^-$, accompanied by a $[M+Na-2H]^-$ ion peak of very low intensity. The simplicity of the negative ion MALDI mass spectra allows unambiguous molecular weight determinations of unknown samples.

In the positive and negative MALDI spectra for the pentaglycine-substituted monomer shown in Figure 3.4, $[M+Na]^+$ and $[M-H]^-$ ion signals, respectively, are predominant. Some ion signals from prompt fragmentation (i.e., fragmentation occurring in the ion source) (25), denoted as $[M-GlcNAc+Na]^+$ and $[M-GlcNAc-H]^-$, resulting from



(a)



Figure 3.2: Mucopeptides of *Staphylococcus aureus* strain 8006. (a) Mucopeptide monomer (b) Mucopeptide dimer (c) Mucopeptide trimer (d) Mucopeptide tetramer (e) Mucopeptide pentamer (f) Mucopeptide hexamer (g) Mucopeptide heptamer (h) Mucopeptide octamer (i) Mucopeptide nonamer (j) Mucopeptide decamer. Abbreviations were used: Glc, glucose; Mur, muramic acid; iGln, isoglutamine; Lys, lysine; Ala, alanine.

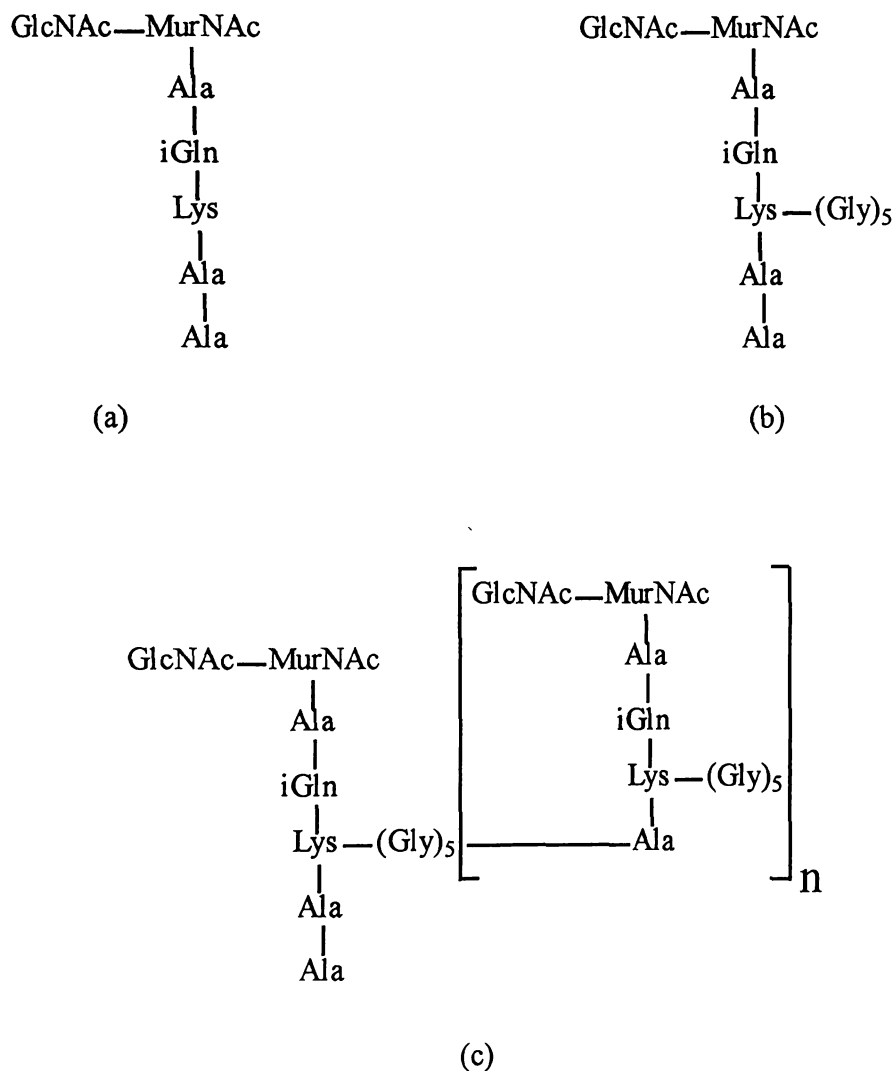


Figure 3.2: Muropeptides derived from peptidoglycan of the highly methicillin-resistant *Staphylococcus aureus* strain COL. (a) unsubstituted monomer; (b) pentaglycine-substituted monomer (c) muropeptide oligomers ($n = 1, 2, 3, \dots$). The following abbreviations were used: GlcNAc, N-acetylglucosamine; MurNAc, N-acetylmuramic acid; Ala, alanine; iGln, isoglutamine; Lys, lysine; Gly, glycine.



Fig. 3.3: Linear MALDI mass spectra of the polymer obtained in (a) positive (b) negative ion mode. The amount of analyte loaded was 100 ng.

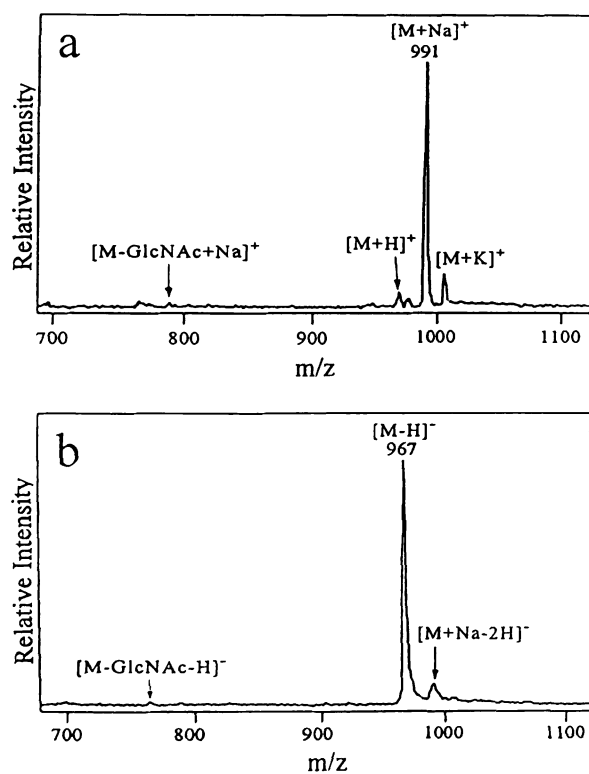


Figure 3.3: Linear MALDI mass spectra of an unsubstituted monomer obtained in (a) positive (b) negative ion mode with CMBT as matrix. The amount of analyte loaded on the sample plate was ca. 2 picomoles.

to loss of *N*-acetylglucosamine

competition, only small frag-

ments of the unsubstituted

Figure 3.4 are likely to be

to glycine chain that coelute

has previously been de-

tripeptides all elute within

to m/z values of these peaks

to and di-glycine substituted

With much lower se-

mass spectrometry (FAB-MS)

monomers and dimers. For

tripeptide oligomers large

advantage in the determination

with the positive ion MALDI

monomer. $[M+Na]^+$ ion

tripeptide oligomer. The

increases with the size of the

oligomers contain more terms

peaks are separated by 116

the loss of *N*-acetylglucosamine (GlcNAc) residue, are also observed in the spectra. In comparison, only small fragment peaks for the loss of GlcNAc were observed in the spectra of the unsubstituted monomer (Figure 3.3). Other peaks labeled with an asterisk in Figure 3.4 are likely to be related mucopeptide contaminants differing in the length of the glycine chain that coeluted in the HPLC fraction, rather than mass spectral fragments. It has previously been demonstrated that penta-, tetra-, and tri- glycine substituted mucopeptides all elute within one minute of each other under these HPLC conditions (10). The m/z values of these peaks in the spectra correspond to those expected for the tetra-, tri-, and di- glycine substituted mucopeptides.

With much lower sensitivity compared to MALDI-MS, fast atom bombardment mass spectrometry (FAB-MS) is able to determine the molecular masses of mucopeptide monomers and dimers. However, it is difficult to obtain signals by FAB-MS for mucopeptide oligomers larger than dimers. MALDI-TOF instruments have an inherent advantage in the determination of higher mass oligomers. This is illustrated in Figure 3.5 with the positive ion MALDI mass spectra of mucopeptides ranging from a dimer to a pentamer. $[M+Na]^+$ ion peaks are the dominant species in the spectrum of each mucopeptide oligomer. The intensity of the fragment ion peak, $[M-GlcNAc+Na]^+$, also increases with the size of the mucopeptides. This is not unexpected given that higher oligomers contain more terminal GlcNAc groups. Between adjacent oligomers, $[M+Na]^+$ peaks are separated by 1164 mass units, which characterizes the repeating units of the

a

Relative Intensity

b

Relative Intensity

Figure 3.4: Linear MALDI mass spectra of (a) positive and (b) negative ion mode. The spectra were obtained on the sample plate with a laser wavelength of 337 nm. The x-axis represents the mass-to-charge ratio (m/z) and the y-axis represents the relative intensity. The spectra show a series of peaks corresponding to the molecular ions of the sample, with the base peak at m/z 100. The peaks are labeled with their corresponding m/z values: 100, 117, 134, 151, 168, 185, 202, 219, 236, 253, 270, 287, 304, 321, 338, 355, 372, 389, 406, 423, 440, 457, 474, 491, 508, 525, 542, 559, 576, 593, 610, 627, 644, 661, 678, 695, 712, 729, 746, 763, 780, 797, 814, 831, 848, 865, 882, 899, 916, 933, 950, 967, 984, 1001.

100, 117, 134, 151, 168, 185, 202, 219, 236, 253, 270, 287, 304, 321, 338, 355, 372, 389, 406, 423, 440, 457, 474, 491, 508, 525, 542, 559, 576, 593, 610, 627, 644, 661, 678, 695, 712, 729, 746, 763, 780, 797, 814, 831, 848, 865, 882, 899, 916, 933, 950, 967, 984, 1001.

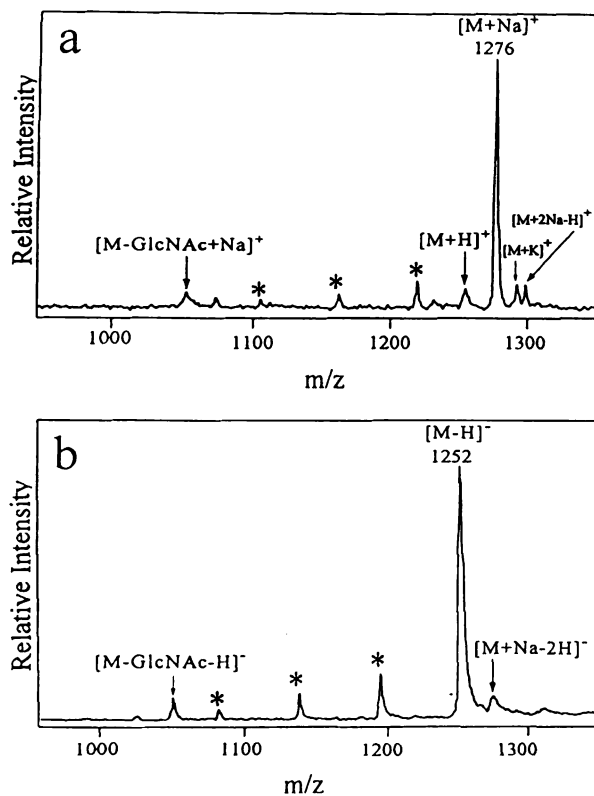


Figure 3.4: Linear MALDI mass spectra of pentaglycine-substituted monomer obtained in (a) positive and (b) negative ion mode with CMBT as matrix. The amount of analyte loaded on the sample plate was approximately 2 picomoles. Peaks labeled with an asterisk are separated by 57 mass units representing mucopeptide impurities in the sample that are tetra-, tri-, di- glycine substituted.

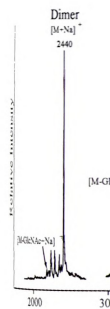


Figure 1.5: Composite positive ion mass spectrum from monomer to pentamer. The sample plate was between 1

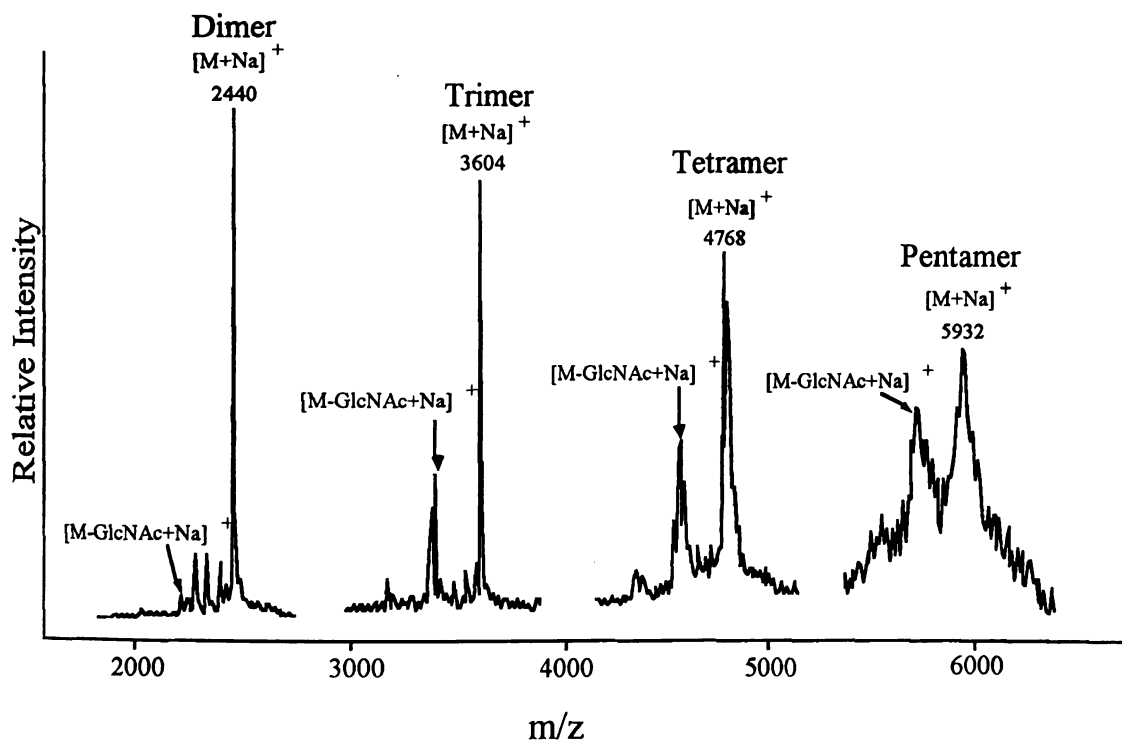


Figure 3.5: Composite positive ion linear MALDI mass spectra of muopeptides ranging from monomer to pentamer with CMBT as matrix. The amount of analyte loaded on the sample plate was between 1 to 5 picomoles.

glycan structure.

11. Analysis by MALDI-FT

MALDI-PSD fragments
oligosaccharides for structural
analysis of the analyte samples.

MALDI-PSD Analysis of an

In positive ion mode
mass spectrometer for an unsubstituted
oligosaccharide relative to the protonated molecular ion
shown in Figure 3.6, all assignments. All peaks labeled
formed before the analysis. Peaks arise by the loss of the
oligosaccharide at m/z 920, 849, and 721, and
nomenclature conventions. The mechanism of formation of the
oligosaccharide group after the analysis (26, 28-30). The
PSD spectrum. These ions provide the sequence of the

peptidoglycan structure.

3.2. Analysis by MALDI-PSD

MALDI-PSD fragment ion analysis has been performed on the monomeric mucopeptides for structural elucidation. The sensitivity was such that typically only 1-10 picomoles of the analyte sample were loaded on the sample plate for the PSD experiment.

MALDI-PSD Analysis of an Unsubstituted Monomer

In positive ion mode analysis, the $[M+Na]^+$ ion at m/z 991 was selected as the precursor for an unsubstituted mucopeptide monomer because of its high abundance relative to the protonated molecules, $[M+H]^+$. The post-source decay (PSD) spectrum is shown in Figure 3.6, along with its corresponding structure and fragment mass assignments. All peaks labeled in the spectrum are sodium-containing species. As reported before in the analysis of alkali metal-cationized peptides, some major fragment ions arise by the loss of the C-terminal amino acid residues (26). These ions, including m/z 920, 849, and 721, are designated as $[b_{n-1}+Na+OH]^+$ according to current nomenclatural conventions (26-27), but are labeled as "b+18" in the figure for simplicity. The mechanism of formation of ions of this type, involving retention of the C-terminal hydroxyl group after the rearrangement, has been described in detail by several laboratories (26, 28-30). Two $[a_n+Na-H]^+$ ions at m/z 874 and m/z 546, also appear in the PSD spectrum. These ions, along with two $[b_n+Na-H]^+$ ions at m/z 831 and m/z 703, provide the sequence of the peptide portion of the molecule. Another prominent fragment

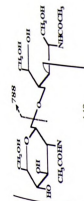
$[M+Na]^+$
991

24

CHL -

Na^+
(CH₃)₂
CHL

CONH



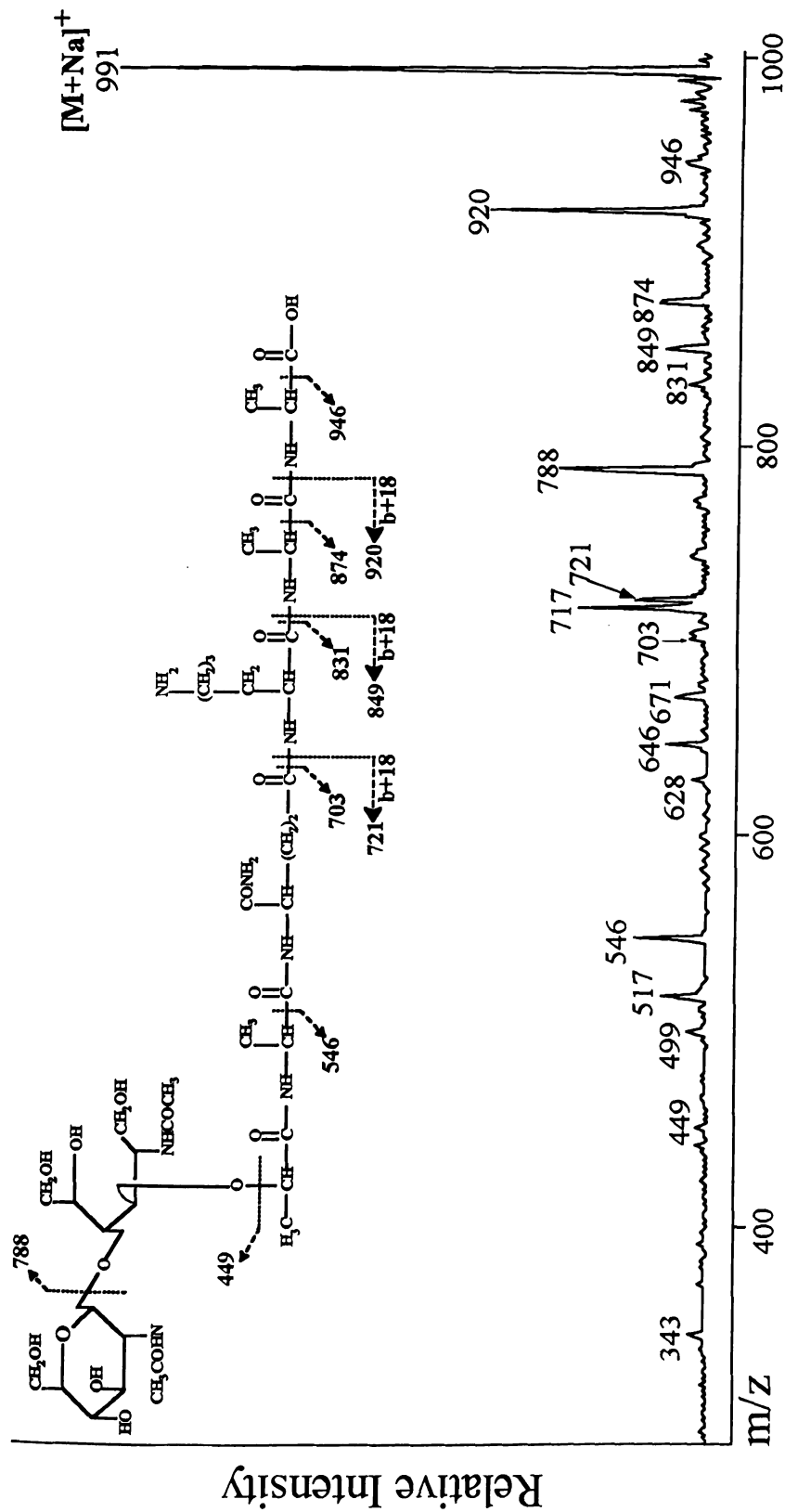


Figure 3.6: Positive ion MALDI-PSD mass spectrum of an unsubstituted monomer. Matrix: CMBT. Sample loading: 2 picomoles.

at m/z 788, resulting
the moiety in the mol
units, also accompani
ing m/z 920, 874, 84
717, 671, 646, 628, 51
saccharide of this mo
moiety.

AQD-PSD Analysis of a

The PSD spectrum

attached via the e

substituted monomer. T

25, 1134, and 721 from

lated in Figure 3.7. Se

^{16}O - ^{16}H ion at m/z 703

the fragment ion at

the neutral loss of GlcNAc

for the pentaglycine

956, 885, 517, and 49

erved along the pentagl

Structural informat

found in the negative io

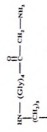
selected as the precu

ion at m/z 788, resulting from a loss of 203 mass units from $[M+Na]^+$, confirms the GlcNAc moiety in the molecule. This fragmentation process, resulting in a loss of 203 mass units, also accompanied most of the peptide sequence fragments described earlier, including m/z 920, 874, 849, 831, 721, 703, and 546, producing a related series of ions at m/z 717, 671, 646, 628, 517, 499, and 343. The fragment peak at m/z 449 corresponds to the disaccharide of this monomer, with the loss of the three carbon unit at C-3 of the muramyl moiety.

MALDI-PSD Analysis of a Pentaglycine-substituted Monomer

The PSD spectrum of the $[M+Na]^+$ ion of a monomer containing a pentaglycine moiety attached via the ϵ amino group of the lysine residue is very similar to that of unsubstituted monomer. The peaks representing $[b_{n-1}+Na+OH]^+$ ("b+18") ions at m/z 1205, 1134, and 721 from the stem peptide are also prominent in the PSD spectrum as illustrated in Figure 3.7. Several $[a_n+Na-H]^+$ ions at m/z 1159, 1088 and 546, as well as a $[b_n+Na-H]^+$ ion at m/z 703 further characterize the peptide sequence of the molecule. As before, the fragment ion at m/z 1073, originates from $[M+Na]^+$ after a loss of GlcNAc. The neutral loss of GlcNAc (-203) also occurred for most of the fragment ions described earlier for the pentaglycine-substituted monomer, resulting a series of ions at m/z 1002, 931, 956, 885, 517, and 499. It is interesting to note that no obvious fragmentation was observed along the pentaglycine chain in the positive ion PSD spectrum.

Structural information concerning the pentaglycine moiety in this muropeptide can be found in the negative ion PSD spectrum, as shown in Figure 3.8. Here, the $[M-H]^-$ ion was selected as the precursor. Most of the fragments in the negative mode PSD spectrum



$[M+Na]^+$
1276

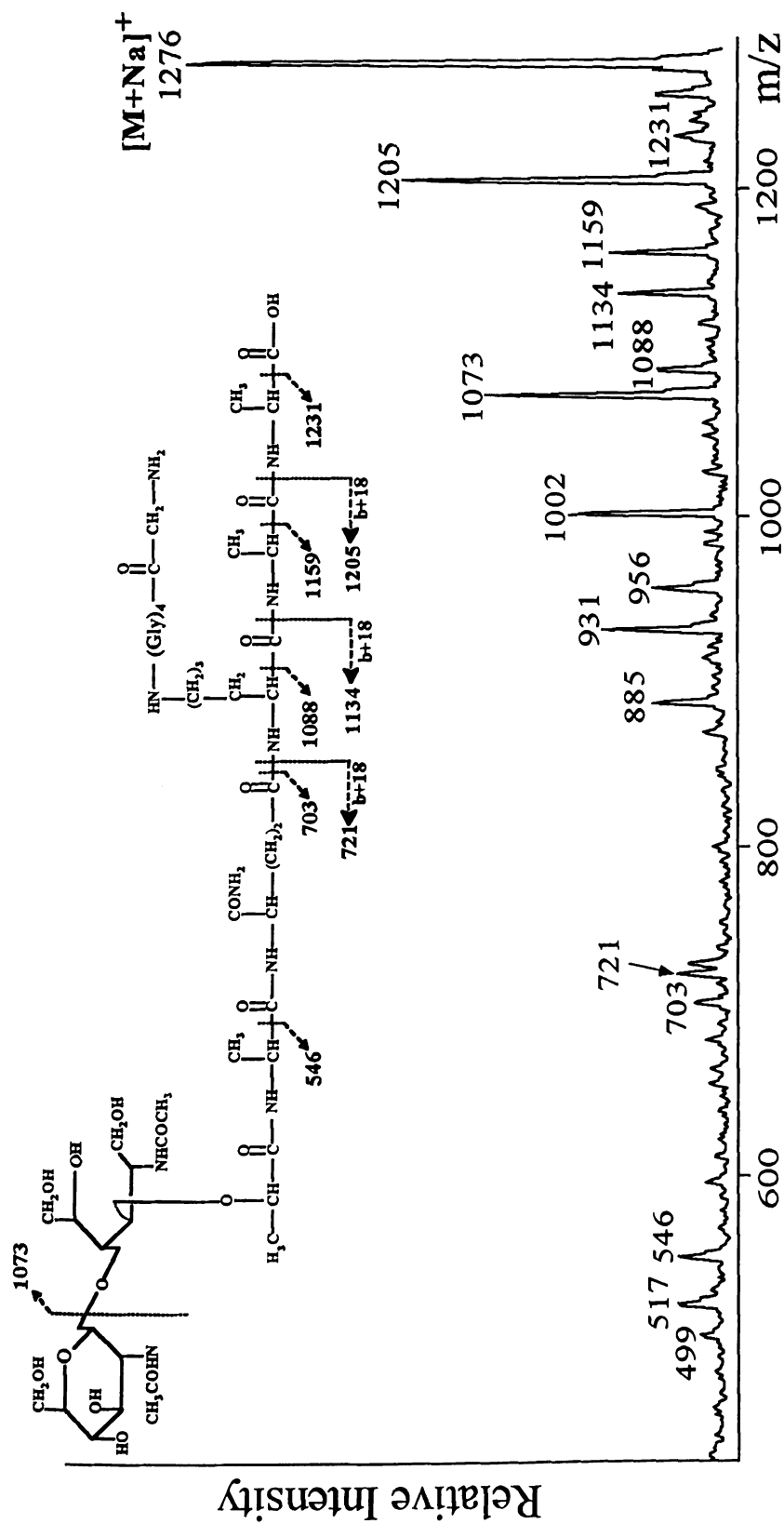
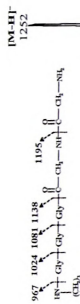
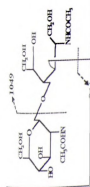


Figure 3.7: Positive ion MALDI-PSD mass spectrum of a pentaglycine-substituted monomer. Matrix: CMBT; Sample loading: 2 picomoles.



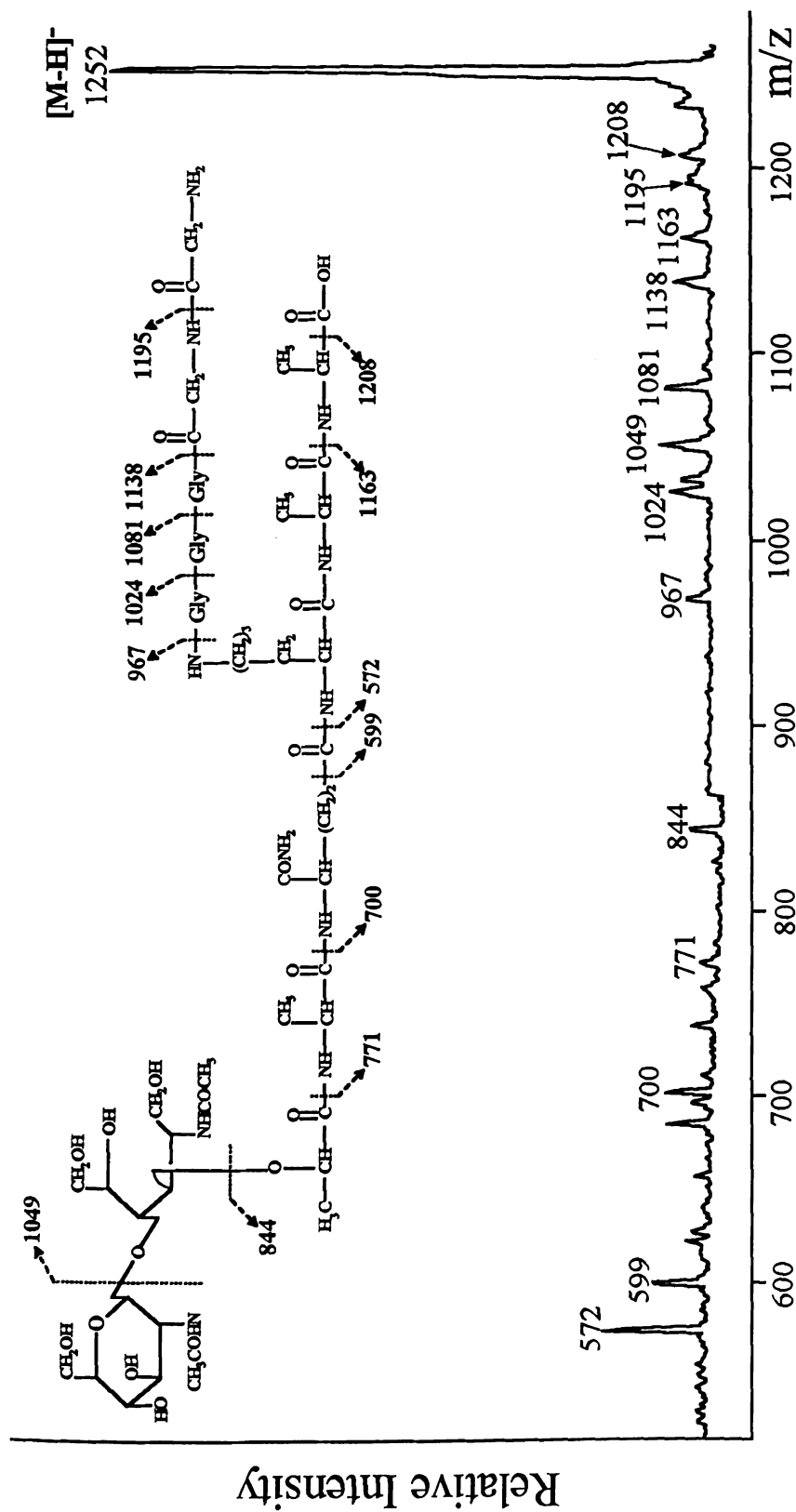


Figure 3.8: Negative ion MALDI-PSD mass spectrum of a pentaglycine-substituted monomer. Matrix: CMBT; Sample loading: 2 picomoles.

gear to be formed with
peptide chain. Losses of
to resulting fragment ion
confirming the pentaglycine
described by Bowie *et al.*
fragment ion peaks at m/z
peptide backbone. The loss
corresponds to the cleavage
of the sugar, MurNAc, with
a difference of 205 mass units
observed in the negative ion
mode. We experience this observa-

tion in our MS/MS analysis of

3.2. Lysozyme Digestion

MALDI-PSD analysis
of cross-linking information
target. We have employed
oligos to characterize oligo
Staphylococcus aureus (33)
reported to liberate N-termini
from *Staphylococcus*

appear to be formed with charge retention at the C-terminal end with respect to the glycine chain. Losses of one to five glycine residues occurred from the $[M-H]^-$ ion, and the resulting fragment ion peaks were found at m/z 1195, 1138, 1081, 1024, and 967, confirming the pentaglycine sequence. This type of fragmentation agrees well with that observed by Bowie *et al.* (31, 32). A similar fragmentation mechanism also produces fragment ion peaks at m/z 771, 700, and 572, providing the primary sequence of the peptide backbone. The loss of 203 Da from $[M-H]^-$, resulting in a fragment at m/z 1049 corresponds to the cleavage at the glycosidic bond and the loss of GlcNAc. The loss of the second sugar, MurNAc, was indicated by a fragment ion peak at m/z 844, representing a difference of 205 mass units from the peak at m/z 1049. The extent of fragmentation was diminished in the negative ion mode PSD analysis, relative to that in the positive mode. In our experience this observation is typical for other peptides and muropeptides in negative ion mode MS/MS analysis (including PSD).

3.3. Lysostaphin Digestion and Mass Mapping Analysis

MALDI-PSD analysis of higher oligomers does not provide complete sequence and cross-linking information. These spectra can also be rather complex and difficult to interpret. We have employed a specific protease, lysostaphin, combined with MALDI-MS analyses to characterize oligomeric muropeptides. Lysostaphin is an antibiotic that lyses *Staphylococcus aureus* (33). The active principal is enzymic in nature and has been reported to liberate N-terminal glycine and alanine (but not C-terminal alanine in the stem peptide) from *Staphylococcus aureus* cell wall (34). It was also observed that

lysine cross-linked

of this unique property, ly

alysis, and was found t

oligomeric isomers.

Two mucopeptide

ature for each was su

ification. The first mu

Staphylococcus aureus str.

both positive and neg

spectrum of the lysostaphin

the spectrum correspond

to cleavage at different

knowledge, this is the first

to specificity of lysostaphin

olysis, based on the int

Figure 3.9(b)) are consister

ities and observed values

10. The most favorable hy

ones represented by pro

characterizing both sides o

chemical composition infor

lysine moiety with a f

pentaglycine cross-linked mucopeptides can be hydrolyzed by this enzyme (34). Because of this unique property, lysostaphin was used in the staphylococcal mucopeptide oligomer analysis, and was found to be particularly useful for characterizing glycine cross-linked oligomeric isomers.

Two mucopeptide dimers were subjected to lysostaphin digestion and the reaction mixture for each was subjected directly to analysis by MALDI-MS without further purification. The first mucopeptide dimer, I, was from the highly methicillin-resistant *Staphylococcus aureus* strain COL. The molecular weight was determined to be 2417 Da by both positive and negative mode MALDI-MS. The negative ion mode MALDI spectrum of the lysostaphin digest of mucopeptide I is shown in Figure 3.9(a). The peaks in the spectrum correspond to the product of hydrolysis, *i.e.*, smaller mucopeptides, after the cleavage at different sites between the Gly-Gly bonds of the dimer. To our knowledge, this is the first time that direct mass spectral evidence has been obtained for the specificity of lysostaphin cleavage of mucopeptide oligomers. The favorable sites of hydrolysis, based on the intensities of peaks in the MALDI mass spectrum of the digest (Figure 3.9(b)) are consistent with those predicted from previous studies (34). Calculated values and observed values of $[M-H]^-$ for resulting mucopeptides are summarized in Table 3.1. The most favorable hydrolysis is at site A, as indicated in Figure 3.9(b), producing two ions represented by prominent peaks at m/z 1195 and m/z 1238 in the spectrum, characterizing both sides of the pentaglycine bridge. Other peaks also provide the chemical composition information of the cross-linking pentaglycine bridge and the pentaglycine moiety with a free N-terminal, attached to one of the monomer units of the

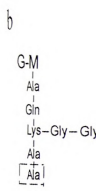


Figure 3.9: (a) Negative-ion mass spectrum of (b) mucopeptide. The mucopeptide was used directly for the analysis. The arrows in Fig. 3.9 indicate the positions of the peaks. The intensities of peak 1067 and 1068 are the highest. Thicker arrows indicate the positions of the peaks.

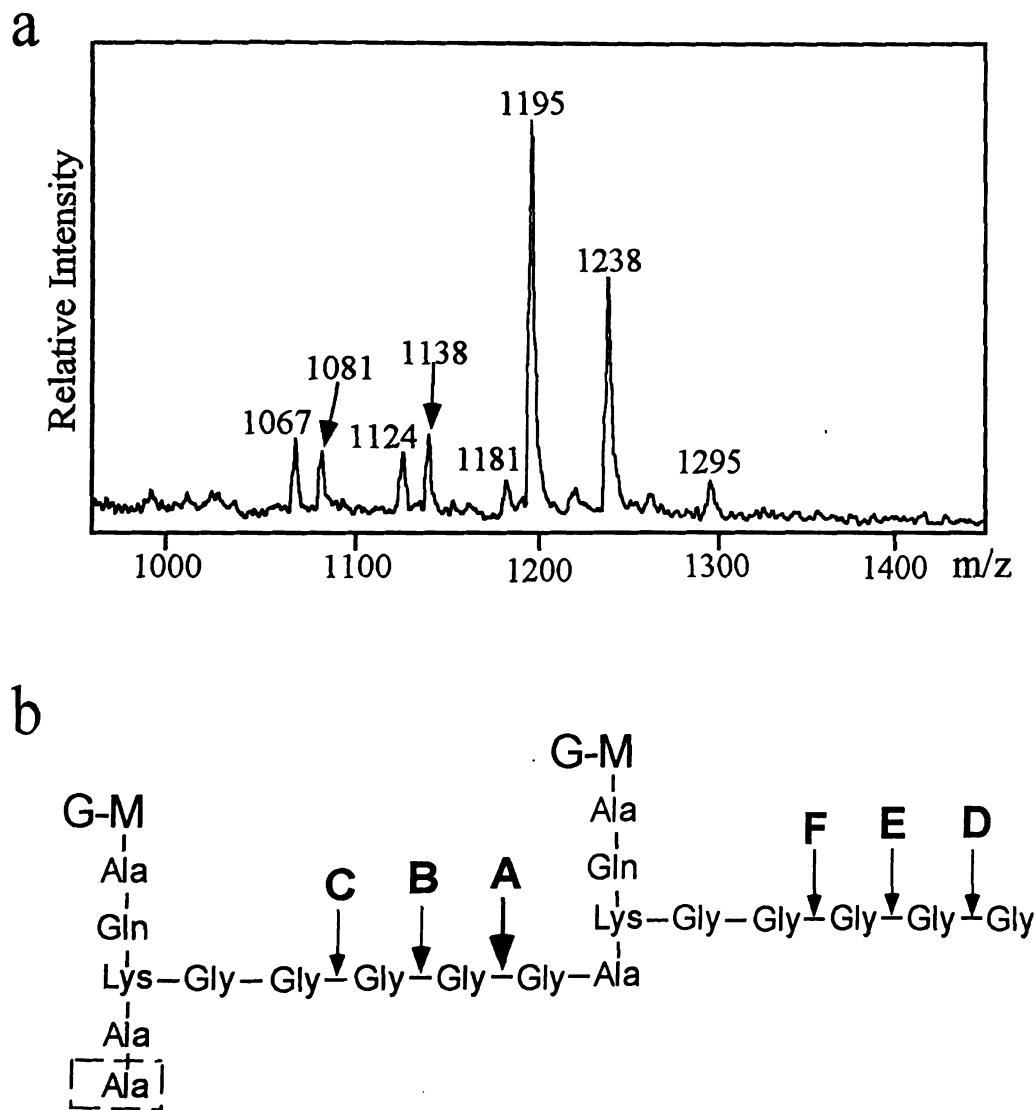


Figure 3.9: (a) Negative ion MALDI mass spectrum of the lysostaphin digest of (b) mucopeptide dimer I. Matrix: CMBT. The digestion mixture was used directly for MALDI-MS analysis without further purification. The arrows in Fig. 3.9b indicate the favorable sites of hydrolysis, based on the intensities of peaks in the MALDI mass spectrum of the lysostaphin digest. Thicker arrows indicate more favorable hydrolysis sites.

muropeptide. The absence
to conclusion that lysos-
second glycine residue of

When this strain was
phone, the muropeptide
compared with muropeptid-

by MALDI-MS. Infor-
muropeptide led to the d-
spectrum of the lysostaphin

to 1181 and m/z 1238
fragments obtained after
1106)). The masses of

shown on the left side of
hydrolysis peaks, as sum-

in conjunction with PSD
including other types of s-

muropeptide oligomers from

muropeptide. The absence of peaks with a molecular weight less than 1024 Da supports the conclusion that lysostaphin is unable to hydrolyze the bond between the first and second glycine residue of the cross-linking bridge.

When this strain was grown in the presence of high concentrations of exogenous glycine, the muropeptide composition of peptidoglycan changed (35). The dimer I was replaced with muropeptide dimer II having a molecular weight of 2403 Da as determined by MALDI-MS. Information obtained from the lysostaphin hydrolysate of this muropeptide led to the determination of its structure. The negative ion MALDI mass spectrum of the lysostaphin digest is shown in Figure 3.10(a). Two prominent peaks at m/z 1181 and m/z 1238 were detected in the MALDI-MS spectrum, representing the fragments obtained after cleavage at the most favorable hydrolysis site (site A, Figure 3.10(b)). The masses of these peaks indicate that amino acid substitution (Ala \rightarrow Gly) occurred on the left side of the muropeptide dimer as illustrated in Figure 3.10(b). Other hydrolysate peaks, as summarized in Table 3.2, also verified this amino acid substitution. In conjunction with PSD analysis, lysostaphin digestion may be used in the future for probing other types of structural modifications occurring in the glycine cross-linked muropeptide oligomers from *Staphylococcus aureus*.

Table 3.1: MALDI Mass Mapping Analysis of Lysostaphin Digest of Muropeptide Dimer I

cleavage site	Resulting muropeptide	$[M-H]^-$ (calc)	$[M-H]^-$ (obs)
---------------	-----------------------	------------------	-----------------

Table 3.1: MALDI Mass Mapping Analysis of Lysostaphin Digest of Muropeptide Dimer I

cleavage site	Resulting muropeptide	[M-H] ⁻ (calc)	[M-H] ⁻ (obs)
A	Left of A site	1195	1195
	Right of A site	1238	1238
B	Left of B site	1138	1138
	Right of B site	1295	1295
C	Left of C site	1081	1081
	Right of C site	1352	1352
A and D	Between A and D	1181	1181
A and E	Between A and E	1124	1124
A and F	Between A and F	1067	1067



b

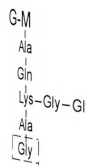


Figure 3.10: (a) Negative ion mass spectrum of (b) mucopeptide

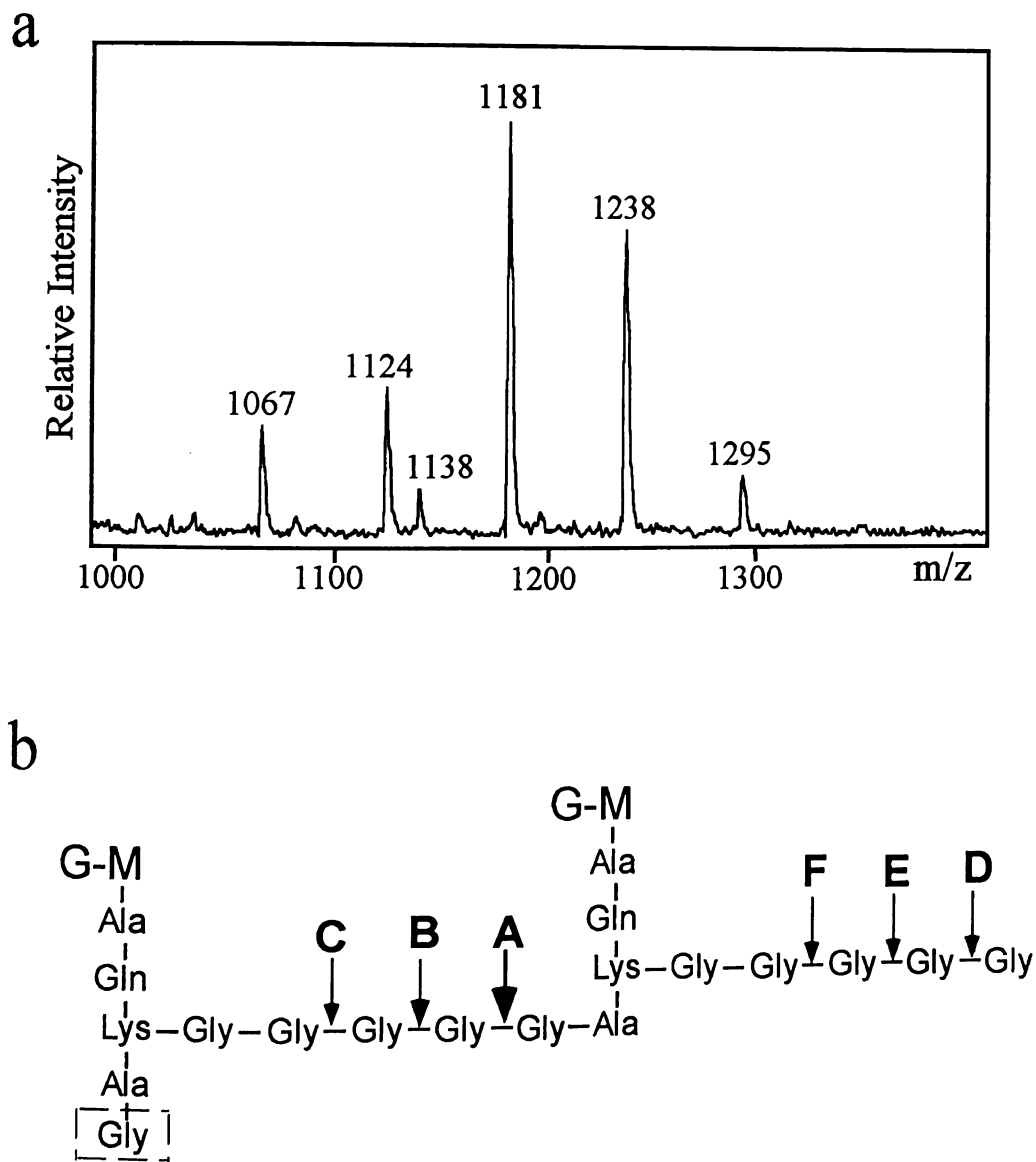


Figure 3.10: (a) Negative ion MALDI mass spectrum of the lysostaphin digest of (b) mucopeptide dimer II. Matrix: CMBT.

Table 3.2: MALDI Mass Mapping Analysis of Lysostaphin Digest of Mucopeptide Dimer II

cleavage site	Resulting mucopeptide	[M-H] ⁺ (calc)	[M-H] ⁺ (obs)
---------------	-----------------------	---------------------------	--------------------------

Table 3.2: MALDI Mass Mapping Analysis of Lysostaphin Digest of Muropeptide Dimer II

cleavage site	Resulting muropeptide	$[M-H]^-$ (calc)	$[M-H]^-$ (obs)
A	Left of A site	1181	1181
	Right of A site	1238	1238
B	Left of B site	1124	1124
	Right of B site	1295	1295
C	Left of C site	1067	1067
	Right of C site	1352	n.d.
A and D	Between A and D	1181	1181
A and E	Between A and E	1124	1124
A and F	Between A and F	1067	1067

Note. n.d.: not detected

Conclusion

In conclusion, the
described for micropeptide
matrix for MALDI-MS.
analysis after lysostaphin
monomers and oligomers
cross-linking patterns.
promising for the study of
bacterial species, although
Staphylococcus aureus
understandings of peptidomimetic

4. Conclusion

In conclusion, this chapter has demonstrated that sensitive detection can be achieved for muropeptides in both positive mode and negative mode with CMBT as matrix for MALDI-MS. Post-source decay (PSD) analyses, together with mass mapping analyses after lysostaphin digestion, reveal structural information on muropeptide monomers and oligomers including peptide sequence, oligosaccharide composition, and cross-linking patterns. The combination of MALDI-MS and PSD analysis is also promising for the study of other structural modifications of peptidoglycans from different bacterial species, although the utility of lysostaphin digestion is restricted only to *Staphylococcus aureus* muropeptides. These approaches can provide better understandings of peptidoglycan metabolism and clarify resistance mechanisms.

References:

- Schäfer, K. H., and K.
- Garsen, J. M. (197
- Ess.), vol 4, pp. 463-
- Temsend, D. E., Asl
- Grob, W. B., (1987)
- Nitzberg, D. R. (1994
- Back-Sague, C. M., C
- Acta Control Hosp.*
- Martin, S. A., Biemar
- Peptidoglycans* (Seidl
- Berlin-New York.
- Martin, S. A., Rosent
- 5326
- Pittenauer E., Allmaier
- de Pedro M. A., Hölt
- York.
- de Jonge, B. L. M., C
- 267, 11248-11254.
- de Jonge, B. L. M., C
- 267, 11255-11259.
- Mann, M., Shen, S., a
- Sciences (Gross, M. L.

References:

1. Schleifer, K. H., and Kandler, O. (1972) *Bacteriol. Rev.* **36**, 407-477.
2. Ghuysen, J. M. (1977) in *Cell Surface Reviews* (Poste, G., and Nicholson, G. L., Eds.), vol. 4, pp. 463-595, Elsevier/North Holland Publishing Co., Amsterdam.
3. Townsend, D. E., Ashdown, N., Bolton, J., Duckworth, G., Moorhouse, E. C., and Grubb, W. B., (1987) *J. Hosp. Infect.* **9**, 60-71.
4. Schaberg, D. R. (1994) *Ann. Emerg. Med.* **24**, 462-464.
5. Beck-Sague, C. M., Chong, W. H., Roy, C., Anderson, R., and Jarvis, W. R. (1992) *Infect. Control Hosp. Epidemiol.* **13**, 526-534.
6. Martin, S. A., Biemann, K., and Rosenthal, R. S. (1986) in *Biological Properties of Peptidoglycans* (Seidl, P. H., and Schleifer, K. H., Eds.), pp. 49-54, Water de Gruyter, Berlin-New York.
7. Martin, S. A., Rosenthal, R. S., and Biemann, K. (1987) *J. Biol. Chem.* **262**, 7514-7526.
8. Pittenauer E., Allmaier G., and Schmid, E. R. (1993) in *Bacterial Growth and Lysis* (de Pedro M. A., Höltje J. -V., Löffelhardt, W., Eds.), pp. 39-46, Plenum Press, New York.
9. de Jonge, B. L. M., Chang, Y. -S., Gage D. A., and Tomasz, A. (1992) *J. Biol. Chem.* **267**, 11248-11254.
10. de Jonge, B. L. M., Chang, Y. -S., Gage D. A., and Tomasz, A. (1992) *J. Biol. Chem.* **267**, 11255-11259.
11. Mann, M., Shen, S., and Fenn, J. B. (1992) in *Mass Spectrometry in the Biological Sciences* (Gross, M. L., Ed.), pp. 145-163, Kluwer, Dordrecht.

Karas, M., and Hillen

Seavis, R. C., and Ch

Kuhl, B., Steup, M.,

1996

Feng, K., Allman, S.

1996

Kimzer G., and Schn

Häbe J.-V., Löffelhar

Kaufmann, R., Speng

Spectrom. 7, 902-910

Kaufmann, R., Kirsch

Processes 131, 355-38

Huberty, M. C., Vath

1999

Spengler, B., Kirsch,

Rouse, J. C., Yu, W.

335.

Glauner, B. (1988) A

Garcia-Bustos, J. F.,

178-182.

Xu N., Huang Z. 1

Spectrom. 8, 116-1

12. Karas, M., and Hillenkamp F. (1988) *Anal Chem.* **60**, 2299-2301.
13. Beavis, R. C., and Chait, B. T. (1990) *Proc. Natl. Acad. Sci.* **87**, 6873-6878.
14. Stahl, B., Steup, M., Karas, M., and Hillenkamp, F. (1991) *Anal. Chem.* **63**, 1463-1466.
15. Tang, K., Allman, S. L., Chen, C. H. (1992) *Rapid Commun. Mass Spectrom.* **6**, 365-368.
16. Allmaier G., and Schmid, E. R. (1993) in *Bacterial Growth and Lysis* (de Pedro M. A., Höltje J.-V., Löffelhardt, W., Eds.), pp. 23-30, Plenum Press, New York.
17. Kaufmann, R., Spengler, B., and Lützenkirchen, F. (1993) *Rapid Commun. Mass Spectrom.* **7**, 902-910.
18. Kaufmann, R, Kirsch, D., and Spengler, B. (1994) *Int. J. Mass Spectrom. Ion Processes* **131**, 355-385.
19. Huberty, M. C., Vath, J. E., Yu, W., and Martin, S. A. (1993) *Anal. Chem.* **65**, 2791-2799.
20. Spengler, B., Kirsch, D., and Kaufmann, R. (1995) *J. Mass Spectrom.* **30**, 782-787.
21. Rouse, J. C., Yu, W., and Martin, S. A. (1995) *J. Am. Soc. Mass Spectrom.* **6**, 822-835.
22. Glauner, B. (1988) *Anal. Biochem.* **172**, 451-464.
23. Garcia-Bustos, J. F., and Dougherty, T. J. (1987) *Antimicrob. Agents Chemother.* **31**, 178-182.
24. Xu N., Huang Z. H., Watson J. T., and Gage D. A., (1997), *J. Am. Soc. Mass Spectrom.* **8**, 116-124.

15 Scott, P. D. and Viswan

16 Giese, R. P., Cerny,

342

17 Bemann, K. (1988) *B*

18 Tang, X., Ens, W., S

1791-1799.

19 Thorne, G. C., Ballard

249-257.

20 Teesch, L. M., Orlan

3675.

21 Bradford, A. M., Wa

Spectrom. 9, 677-685.

22 Waugh R. L., and Bow

23 Schindler, C. A., and S

24 Browder, H. P., Zygo

Biochem. Biophys. Res

25 de Jonge, B. L. M., Ch

Chemother. 35, 124-12

5. Scott, P. D. and Viswanatham, K., (1994), *Anal. Chem.* **66**, 3727-3733.
6. Grese, R. P., Cerny, R. L., and Gross, M. L. (1989) *J. Am. Chem. Soc.* **111**, 2835-2842.
7. Biemann, K. (1988) *Biomed. Environ. Mass Spectrom.* **16**, 99-111.
8. Tang, X., Ens, W., Standing, K. G., and Westmore, J. B. (1988) *Anal. Chem.* **60**, 1791-1799.
9. Throne, G. C., Ballard, K. D., and Gaskell, S. J. (1990) *J. Am. Soc. Mass Spectrom.* **1**, 249-257.
10. Teesch, L. M., Orlando R. C., and Adams, J. (1991) *J. Am. Chem. Soc.* **113**, 3668-3675.
11. Bradford, A. M., Waugh, R. J., and Bowie, J. H. (1995) *Rapid Commun. Mass Spectrom.* **9**, 677-685.
12. Waugh R. L., and Bowie, J. H. (1994) *Rapid Commun. Mass Spectrom.* **8**, 169-173.
13. Schindler, C. A., and Schuhardt, V. T., (1964) *Proc. Natl. Acad. Sci.* **51**, 414-419.
14. Browder, H. P., Zygmunt, W. A., Young, J. R., and Tavormina, P. A. (1965) *Biochem. Biophys. Res. Commun.* **19**, 2-8.
15. de Jonge, B. L. M., Chang, Y. -S., Xu N., and Gage D. A. (1996) *Antimicrob. Agents Chemother.* **35**, 124-129.

Introduction

Glycoproteins

Research of the
are, in fact, g
carbohydrates (1-2). Glyc
They occur in a
of protein activi
and structural
proteins, are synthesized
enzymatically generated
of nucleic acid to
incomplete leading
glycoproteins therefore h
multiple carbohydrate va
heterogeneity or he
glycoproteins.

Carbohydrates fo
the chain (Asn- or N-lin
(linked). Sequence a
about the attachments o

Chapter 5. Structural Characterization of Glycoproteins by MALDI-MS

1. Introduction

1.1 Glycoproteins

Research of the past 20 to 30 years has demonstrated that most eukaryotic proteins are, in fact, glycoproteins; that is, they are covalently associated with carbohydrates (1-2). Glycoproteins vary in carbohydrate content from $< 1\%$ to $> 90\%$ by weight. They occur in all forms of eukaryotes and have functions that span the entire spectrum of protein activities, including those of enzymes, transport proteins, receptors, hormones, and structural proteins. The polypeptide chains of glycoproteins, like those of all proteins, are synthesized under genetic control. Their carbohydrate chains, in contrast, are enzymatically generated and covalently linked to the polypeptide without the rigid guidance of nucleic acid templates. The processing of proteins by glycosyltransferases is often incomplete leading to the synthesis of non-uniform glycoprotein products. Glycoproteins therefore have variable carbohydrate compositions, which is reflected by the multiple carbohydrate variants at a single glycosylation site. This phenomenon, known as microheterogeneity or heterogeneity, has complicated purification and characterization of glycoproteins.

Carbohydrates form two types of linkage to proteins: (1) through an asparagine side chain (Asn- or *N*-linked); (2) through the side chain of serine (Ser) or threonine (Thr) (*O*-linked). Sequence analyses of glycoproteins have led to the following generalizations about the attachments of carbohydrates in glycoproteins.

In *N*-glycosidic (
 usually β linked to the :
 can in the sequence Asn
 usually proline (Pro) or a
 these linkages usually
 linked to either mannose
 may, in turn, be linked to
 linked oligosaccharides is

(a)

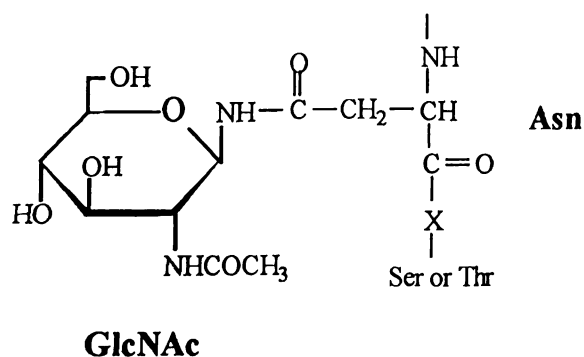
H

(b)

Figure 4.1: *N*-linked ol
 through a β -*N*-acetylglu

In *N*-glycosidic (*N*-linked) attachments, an *N*-acetylglucosamine (GlcNAc) is invariably β linked to the amide nitrogen of an asparagine (Asn) residue of the polypeptide chain in the sequence Asn-X-Ser or Asn-X-Thr, where X is any amino acid residue except possibly proline (Pro) or aspartic acid (Asp) as shown in Figure 4.1. The oligosaccharides in these linkages usually have a distinctive core whose peripheral mannose residues are linked to either mannose (Man) or *N*-acetylglucosamine residues. These latter residues may, in turn, be linked to yet other sugar residues so that an enormous diversity of *N*-linked oligosaccharides is possible.

(a)



(b)

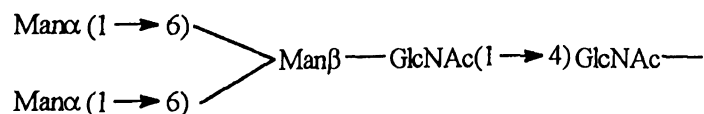


Figure 4.1: *N*-linked oligosaccharides. (a) All *N*-glycosidic protein attachments occur through a β -*N*-acetylglucosamino-Asn bond in which the Asn occurs in the sequence Asn-

Fig. 1 (b) N-linked

structure.

Structures of N-li

mannose-type, comp

Man α 1 \rightarrow 3(Man

branches (Figure

α -mannose res

disaccharides contain

type structures ha

disaccharides, and mos

4 to the β -linked man

The most commo

β -galactosyl-(1-3)-

sine or threonine as sh

ose form α -O-glycos

ends to the 5-hydroxyl

ditional generalization

They vary in size from

disaccharide units in pr

X-Ser/Thr. (b) *N*-linked oligosaccharides usually have the branched $(\text{Man})_3(\text{GlcNAc})_2$ core structure.

Structures of *N*-linked oligosaccharides fall into three main categories termed high-mannose-type, complex-type, and hybrid-type. They share the common core structure $\text{Man}\alpha 1 \rightarrow 3(\text{Man}\alpha 1 \rightarrow 6)\text{Man}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow 4\text{GlcNAc-Asn}$, but differ in their outer branches (Figure 4.2). High-mannose-type oligosaccharides have two to six additional α -mannose residues attached to the core structure. Typical complex-type oligosaccharides contain two to four outer branches with a sialyllactosamine sequence. Hybrid-type structures have the features of both high-mannose-type and complex-type oligosaccharides, and most of them contain bisecting *N*-acetylglucosamine, which is linked $\beta 1-4$ to the β -linked mannose residue of the core structure (Figure 4.2).

The most common *O*-glycosidic (*O*-linked) attachment involves the disaccharide core β -galactosyl-(1-3)- α -*N*-acetylgalactosamine α linked to the OH group of either serine or threonine as shown in Figure 4.3. Less commonly, galactose (Gal), mannose, or xylose form α -*O*-glycosides with serine or threonine. Galactose also forms *O*-glycosidic bonds to the 5-hydroxylysyl residues of collagen. However, there seem to be few, if any, additional generalizations that can be made about *O*-glycosidically linked oligosaccharides. They vary in size from a single galactose residue in collagen to the chains of up to 1000 disaccharide units in proteoglycans.

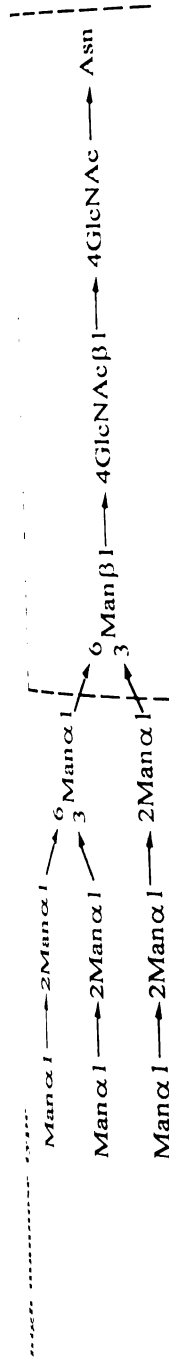




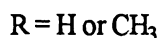
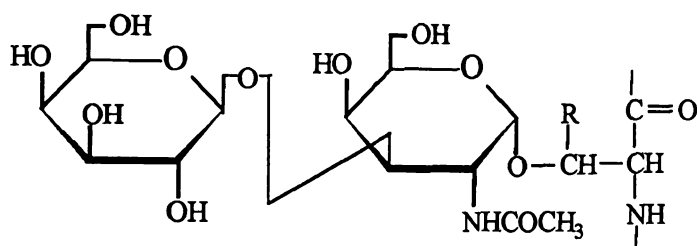
Figure 4.2: Structures of major types of *N*-linked oligosaccharides. The box area encloses the core structure common to all *N*-linked oligosaccharides.



β -G

Figure 4.3: Most common

The carbohydrate
 having integral rol
 carbohydrates have been
 biological activity, recog
 distribution in tissues (1-
 properties of gl
 collatory half-life, res
 increasing scrutiny with
 molecules for therapeuti
 roles of O-linked carboh
 glycosylation is necessar
 such as erythropoietin a
 human granulocyte-macr
 For production
 protein, the specific cel



β -Galactosyl-(1 \rightarrow 3)- α -*N*-acetylgalactosyl Ser/Thr

Figure 4.3: Most common *O*-glycosidic attachments of oligosaccharides to glycoproteins.

The carbohydrate chains of glycoproteins have begun to be widely acknowledged as having integral roles in the functional properties of glycoproteins. *N*-linked carbohydrates have been implicated in a wide variety of functions including modulation of biological activity, recognition between cells, interaction between host and pathogen, and distribution in tissues (1-4). The profound effects that carbohydrate moieties can have on other properties of glycoproteins such as solubility, antigenicity, immunogenicity, circulatory half-life, resistance to proteolysis, and thermal stability have come under increasing scrutiny with the desire of the biotechnology industry to produce these molecules for therapeutic use in humans (4-5). Although the biological and functional roles of *O*-linked carbohydrates are less well understood, it has been reported that *O*-glycosylation is necessary for efficient biosynthesis and secretion of certain glycoproteins such as erythropoietin and for prevention of antibody-based clearance of recombinant human granulocyte-macrophage colony-stimulating factor (1).

For production of glycoproteins in recombinant systems, the structure of the protein, the specific cell type employed, and the bioprocess environment all influence

and glycosylation sites
site. Glycoprotein
removal of specific oligo
structural classes from
glycoforms, as well as the
and appreciation of their

12 Glycoprotein analysis

In recent years, un
has been of increasing im
biological functions and
fundamental to this und
glycosylation, determine
the carbohydrate moieties
and the roles of the carb
This is partly caused by la
in small quantities and
carbohydrates, such as N
require large amounts of
a mixture of carbohydr
is currently being
chromatography with pu
matrix-assisted la

which glycosylation sites are utilized and the range of carbohydrate structures found at each site. Glycoprotein bioactivity or clearance can be affected by modification or removal of specific oligosaccharide structural classes from specific oligosaccharide structural classes from specific attachment sites. Thus the precise structures of the glycoforms, as well as their protein sequence context, must be understood in order to gain a full appreciation of their structure/function relationship.

1.2 Glycoprotein analysis

In recent years, understanding and elucidating the role of glycosylation of proteins has been of increasing interest due to this posttranslational modification's influence over biological functions and implications on the therapeutic usage of recombinant proteins. Fundamental to this understanding is the necessity to recognize the presence of protein glycosylation, determine the site(s) of protein glycosylation, and elucidate the structure of the carbohydrate moieties. In many proteins, however, the carbohydrate heterogeneity and the roles of the carbohydrates are diverse, complex and not very well understood. This is partly caused by lack of information because many glycoproteins are available only in small quantities and because traditional methods for structural analysis of carbohydrates, such as NMR and fast atom bombardment (FAB) mass spectrometry (MS), require large amounts of the glycosylated protein and mainly show the major components in a mixture of carbohydrates (6). More sensitive methods for analysis of carbohydrates are currently being developed, including high-performance anion-exchange chromatography with pulsed amperometric detection (7), capillary zone electrophoresis (8), matrix-assisted laser desorption ionization (MALDI) MS (9-12), liquid

chromatography/electrospray

is reagent array analysis

MALDI-MS has

composition determinat

proprotein (9-12). MA

propeptides which are

methods. With knowl

sequences, glycosylation

mass obtained from pro

for microheterogeneity

Information about

gated from mass differ

2 exoglycosidase). For

value. MALDI-MS has

structures of specific exog

exoglycosidases sequenti

was first shown by Mizu

has further developed

exoglycosidase mixtures

for a set of end product

chromatography to obtain

original structure. This ap

mounting methods (ext

chromatography/electrospray ionization (ESI) MS (13-16) and enzymatic methods such as the reagent array analysis method (17).

MALDI-MS has been gaining increasing acceptance for the detection, localization, composition determination, and sequence analysis of carbohydrates present in a glycoprotein (9-12). MALDI-MS experiments are typically performed on HPLC separated glycopeptides which are prepared from glycoproteins by proteolytic digestions or chemical methods. With knowledge of a protein's primary structure available from DNA sequences, glycosylation modifications can often be inferred on the basis of anomalous masses obtained from proteolytic peptides. Glycopeptide fractions also can be identified by their microheterogeneity as revealed in MALDI-MS spectra.

Information about the type of modifications present and their compositions can be gained from mass differences after digestion with a specific glycosidase (endoglycosidase or exoglycosidase). For example, a mass shift of 291 Da corresponds to a sialic acid residue. MALDI-MS has been exploited to monitor the degradation of glycopeptides by mixtures of specific exoglycosidases (9-10, 12, 18). The concept of using highly specific exoglycosidases sequentially to remove monosaccharide residues from oligosaccharides was first shown by Mizuochi et al. (19). A related non-mass spectrometric approach has been further developed into the reagent-array analysis method whereby a set of exoglycosidase mixtures are incubated with a chemically radiolabeled oligosaccharide to give a set of end products (17). These are pooled and analyzed by Bio-Gel P4 column chromatography to obtain a specific fingerprint which can lead to identification of the original structure. This approach is dependent on isolation of oligosaccharides using time-consuming methods (extensive dialysis, hydrazinolysis, radiolabeling, and a range of

chromatographic and elec

ic protein moiety.

oligosaccharides on the b

hours per run). Furt

ly with oligosacc

lary mixtures. In con

resorve and quick m

combination of exoglyc

structural details of olig

(3), recombinant human

of metalloproteinases (18

12 Fractionation and

immobilized column chr

The use of lectin i

or only for the purificati

for separating and de

olypeptides (20-22). L

agglutinate cells and/or

fractionation techniques s

and normal-phase chro

carbohydrate-binding sp

chromatography, the mi

chromatographic and electrophoretic purification steps) which also result in destruction of the protein moiety. In addition Bio-Gel P4 chromatography, which separates oligosaccharides on the basis of hydrodynamic volume, is a relatively slow technique (6 to 24 hours per run). Furthermore, as it is a low-resolution approach, it is possible work only with pure oligosaccharides and interpretation becomes difficult even when there are binary mixtures. In contrast, analogous MALDI-MS based procedures provide a more sensitive and quick means for detection of the degradation products (18). The combination of exoglycosidase digestion and MALDI-MS has been used to obtain structural details of oligosaccharides derived from glycoproteins such as bovine fetuin (10), recombinant human macrophage colony stimulating factor (12), and tissue inhibitor of metalloproteinases (18).

1.3 Fractionation and structural assessment of glycoconjugates by lectin immobilized column chromatography

The use of lectin immobilized column chromatography has been recently evaluated not only for the purification of glycoproteins with pathologically altered sugar chains but also for separating and determining structure of glycoprotein-derived oligosaccharides and glycopeptides (20-22). Lectins are proteins or glycoproteins of non-immune origin that agglutinate cells and/or precipitate specific complex carbohydrates. Different from fractionation techniques such as gel permeation, ion-exchange, partition, reversed-phase and normal-phase chromatography, lectin column chromatography is based on the carbohydrate-binding specificities of lectins. In typical lectin affinity column chromatography, the mixture of glycoconjugates to be resolved is applied to the

immobilized lectin column
lectin and components v
carbohydrate containing
with the specific hapten

The first trial of
glycoproteins was repo
radioabeled Asn-linked
4-Sepharose column m
monosyl residues must
lectin column. This fin
type *N*-linked sugar ch
pentasaccharide, Man₅
core. Since the core ha
column. Because the f
can still interact with
abolishes the interactio
type sugar chain binds
outer chains pass th
oligosaccharides, whic
bind more tightly to a
4-Sepharose column, Og
membrane of polyoma
only enriched in sialic

immobilized lectin column. The glycoconjugates of interest are selectively bound to the lectin and components without affinity for the lectin are then washed away. The bound carbohydrate containing components are dissociated from the lectin by competitive elution with the specific hapten sugar for the lectin.

The first trial of a lectin column for the structural study of the sugar chains of glycoproteins was reported by Ogata et al. (23). Study of the behavior of various radiolabeled Asn-linked oligosaccharides and oligomannosides in a concanavalin A (Con A)-Sephacrose column revealed that at least two nonsubstituted or 2-*O*-substituted α -mannosyl residues must be present in a sugar chain to be retained on the immobilized lectin column. This finding was particularly useful for the structural analysis of complex-type *N*-linked sugar chains. For example, all complex-type oligosaccharides contain a pentasaccharide, $\text{Man}\alpha 1 \rightarrow 6(\text{Man}\alpha 1 \rightarrow 3)\text{Man}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow 4\text{GlcNAc}$ as a common core. Since the core has two α -mannosyl residues, this core binds to a Con A-Sepharose column. Because the first outer chain is linked at the C-2 position, the mannose residue can still interact with Con A. However, substitution with two outer chains totally abolishes the interaction of the α -mannosyl residue. Therefore, a biantennary complex-type sugar chain binds to a Con A-Sepharose column, but those with more than three outer chains pass through the column without interaction. High mannose-type oligosaccharides, which have many $\text{Man}\alpha 1 \rightarrow$ residues and $\rightarrow 2\text{Man}\alpha 1 \rightarrow$ residues, will bind more tightly to a Con A-Sepharose column. With use of this character of a Con A-Sepharose column, Ogata et al. found that the glycopeptides, obtained from the plasma membrane of polyomavirus-transfected baby hamster kidney cells by pronase digestion, are not only enriched in sialic acid residues, but contain more branched structures than those from

day hamster cells (24
transitory complex-t
They also reported that
column with 10mM α -
can be eluted with 250 m
Complex-type s
hexosyl residue linked a
the trimannosyl core.
transplanted at their core
are easily be separated
immobilized *Aleuria*
trifucosylated core p
chains with a fucosylate
15 mM fucose (Fuc)
branching nor by the p
can be used as a more
immobilized lentil and
linked to the core por
binding. The Fuc α 1
Gal β 1 \rightarrow 3(Fuc α 1 \rightarrow 4)G
Fuc α 1 \rightarrow 2Gal β 1 \rightarrow 3Glc
GlcNAc groups show
with more than two of

baby hamster cells (24). Later, Narasimhan et al. added the evidence that bisected biantennary complex-type sugar chains do not bind to a Con A-Sepharose column (25). They also reported that biantennary complex-type oligosaccharides can be eluted from the column with 10mM α -methylglucopyranoside, and high mannose-type oligosaccharides can be eluted with 250 mM α -methylmannopyranoside.

Complex-type sugar chains can be classified by the presence or absence of an α -fucosyl residue linked at the C-6 position of the proximal *N*-acetylglucosamine residue of the trimannosyl core. It was found that some of the hybrid-type sugar chains are also fucosylated at their core (26). The fucosylated complex-type and hybrid-type sugar chains can easily be separated from their nonfucosylated counterparts by passing through an immobilized *Aleuria aurantia* lectin (AAL) column (27). Sugar chains with a nonfucosylated core pass through the column without interaction. In contrast, sugar chains with a fucosylated core bind to the column and are eluted with a buffer containing 0.5 mM fucose (Fuc) (28). Because the binding is affected neither by the status of branching nor by the presence of bisecting *N*-acetylglucosamine residue, the fractionation can be used as a more reliable method to determine the fucosylated sugar chains than immobilized lentil and pea lectin columns. In these cases, not only the fucosyl residue linked to the core portion, but the structure of the outer chains are required for the binding. The $\text{Fuc}\alpha 1\rightarrow 2\text{Gal}1\rightarrow 4\text{GlcNAc}$, the $\text{Gal}\beta 1\rightarrow 4(\text{Fuc}\alpha 1\rightarrow 3)\text{GlcNAc}$, and the $\text{Gal}\beta 1\rightarrow 3(\text{Fuc}\alpha 1\rightarrow 4)\text{GlcNAc}$ groups weakly interact with an AAL column, whereas the $\text{Fuc}\alpha 1\rightarrow 2\text{Gal}\beta 1\rightarrow 3\text{GlcNAc}$, and the $\text{Gal}\beta 1\rightarrow 4\text{GlcNAc}\beta 1\rightarrow 3\text{Gal}\beta 1\rightarrow 4(\text{Fuc}\alpha 1\rightarrow 3)\text{GlcNAc}$ groups show almost no interaction with the matrix. However, oligosaccharide with more than two of these groups bind more strongly than those with only one of the

groups. Therefore,

chromatography for the

Because most of

complex-type sugar chains

which are widely found

particular use. Many gly-

to binding specificity

not intensively. So

monoglucosamine resid-

oligosaccharides with

epitopeagglutinin (E-PA)

existence and indicated t-

Another variation

Although mono- and

branched complex-type

extramembrary oligosac-

Quercus stramonium ag-

highly branched oligosa-

oligosaccharides in a D

for Gal β 1 \rightarrow 4GlcNAc β

column, as long as t-

Oligosaccharides that co-

groups. Therefore, care must be taken in interpreting the results of affinity chromatography for the structural determination of oligosaccharide.

Because most of the *N*-linked sugar chains of glycoproteins are included in the complex-type sugar chain subclass, lectins that interact with the β -galactosyl residues, which are widely found in the outer chain moieties of complex-type sugar chains, are of particular use. Many galactose-binding lectins have been reported (29-30). Among them, the binding specificity of *Ricinus communis* agglutinin I (RCA I) has been investigated most intensively. Several complex-type oligosaccharide contain a bisecting *N*-acetylglucosamine residue. Irimura et al. reported that bisected biantennary complex-type oligosaccharides with the $\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow$ outer chains binds to an erythroagglutinin (E-PHA)-agarose column (31). Cummings and Kornfeld confirmed this evidence and indicated that none of nonbisected sugar chains will bind to the column (32).

Another variation in the complex-type sugar chains is in their branching pattern. Although mono- and biantennary sugar chains can be separated from other highly branched complex-type sugar chains by a Con A-Sepharose column, tri- and tetraantennary oligosaccharides cannot be separated by this column. An immobilized *Datura stramonium* agglutinin (DSA) column can be effectively used to fractionate these highly branched oligosaccharides (33-36). A study of the behavior of 34 complex-type oligosaccharides in a DSA-Sepharose column showed that oligosaccharides that contain the $\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow 4$ ($\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow 2$)Man group are retarded by the column, as long as the pentasaccharide group is not substituted by other sugars. Oligosaccharides that contain unsubstituted $\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}\beta 1 \rightarrow 6(\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}\beta$

(-4)Man group binc
polyglucosamine oligo
structure of the inner c
size. Therefore,
oligosaccharides into t
oligosaccharides; the "n
;4 branch; and the "bo
trond and tetraantenna

One lectin of th
meryfolia), which has
[1-3) N-acetylgalact
glycosidase exists, lec
technique for isolating
localizing sites of attac
oligosaccharides by jac
of IgA and C1 inhibito
glycoproteins to jacali
other proteins that con
only a single O-link
oligosaccharides (37-38
a useful technique for
greatly assist in localizi
has been a difficult ana

1→2)Man group bind to the column and are eluted with buffer containing *N*-acetylglucosamine oligomers. The binding characteristic was not affected by either the structure of the inner core portion or by the presence of bisecting *N*-acetylglucosamine residue. Therefore, the column can fractionate a mixture of complex-type oligosaccharides into three fractions: the "pass-through fraction" contains biantennary oligosaccharides; the "retarded fraction" contains triantennary oligosaccharides, with a C-2, 4 branch; and the "bound fraction" contains triantennary oligosaccharides, with a C-2, 6 branch and tetraantennary oligosaccharides.

One lectin of the greatest interest is jacalin, a lectin from jackfruit (*Artocarpus integrifolia*), which has high specificity and affinity for the core disaccharide, galactosyl $\beta(1\rightarrow3)$ *N*-acetylgalactosamine, in *O*-linked oligosaccharides (37-38). As no universal *O*-glycosidase exists, lectin affinity chromatography using jacalin-agarose is a useful technique for isolating glycopeptides containing *O*-linked oligosaccharides and thereby localizing sites of attachment of these oligosaccharides. The specific binding of *O*-linked oligosaccharides by jacalin coupled to agarose has been exploited for affinity purifications of IgA and C1 inhibitor from human plasma. Recent characterization of the binding of glycoproteins to jacalin-agarose indicates that this immobilized lectin specifically binds other proteins that contain of *O*-linked oligosaccharides, including glycoproteins bearing only a single *O*-linked oligosaccharide and those bearing sialic acid-containing oligosaccharides (37-38). Lectin affinity chromatography on jacalin-agarose also provides a useful technique for purifying glycopeptides from glycoprotein digests. This would greatly assist in localizing peptides with *O*-linked oligosaccharides. In many instances, this has been a difficult analytical problem even for extensively characterized proteins because,

like the case of *N*-lin

lated glycosylation ba

This chapter d

oocyte zona pellucida

fusion, and MALDI

into-carbohydrate aff

glycopeptides from try

of glycoproteins. Lect

14-kDa protein with

osides (Ca^{++} and M

glycoprotein with bindi

1. Experimental

Materials

Rhönuclease E

were purchased from S

porine oocyte zona

Ureawicz (Department

Medicine). Jacalin-ag

depending grade t

Hexaminidase (from

unlike the case of *N*-linked oligosaccharides, it is not possible to reliably predict sites of *O*-linked glycosylation based on the amino acid sequence of a protein.

This chapter describes our efforts in the structural characterization of porcine oocyte zona pellucida 3 α -glycoprotein (ZP3 α) using HPLC separation, exoglycosidase digestion, and MALDI-MS analysis. We also have investigated the utility of some specific lectin-carbohydrate affinity interactions in combination with MALDI-MS for identifying glycopeptides from tryptic digest and determining site-specific carbohydrate heterogeneity of glycoproteins. Lectins studied in this chapter include: (1) Concanavalin A (Con A): an 104-KDa protein with binding capability for terminal α -D-mannosyl and α -D-glucosyl residues (Ca^{++} and Mn^{++} are required for its activity), and (2) Jacalin: a 50-KDa glycoprotein with binding capability for only *O*-glycosidically linked oligosaccharides.

2. Experimental

Materials

Ribonuclease B (Type III-B, from bovine pancrease), avidin, α -methylgalactoside were purchased from Sigma (St. Louis, MO). Endo- β -galactosidase digested and purified porcine oocyte zona pellucida glycoprotein 3 α (ZP3 α) was provided by Dr. E. C. Yurewicz (Department of Obstetrics and Gynecology, Wayne State University School of Medicine). Jacalin-agarose was obtained from Vector Laboratories (Watertown, MA). Sequencing grade trypsin was from Boehringer Mannheim (Indianapolis, IN). Neuraminidase (from *Arthrobacter ureafaciens*), β -galactosidase (from *Diplococcus*

pneumonia), and *N*-a

lysoSystems (Bedfo

MALDI-MS

MALDI-MS e

pectrometer (PerSept

m, 2,5-Dihydroxybe

throughout this chap

experiments. A delay t

HPLC of digested glyco

Microbore rev

UMA (M6Chrom BioF

m, 5- μ m particle, 3

45%-45% acetonitrile

was maintained at 50

Individual fractions we

Oligoprotein reduction

The reduction

ly digestion. In this st

the HCl buffer (pH 8

ridon was mixed with

pneumonia), and *N*-acetylhexosaminidase (from Jack bean) were purchased from Oxford GlycoSystems (Bedford, MA).

MALDI-MS

MALDI-MS experiments were carried out on a Voyager-Elite time-of-flight mass spectrometer (PerSeptive Biosystems Inc.) equipped with a nitrogen laser emitting at 337 nm. 2,5-Dihydroxybenzoic acid (DHB, recrystallized) was used as the MALDI matrix throughout this chapter. The ion source accelerating voltage was 20 kV in all TOF experiments. A delay time of 50 ns was used before the ion extraction.

HPLC of digested glycopeptides

Microbore reverse phase HPLC experiments were performed using a MiChrom UMA (MiChrom BioResources) instrument equipped with a PLRP-S column (1.0 × 150 mm, 5-μm particle, 300-Å pore size). Elution was achieved with a linear gradient of 4.5%-45% acetonitrile in 0.1% TFA mobile phase over a period of 80 min. The flow rate was maintained at 50 μl/min and effluent was monitored continuously at 214 nm. Individual fractions were collected and used for further MALDI-MS analysis.

Glycoprotein reduction and carboxymethylation

The reduction and S-carboxymethylation for glycoproteins were performed prior to digestion. In this study, 300 pmol ZP3α glycoprotein was dissolved in 200 μl 0.5 M Tris-HCl buffer (pH 8.2) containing 2 mM EDTA and 6.0 M guanidine-HCl. This solution was mixed with 10 μl of 0.4 M dithiothreitol (DTT) and incubated at 37°C for 3

1245 A volume of 2

1247 Avidin and rib

1249 The reduce

1251 tubing (3,500 l

1253 Digestion with Trypsin

For preparation

1257 Ribonuclease B,

1259 sequencing grade try

1261 incubation was perform

1263 Digestion with exoglyc

Glycopeptides

1267 Glycosidases were t

1269 Lectin (Jacalin) affinity

Prior to affinity

1273 dissolved in 50 mM

1275 0.02% sodium azide (

1277 concentration of 2 mM

1279 trypsin activity. The

1281 equilibrated with buffer

1283 sample has entered the

hours. A volume of 20 μ l of 0.8 M iodoacetic acid was then added and incubated for 1 hour. Avidin and ribonuclease B samples (100 pmol each) were treated in an identical manner. The reduced and carboxymethylated samples were desalted overnight using dialysis tubing (3,500 MW cut-off) from Fisher and dried for further experiments.

Digestion with Trypsin

For preparation of tryptic digests, reduced and carboxymethylated preparations of ZP3 α , ribonuclease B, avidin were dissolved in 100 mM ammonium bicarbonate (pH 8.0). Sequencing grade trypsin was added at an enzyme:substrate ratio of 1:50 (w/w) and incubation was performed at 37°C for 24 hours.

Digestion with exoglycosidases

Glycopeptides were dissolved in 25 mM ammonium acetate buffer (pH 5.5). Exoglycosidases were then added and incubations were performed at 37°C for 24 hours.

Lectin (jacalin) affinity chromatography

Prior to affinity chromatography on jacalin-agarose, tryptic digests were dried and redissolved in 50 mM sodium phosphate buffer, pH 7.0, containing 500 mM NaCl and 0.02% sodium azide (buffer A). Phenylmethylsulfonyl fluoride was added to a final concentration of 2 mM and the sample incubated at 4°C for 30 min to inactivate residual trypsin activity. The sample was then applied to a minicolumn of jacalin-agarose equilibrated with buffer A using a ratio of 1 mg digest per ml of settled resin. After the sample has entered the column, flow was stopped and the column incubated for 30 min to

asure binding of O-
liquots of cold (4°C)
are subsequently elu
-methylgalactoside.

1. Results and Discus

1.1. Structural cha

2P3a)

The zona pelli
glycoconjugates which
programmed manner
translational processi
serine/threonine residu
heterogeneous with re
residues representing
intracellular trafficking
properties of zona g
presumably as ligands
surface (43).

The pig represe
the structural and func
zona is comprised of th

ensure binding of *O*-glycopeptides. The column was washed with 20 column-volume aliquots of cold (4°C) buffer A and each fraction collected separately. *O*-glycopeptides were subsequently eluted with ten column-volume aliquots of buffer A containing 50 mM α -methylgalactoside.

3. Results and Discussion

3. 1. Structural characterization of porcine oocyte zona pellucida glycoprotein 3 α (ZP3 α)

The zona pellucida of mammalian oocyte are comprised of sex and gamete-specific glycoconjugates which are expressed during oocyte growth in a developmentally programmed manner (39-41). Zona glycoproteins are characterized by extensive post-translational processing, including *N*- and *O*-glycosylation on asparagine and serine/threonine residues, respectively. Zona sugar chains examined to date are highly heterogeneous with respect to structure, size, and charge with polylatosamines and sulfate residues representing distinguishing characteristics (40, 42). In addition to a role in intracellular trafficking, zona carbohydrates strongly influence the physicochemical properties of zona glycoproteins and participate in sperm-zona recognition events, presumably as ligands for complementary carbohydrate-binding proteins on the sperm surface (43).

The pig represents an attractive alternative to rodent species for investigations of the structural and functional properties of zona pellucida glycoproteins. The pig oocyte zona is comprised of three structurally and immunologically distinct glycoproteins, which

ZP1 (82-KDa prote

among them, ZP3 α is

in this study. The ar

glycosylation sites at As

and Thr 304 ha

glycosylation analysis.

characterize the site-spe

Reduced and cau

peptides and glycopept

ELC fractions by MA

fractions. These fou

HPLC chromatogr

obtained from one of t

the mass, these peptid

the carbohydrates was

the 203 Da correspond

Isolation and id

glycoprotein can be a

the reverse phase H

exactly to a reverse

gradient program use

single peak was found

are ZP1 (82-KDa protein), ZP3 α (55-KDa protein), and ZP3 β (55-KDa protein) (44-47). Among them, ZP3 α is a highly glycosylated protein and was used as a model compound for this study. The amino acid sequence of ZP3 α is shown in Figure 4.4. Three *N*-glycosylation sites at Asn 203, Asn 220, and Asn 333 and two *O*-glycosylation sites at Ser 293 and Thr 304 have been found before for this protein by HPLC and Edman degradation analysis. The main objective for our study on ZP3 α glycoprotein is to characterize the site-specific microheterogeneity and carbohydrate structure.

Reduced and carboxymethylated ZP3 α glycoprotein was digested and the resulting peptides and glycopeptides were separated by reverse phase HPLC. Analysis of the HPLC fractions by MALDI-MS showed the presence of glycosylated peptides in four of the fractions. These four fractions are labeled as Asn 333, Asn 220, Asn 203, and OGP in the HPLC chromatogram as shown in Figure 4.5(a). The MALDI mass spectrum obtained from one of the four fractions, Asn 220, was shown in Figure 4.6. Based on their mass, these peptides were identified as glycosylated peptides. Microheterogeneity of the carbohydrates was revealed by clusters of peaks separated by a mass difference of 162 and 203 Da corresponding to hexose and *N*-acetylhexosamine residues, respectively.

Isolation and identification of *O*-linked glycopeptides from tryptic digest of ZP3 α glycoprotein can be accomplished by sequential lectin (jacalin) affinity chromatography and reverse phase HPLC. The affinity purified *O*-linked glycopeptides were applied directly to a reverse phase HPLC column for desalting and purification with the same gradient program used for the separation of the tryptic digest of ZP3 α glycoprotein. A single peak was found in the HPLC run as shown in Figure 4.5(b). This peak, at the same

GAATTCGGGGTGA

CGCTGTGTCTTCTC
P L C L A L

CAAGAGCTTTCATT
P S S F H P

ACAGGGGGGGGGGG
P R G R L

AGTCATGGGAGTGG
S S H G V E

ATGGACTTGAAGAAG
I G L E E A

GAATCTAGCTCTTG
C F L A L D

ATCTCCACCCATCA
L P P P I T

CTTACTATGGAACA
C Y Y G N T

GTGAGTCACCTCCAC
V T S P P I

ATGAAACACACACTT
M E T H T F

ACCAAGGGGTATATC
N Q A V Y F

CGAGACAGCATCTTC
R D S I F F

CAAGTTTCACTCTCT
Q V F T L L

AAGATGAACGCTAT
K D E R Y

TATGTGAGGTCTCT
Y V E V S

CCCGCATGAGCCCC
P G H S P

TACCAGACCAACTT
Y Q T K L

GTTCACCTTCAG
V S T F S

GCATCGTCTGCA
A S V C K

TCTGACATCCATT
S D I H F

CGGACTCTTCAG
R D S S E

CTCTGGGAAGCT
L L G S I

CGACCAAAATGTC

Figure 4.4: Amino
[P3a]

GAATTCGGGGTGAAGTACCTGTTCTCCGACGGCGCTATGTGGTTGCGGGCCGTCCATCTGGCTCTGCTTT	70
M W L R P S I W L C F	11
CCGCTGTGTCTTGCTCTGCCAGGCCAGTCTCAGCCCCAAGCAGCAGATGACCTTGGTGGCCTCTACTGTGGG	142
P L C L A L P G Q S Q P K A A D D L G G L Y C G	35
CCAAGCAGCTTTTCATTTCTCCATAAATCTTCTCAGCCAGGACACAGCAACTCCTCTGCACTGGTGGTTTGG	214
P S S F H P S I N L L S Q D T A T P P A L V V W	59
GACAGGCGCGGGCGGCTGCACAAGCTGCAGAATGACTCTGGCTGTGGCACGTGGGTCCACAAGGGCCCAGGC	286
D R R G R L H K L Q N D S G C G T W V H K G P G	83
AGCTCCATGGGAGTGAAGCATCCTACAGAGGCTGCTATGTGACTGAGTGGGACTCTCACTACCTCATGCCC	358
S S H G V E A S Y R G C Y V T E W D S H Y L M P	107
ATTGGACTTGAAGAAGCAGATGCAGGTGGACACAGAACAGTCACAGAGACGAAACTGTTTAAAGTGCCTGTG	430
I G L E E A D A G G H R T V T E T K L F K C P V	131
GATTTCTAGCTCTTGATGTTCCAAACATTGGCCCTTTGTGATGCTGTCCAGTGTGGGACCGATTGCCATGT	502
D F L A L D V P T I G L C D A V P V W D R L P C	155
GCTCTCCACCCATCACTCAAGGAGAATGAAGCAGCTTGGCTGCTGTACAACCTCGGAAGAGGTCCCTTCT	574
A P P P I T Q G E C K Q L G C C Y N S E E V P S	179
TGTTACTATGGAACACAGTGACCTCACGCTGTACCCAAGATGGCCACTTCTCCATCGCTGTGTCTCGCAAT	646
C Y Y G N T V T S R C T Q D G H F S I A V S R N	203
GTGACCTCACCTCCACTGCTCTGGGATTCTGTGCACCTGGCCTTCAGAAATGACAGTGAATGTAACCTGTG	718
V T S P P L L W D S V H L A F R N D S E C K P V	227
ATGGAACACACACTTTTGTCTCTTCCGGTTTCCATTTAGTTCTGTGGGACTGCAAAACGGGTAACCTGGG	790
H E T H T F V L F R F P F S S C G T A K R V T G	251
AACCAGGCGGTATATGAAAATGAGCTGGTAGCAGCTCGGGATGTGAGGACTTGGAGCCATGGTTCTATTACC	862
N Q A V Y E N E L V A A R D V R T W S H G S I T	275
CGAGACAGCATCTTCAGGCTTCGAGTCAGTTGTATCTACTCTGTAAGTAGCAGTGTCTCCAGTTAACATC	934
R D S I F R L R V S C I Y S V S S S A L P V N I	299
CAGGTTTTCACTCTCCACACCGCTTCCGGAGACCCACCTGGACCTCTTACTCTGGAGCTTCAGATTGCC	1006
Q V F T L P P P L P E T H P G P L T L E L Q I A	323
AAAGATGAACGCTATGGCTCCTACTACAATGCTAGTGACTACCCGGTGGTGAATTTGCTTCGGGAGCCCATC	1078
K D E R Y G S Y Y N A S D Y P V V K L L R E P I	347
TATGTGGAGGTCTCTATCCGTCACCGAACAGACCCAGTCTCGGGCTGCACCTGCACAGTGTGGGCCACA	1150
Y V E V S I R H R T D P S L G L H L H Q C W A T	371
CCCGCATGAGCCCCCTGCTCCAGCCACAGTGGCCCATGCTAGTCAATGGATGCCCTTACACTGGAGACAAC	1222
P G H S P L L Q P Q W P M L V N G C P Y T G D N	395
TACCAGACCAAACTGATCCCTGTCCAGAAAGCCTCAAACTGCTATTTCTTCTCACTACCAGCGTTTCAGT	1294
Y Q T K L I P V Q K A S N L L F P S H Y Q R F S	419
GTTTCCACCTTCAGTTTGTGGACTCTGTGGCAAAGCAGGCACTCAAGGGACCGGTGTATCTGCATTGTACT	1366
V S T F S F V D S V A K Q A L K G P V Y L H C T	443
GCATCGGTCTGCAAGCCTGCAGGGGCACCGATCTGTGTGACAACTGTCTGTGCCAGACGAAGAAGAAGT	1438
A S V C K P A G A P I C V T T C P A A R R R R S	467
TCTGACATCCATTTTCAGAATGGCACTGCTAGCATTTCTAGCAAGGGTCCCATGATTCTACTCCAAGCCACT	1510
S D I H F Q N G T A S I S S K G P M I L L Q A T	491
CGGGACTCTTCAGAAAGGCTCCATAAATACTCAAGGCCTCTGTAGACTCCCATGCTCTGTGGGTGGCTGGC	1582
R D S S E R L H K Y S R P P V D S H A L W V A G	515
CTCTTGGGAAGCTTAATTATTGGAGCCTTGTAGTGTCTACCTGGTCTTCAGGAAATGGAGATGAGTTACT	1654
L L G S L I I G A L L V S Y L V F R K W R *	536
CAGACCAATGTGTCAATAAAACCAATAAAACAAAACCGGAATTC	1699

Figure 4.4: Amino acid sequence of porcine oocyte zona pellucida Glycoprotein 3 α (ZP3 α)

retention time as the
glycopeptides.

Further MALDI-TOF MS analysis of the heterogeneity of the glycopeptides likely represented the different sites of (Gal β -GalNAc) attachment. The deglycosylated peptides were analyzed using HPLC and ESI-MS. The major peak was at Thr 304. However, the analysis by gel electrophoresis showed a predominant peak at 10 kDa. A predominant peak at 10 kDa is characteristic of O-linked oligosaccharides.

retention time as the OGP fraction in Figure 4.5(a), positively identified the *O*-linked glycopeptides.

Further MALDI-MS analysis (Figure 4.7) of the collected fraction showed heterogeneity of the *O*-linked glycopeptides. The peak at the low-mass end at m/z 5146.9 likely represented the $[M+H]^+$ ion of a glycopeptide with two *O*-linked disaccharide units (Gal β -GalNAc) attached because the m/z difference between this ion and $[M+H]^+$ ion of the deglycosylated peptide (m/z 4416.1) is 730.1. Previous study by Yurewicz's group using HPLC and Edman degradation has confirmed two *O*-glycosylation sites at Ser 293 and Thr 304. However, it was suspected that there may be one more *O*-glycosylation site from analysis by gel electrophoresis on chemically cleaved *O*-linked oligosaccharides (47). A predominant peak at m/z 5716.5 in Figure 4.7 supported this hypothesis since most of *O*-linked oligosaccharides are disaccharides or trisaccharides.



Figure 4.5: (a) HPLC chromatogram of a hydrophobic peptide fraction. (b) HPLC chromatogram of a hydrophilic peptide fraction.

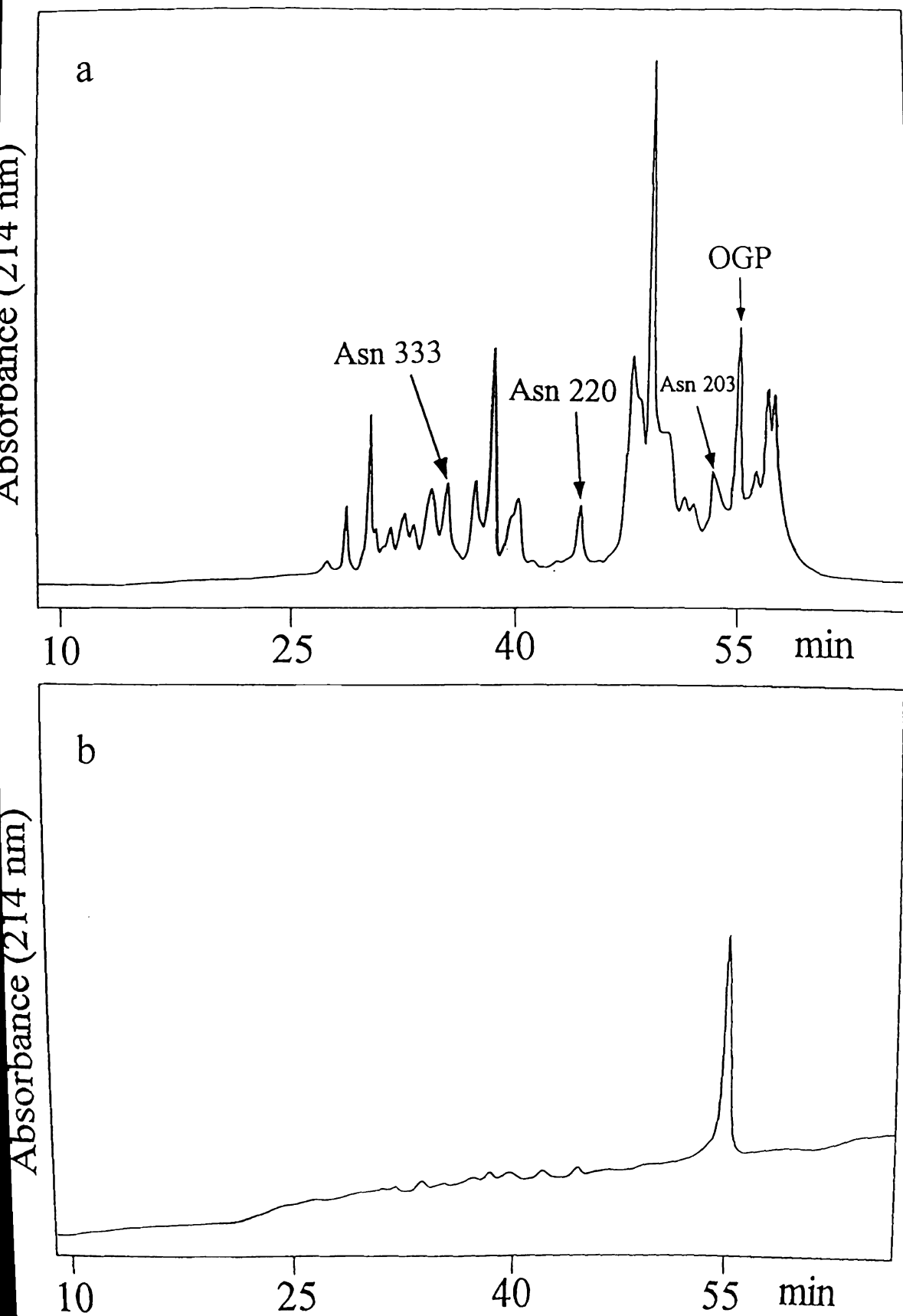


Figure 4.5: (a) HPLC of tryptic digest of ZP3a glycoprotein. (b) HPLC of the peptide fractions from ZP3a glycoprotein after lectin (jacalin) affinity chromatography.



Figure 4.6: *N*-link
Spectrum was obtained

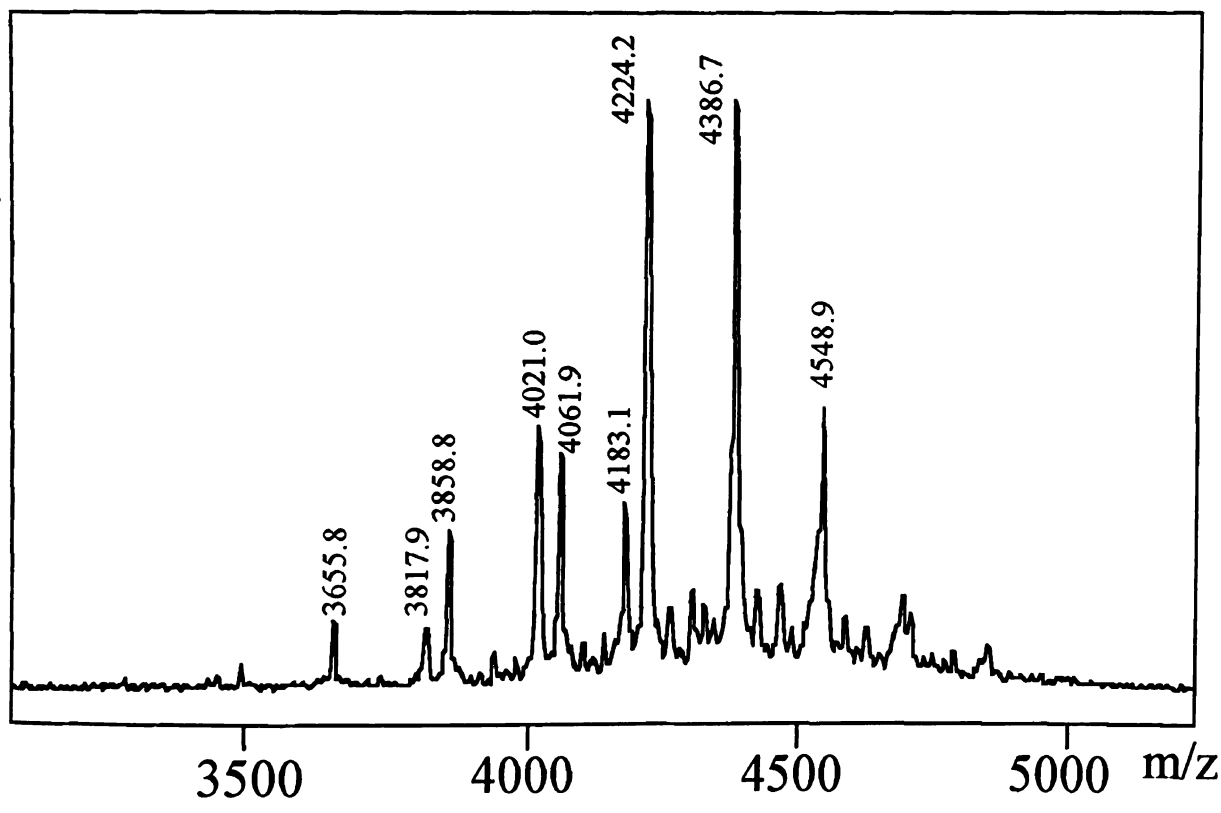


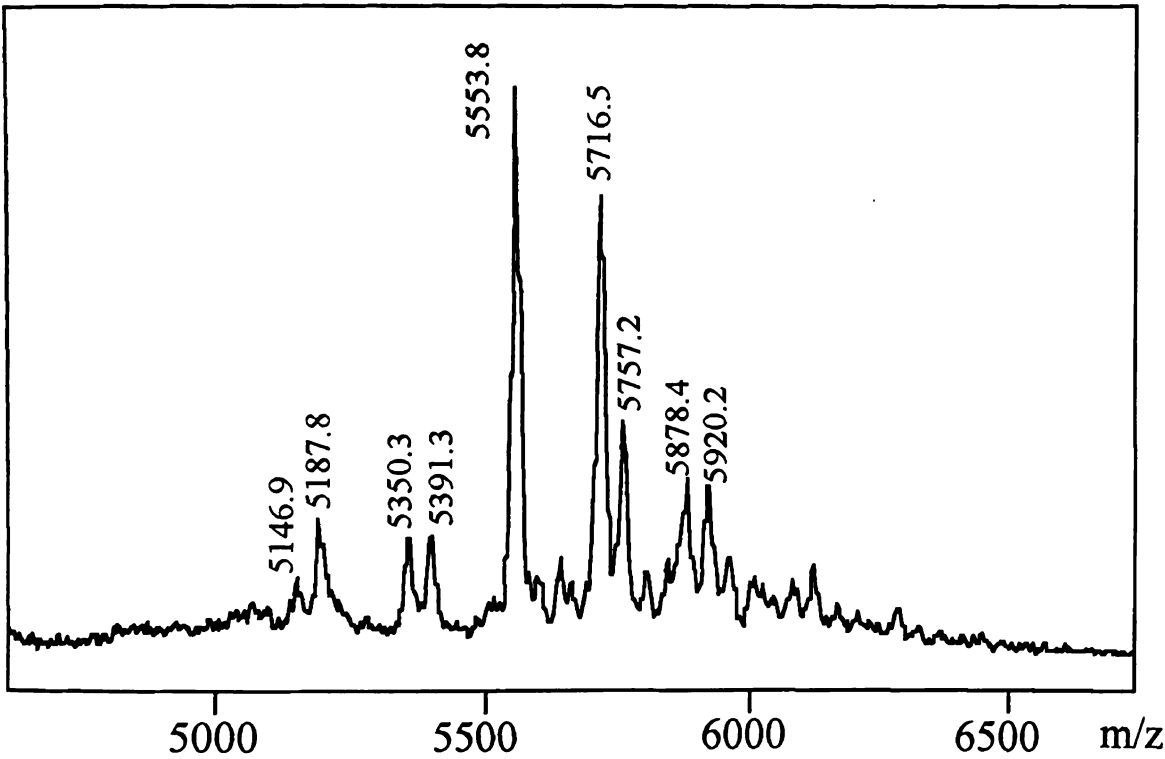
Figure 4.6: *N*-linked glycopeptides at Asn 220 of ZP3a glycoprotein. The MALDI-MS spectrum was obtained from a collected HPLC fraction.



50

VSCIYSVSS
 *. O-glycosyl

Figure 4.7: O-linked
 (x2in) affinity chro



VSCIYSVSSS*ALPVNIQVFTL*PPPLPETHPGPLTLELQIAK
*: O-glycosylation site

Figure 4.7: O-linked glycopeptides of ZP3a glycoprotein after purification by lectin (ConA) affinity chromatography and HPLC.

Three other
determined to be N
Typically, N-linked
glycosidase F (PNG
linked) carbohydrat
treatment to monito
oligosaccharide-bear
of the glycopeptide.

lucions, Asn 203
provides an alternat
glycopeptides, espec

The MALDI
shown previously in
sequential digestion
loses by MALDI-M
revealed by a mass c
followed by N-acetyl
spaced peaks at m/z
spectrum (Figure 4.8)
single peak at m/z 32
[N-H]⁺ ion of a gl
removed. The mass c

Three other glycopeptide fractions, Asn 220, Asn 203, and Asn 333 were determined to be *N*-linked glycopeptides as they were not retained on jacalin-agarose. Typically, *N*-linked glycopeptides can be identified after the treatment with peptide-*N*-glycosidase F (PNGase F), an endoglycosidase that cleaves off asparagine-linked (*N*-glycosylated) carbohydrates. MALDI-MS analysis can be incorporated with PNGase F treatment to monitor the mass shift of the analyte. This enzyme does not function if the oligosaccharide-bearing asparagine residue is located at either N- terminal or C- terminal of the glycopeptide. This happened to be the case for two of the *N*-linked glycopeptide fractions, Asn 203 and Asn 220, from ZP3 α glycoprotein. Jacalin affinity binding provides an alternative approach to distinguish *N*-linked glycopeptides and *O*-linked glycopeptides, especially when PNGase F treatment fails.

The MALDI mass spectrum of one *N*-linked glycopeptide fraction at Asn 220, was shown previously in Figure 4.5. The carbohydrate structure was further characterized by sequential digestion with specific exoglycosidases and by monitoring the resulting mass changes by MALDI-MS. Since the microheterogeneity of this glycopeptide fraction was revealed by a mass difference of 162 and 203 Da, it was digested with β -galactosidase followed by *N*-acetylhexosaminidase. After β -galactosidase digestion, a series of evenly spaced peaks at m/z 3250.4, 3453.1, 3655.8, 3858.6 was observed in the MALDI mass spectrum (Figure 4.8(a)). After β -galactosidase and *N*-acetylhexosaminidase digestion, a single peak at m/z 3250.2 was produced as shown in Figure 4.8(b), which represented the $[M+H]^+$ ion of a glycopeptide with both terminal galactose and *N*-acetylhexosamine removed. The mass difference between this ion and the $[M+H]^+$ ion of the deglycosylated

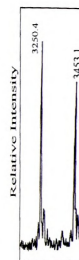


Figure 4.8: N-linked
spectrum was obtained
the β -galactosidase

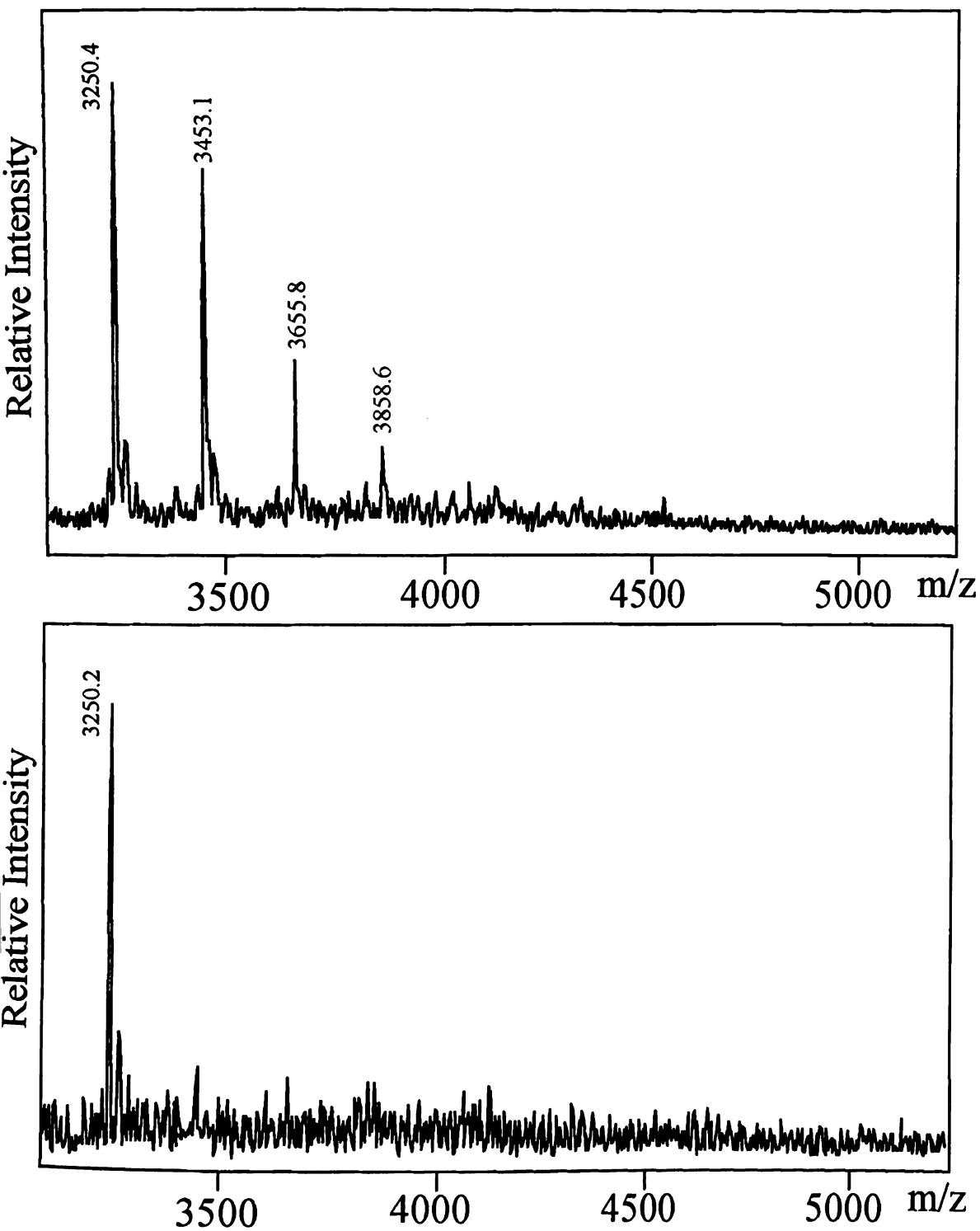


Figure 4.8: N-linked Glycopeptides at Asn 220 of ZP3a Glycoprotein. MALDI-MS spectrum was obtained on the glycopeptide sample (a) after β -galactosidase treatment; (b) after β -galactosidase & *N*-acetylhexoaminidase treatment.

peptide (m/z 2211.5

in be (Man)₂(GlcNAc)

are oligosaccharide

good agreement with

Figure 4.10

glycopeptide fraction

revealed by mass di

107 Da, which illust

and *N*-glycolylneuram

represented the [M+

The peaks for the

glycopeptides were

with much lower inte

oughcosidases: neu

A single peak at m/z

Figure 4.10(b), which

and galactose and *N*

and the [M+H]⁺ ion

confirmed that

(Man)₄(GlcNAc)₂(Fu

oligosaccharide struc

peptide (m/z 2211.5) was 1038.7 Da, which determined the core oligosaccharide structure to be $(\text{Man})_3(\text{GlcNAc})_2(\text{Fuc})$. Other glycopeptide structures were assigned based on this core oligosaccharide structure in Figure 4.9. Observed m/z values of $[\text{M}+\text{H}]^+$ ions were in good agreement with calculated m/z values.

Figure 4.10(a) showed the MALDI mass spectrum of the Asn 333 *N*-linked glycopeptide fraction. The microheterogeneity of this glycopeptide fraction was not only revealed by mass difference of 162 and 203 Da, but also by mass difference of 291 and 307 Da, which illustrate the existence of two types of sialic acid, *N*-acetylneuraminic acid and *N*-glycolylneuraminic acid. The peaks at m/z 3541.1 and 3704.2 in Figure 4.10(a) represented the $[\text{M}+\text{H}]^+$ ions of two *N*-glycolylneuraminic acid containing glycopeptides. The peaks for the $[\text{M}+\text{H}]^+$ ions of two analogous *N*-acetylneuraminic acid- containing glycopeptides were also observed at m/z 3525.0 and 3688.4 (not labeled in the spectrum) with much lower intensity. Glycopeptide fraction at Asn 333 was digested with a series of exoglycosidases: neuraminidase followed by β -galactosidase and *N*-acetylhexosaminidase. A single peak at m/z 2665.4 was produced after the sequential digestion as shown in Figure 4.10(b), which represented the $[\text{M}+\text{H}]^+$ ion of a glycopeptide with terminal sialic acid, galactose and *N*-acetylhexose removed. Again, the mass difference between this ion and the $[\text{M}+\text{H}]^+$ ion of the deglycosylated peptide (m/z 1626.8) was 1038.6 Da, which confirmed that the core oligosaccharide has a fucosylated structure, $(\text{Man})_3(\text{GlcNAc})_2(\text{Fuc})$. Other glycopeptide structures were assigned based on this core oligosaccharide structure in Figure 4.11.

Proposed structure of the glycopeptide	Calculated <i>m/z</i>	Observed <i>m/z</i>	Proposed structure of the glycopeptide	Calculated <i>m/z</i>	Observed <i>m/z</i>
$\begin{array}{c} \text{GN-M} \diagup \text{M-GN-GN-Asn}^{220} \\ \text{GN-M} \diagdown \end{array}$	3656.8	3655.8	$\begin{array}{c} \text{GN-M} \diagup \text{M-GN-GN-F} \\ \text{GN-M} \diagdown \end{array}$	220	

Proposed structure of the glycopeptide	Calculated <i>m/z</i>	Observed <i>m/z</i>	Proposed structure of the glycopeptide	Calculated <i>m/z</i>	Observed <i>m/z</i>
$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	3656.8	3655.8	$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	4063.2	4061.9
$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	3819.0	3817.9	$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	4225.4	4224.2
$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	3860.0	3858.8	$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	4387.5	4386.7
$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	4022.2	4021.0	$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	4549.6	4548.9
$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$	4184.3	4183.1	$\begin{array}{c} \text{GN-M} \\ \diagdown \\ \text{M-GN-Asn}^{220} \\ \diagup \\ \text{GN-M} \end{array}$		

Figure 4.9: N-linked Glycopeptides at Asn 220 of ZP3a Glycoprotein. Observed *m/z* and calculated *m/z* were assigned to the $[M+H]^+$ ions of the glycopeptides.

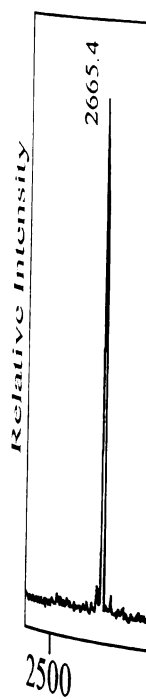
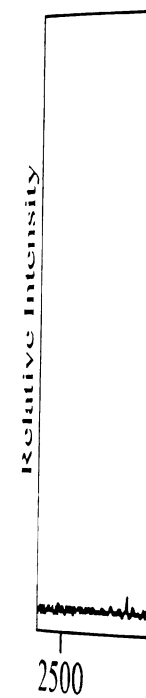


Figure 4.10: N-linked
spectrum was obtained
after β -galact

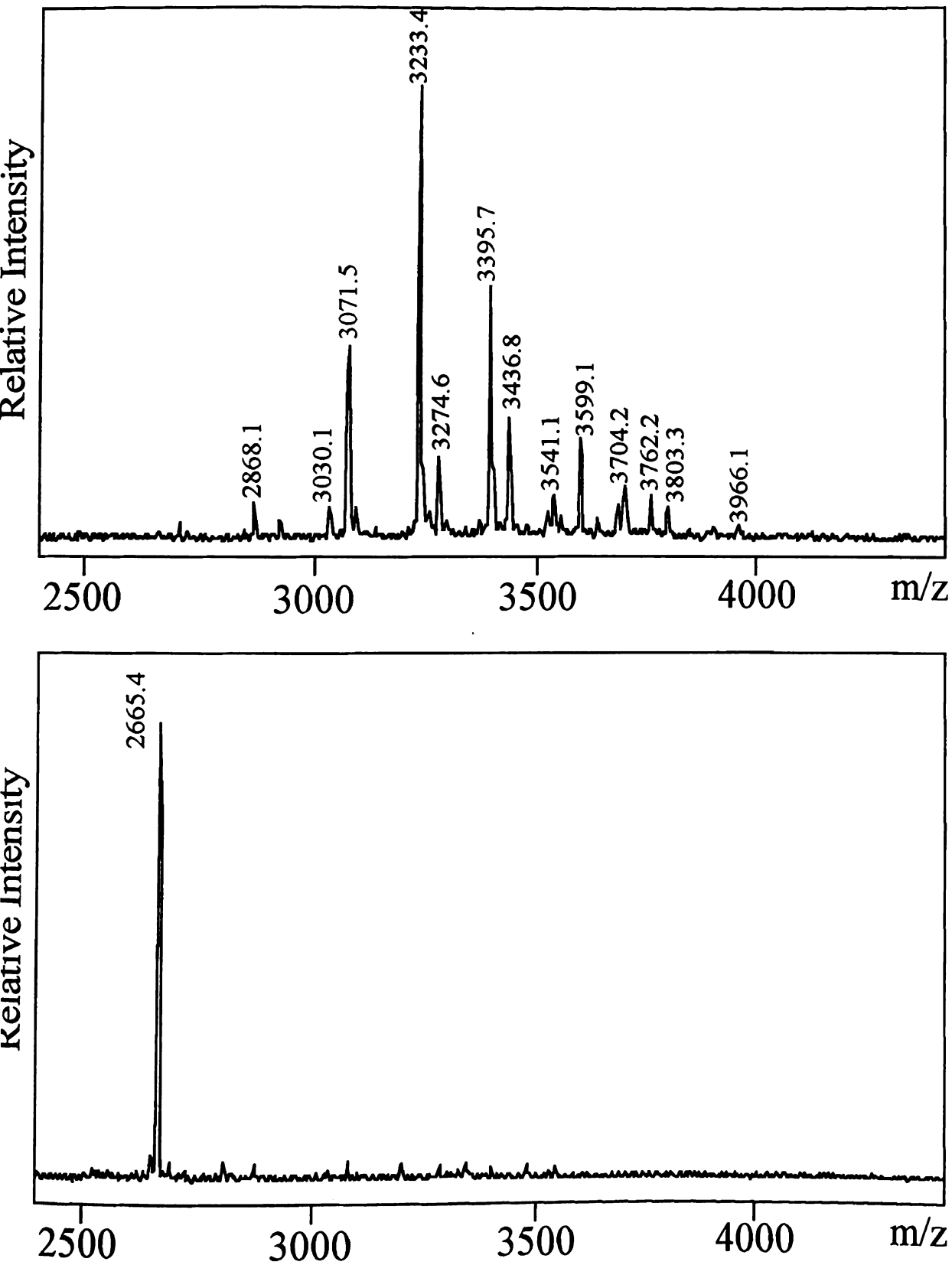


Figure 4.10: N-linked Glycopeptides at Asn 333 of ZP3a Glycoprotein. (a) MALDI-MS spectrum was obtained on original HPLC fraction. (b) MALDI-MS spectrum was obtained after β -galactosidase & N -acetylhexoaminidase treatment

Proposed structure of the Glycopeptide	Calculated <i>m/z</i>	Observed <i>m/z</i>	Proposed structure of the Glycopeptide	Calculated <i>m/z</i>	Observed <i>m/z</i>
$\text{GN-M} \begin{array}{c} \text{M} \\ \text{GN} \end{array} \begin{array}{c} \text{F} \\ \text{GN} \end{array} \text{GN-Asn}^{333}$	2868.9	2868.1	$\text{GN-M} \begin{array}{c} \text{M} \\ \text{GN} \end{array} \begin{array}{c} \text{F} \\ \text{GN} \end{array} \text{GN-Asn}^{333}$	3437.4	3436.8






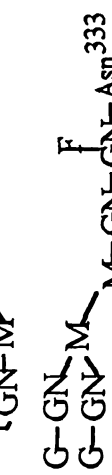
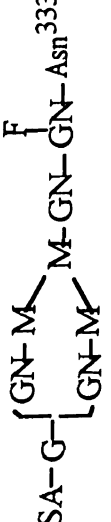



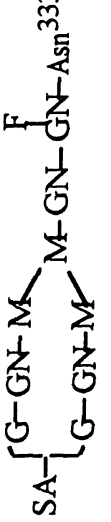

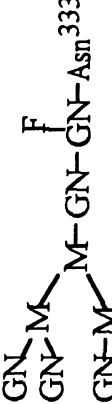
Proposed structure of the glycopeptide	Calculated m/z	Observed m/z	Proposed structure of the glycopeptide	Calculated m/z	Observed m/z
	2868.9	2868.1		3437.4	3436.8
	3072.1	3071.5		3599.6	3599.1
	3234.3	3233.4		3761.7	3762.2
	3541.5	3541.1		3640.6	3640.2
	3396.4	3395.7		3802.8	3803.3
	3703.7	3704.2		3964.9	3966.1
	3275.3	3274.6			

Figure 4.11: *N*-linked glycopeptides at Asn 333 of ZP3a glycoprotein. Observed m/z and calculated m/z were assigned to the $[M+Na]^+$ ions of the glycopeptides.

The MAL

glycoprotein (Fig

Asn 220 or Asn 33

which released m

3450 and m/z 29

mass differences b

m/z 1953.3) were

Asn 203 consists

glycopeptide struct

Figure 4.13 and Fig

3.2. A micro-batch

To minimize

MALDI-MS, a mic

analysis of isolated

this process is illus

centrifuge filter tu

carboxymethylated g

filter tube containing

concentration salt-co

were resuspended in

MALDI-MS analysis

The MALDI mass spectrum of the Asn 203 *N*-linked glycopeptides of ZP3 α coprotein (Figure 4.12(a)) showed more heterogeneity than that of the glycopeptides at 220 or Asn 333. After the treatment with β -galactosidase and *N*-acetylhexoaminidase which released most of the outer branches of these glycopeptides, two peaks at m/z 1953.0 and m/z 2991.5 were found in the MALDI mass spectrum (Figure 4.12(b)). The mass differences between these ions and the $[M+H]^+$ ion of the deglycosylated peptide (m/z 1953.3) were 1038.2 and 891.7 Da, respectively, which indicated that the glycan at Asn 203 consists a mixture of fucosylated and unfucosylated structures. Other glycopeptide structures were assigned based on these core oligosaccharide structures in Figure 4.13 and Figure 4.14.

A micro-batch lectin affinity binding process for direct MALDI-MS analysis

To minimize sample handling between the stages of lectin affinity binding and MALDI-MS, a micro-batch lectin binding process is proposed for direct MALDI-MS analysis of isolated glycopeptides without further HPLC purification. A flow-diagram of the process is illustrated in Figure 4-15. Lectin-immobilized beads were put into a centrifuge filter tube and equilibrated with the binding buffer. Reduced and carboxymethylated glycoproteins were digested with trypsin and applied to the centrifuge filter tube containing lectin beads. After the binding of glycopeptides and lectin, high concentration salt-containing binding buffer was spun off and lectin-immobilized beads were resuspended in matrix (2,5-dihydroxybenzoic acid) solution and subjected to direct MALDI-MS analysis. 2, 5-Dihydroxybenzoic acid was chosen as the matrix in this study

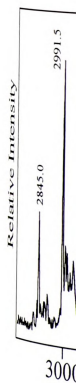


Figure 4.12: N-link
spectrum obtained o
galactosidase & N-a

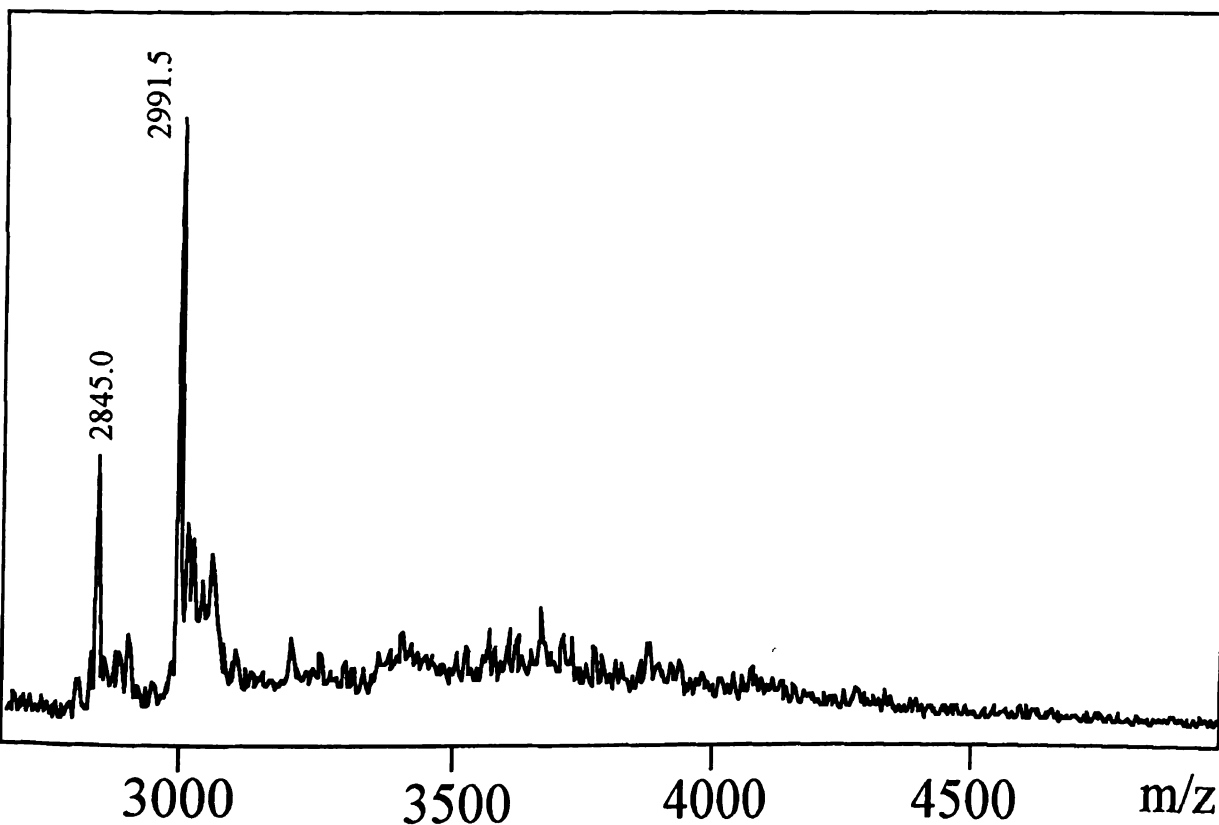
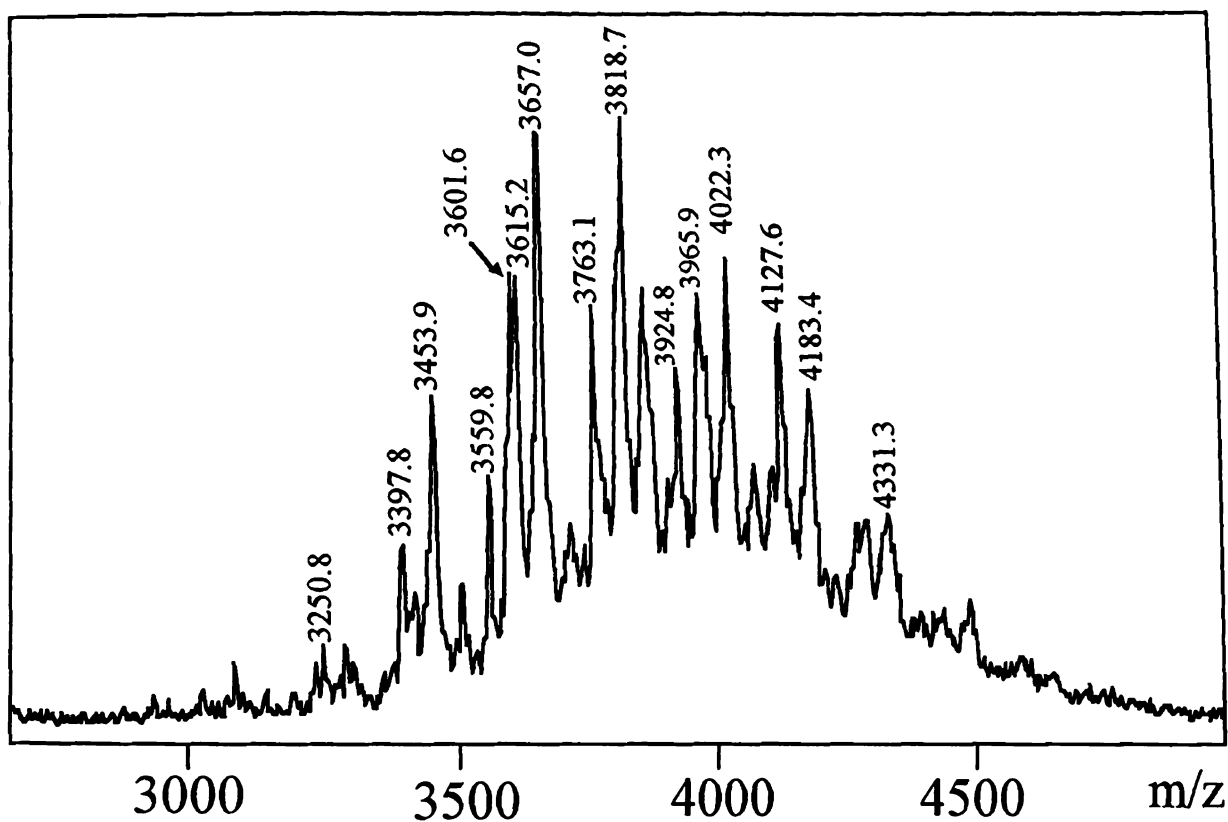
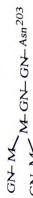


Figure 4.12: N-linked glycopeptides at Asn 203 of ZP3a glycoprotein. (a) MALDI-MS spectrum obtained on original HPLC fraction. (b) MALDI-MS spectrum obtained after β -mannosidase & N -acetylhexosaminidase treatment.

Proposed structure of the Glycopeptide	Calculated m/z	Observed m/z	Proposed structure of the Glycopeptide	Calculated m/z	Observed m/z
---	---------------------	-------------------	---	---------------------	-------------------



Proposed structure of the glycopeptide	Calculated m/z	Observed m/z	Proposed structure of the glycopeptide	Calculated m/z	Observed m/z
	3251.4	3250.8		3820.0	3818.7
	3454.6	3453.9		4023.2	4022.3
	3616.8	3615.2		4185.3	4183.4
	3657.8	3657.0			

Figure 4.13: N-linked glycopeptides (unfucosylated structures) at Asn 203 of ZP3a glycoprotein. Observed m/z and calculated m/z were assigned to the $[M+Na]^+$ ions of the glycopeptides.

Proposed structure of the glycopeptide	Calculated m/z	Observed m/z	Proposed structure of the glycopeptide	Calculated m/z	Observed m/z
---	---------------------	-------------------	---	---------------------	-------------------




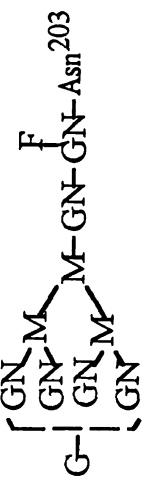
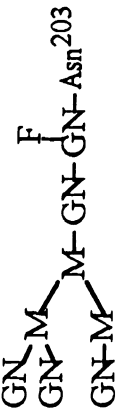
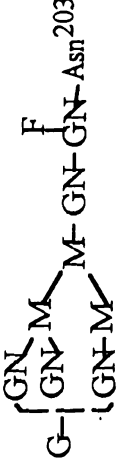
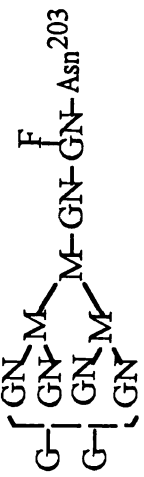

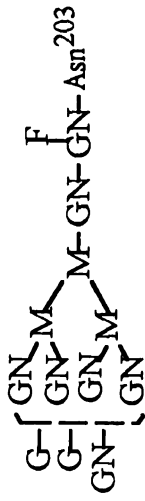
Proposed structure of the glycopeptide	Calculated m/z	Observed m/z	Proposed structure of the glycopeptide	Calculated m/z	Observed m/z
	3397.6	3397.8		3804.0	3805.2
	3559.7	3559.8		3966.1	3965.9
	3600.8	3601.6			
	3762.9	3763.1		4128.3	4127.6
	3925.1	3924.8		4331.4	4331.3

Figure 4.14: *N*-linked glycopeptides (fucosylated structures) at Asn 203 of ZP3a glycoprotein. Observed m/z and calculated m/z were assigned to the $[M+Na]^+$ ions of the glycopeptides.

Trypti
glycop
& S-ca

Figure 4.15: A mic

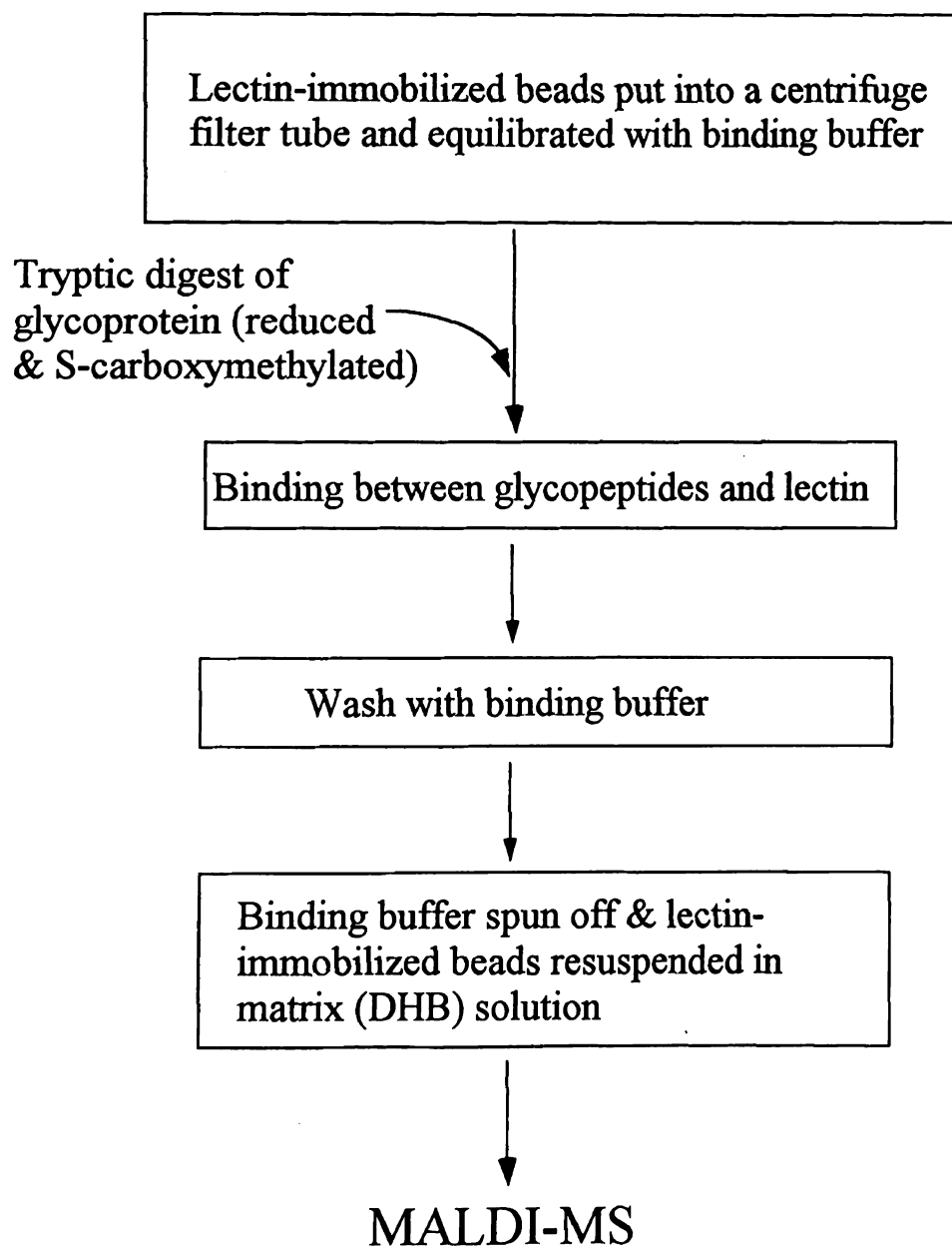


Figure 4.15: A micro-batch lectin affinity binding process for subsequent direct MALDI-MS analysis.

because of its acid

and the lectin so th

This appro

and avidin. Bovin

RNA and also ca

stranded DNA.

glycosylation site a

time and tissues

subunits, each of w

site at asparagine 1

in either avidin or R

Concanaval

and avidin because

Concanavalin A ha

Binding buffer for C

150 mM NaCl and

RNase B or avidin

process whereas ot

buffer.

Direct MAL

416. The glycopep

because of its acid nature, this acidity is able to disrupt the binding between glycopeptides and the lectin so that glycopeptide signal can be obtained by MALDI-MS.

This approach was used to identify glycopeptides derived from Ribonuclease B and avidin. Bovine pancreatic Ribonuclease B (RNase B) is a glycoprotein that cleaves RNA and also can destabilize or unwind the DNA helix by complexing with single-stranded DNA. RNase B comprises 150 amino acid residues and one *N*-linked glycosylation site at asparagine 60 (Asn 60). Avidin is a glycoprotein found in the egg-white and tissues of birds, reptiles and amphibians. Avidin contains four identical subunits, each of which comprises 128 amino acid residues and one *N*-linked glycosylation site at asparagine 17 (Asn 17). No *O*-linked carbohydrate side chains have been observed in either avidin or Ribonuclease B.

Concanavalin A-Sepharose was used to bind glycopeptides derived from RNase B and avidin because these glycopeptides consist high mannose-type oligosaccharides and Concanavalin A has high binding capability for terminal α -D-mannosyl residues. The Binding buffer for Concanavalin A was a sodium phosphate solution (pH 7.4) containing 150 mM NaCl and 1 mM MnCl_2 and 1 mM CaCl_2 . Glycopeptides in the tryptic digest of RNase B or avidin were retained by Concanavalin A-Sepharose beads during the binding process whereas other peptides in the tryptic digest were washed away by the binding buffer.

Direct MALDI-MS analysis of tryptic digest of RNase B was shown in Figure 16. The glycopeptide signals were seriously suppressed by signals of other peptides in

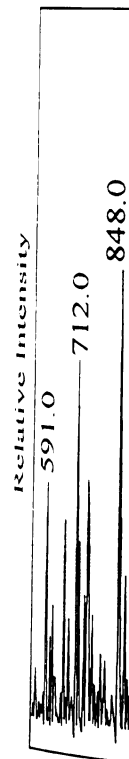


Figure 4.10

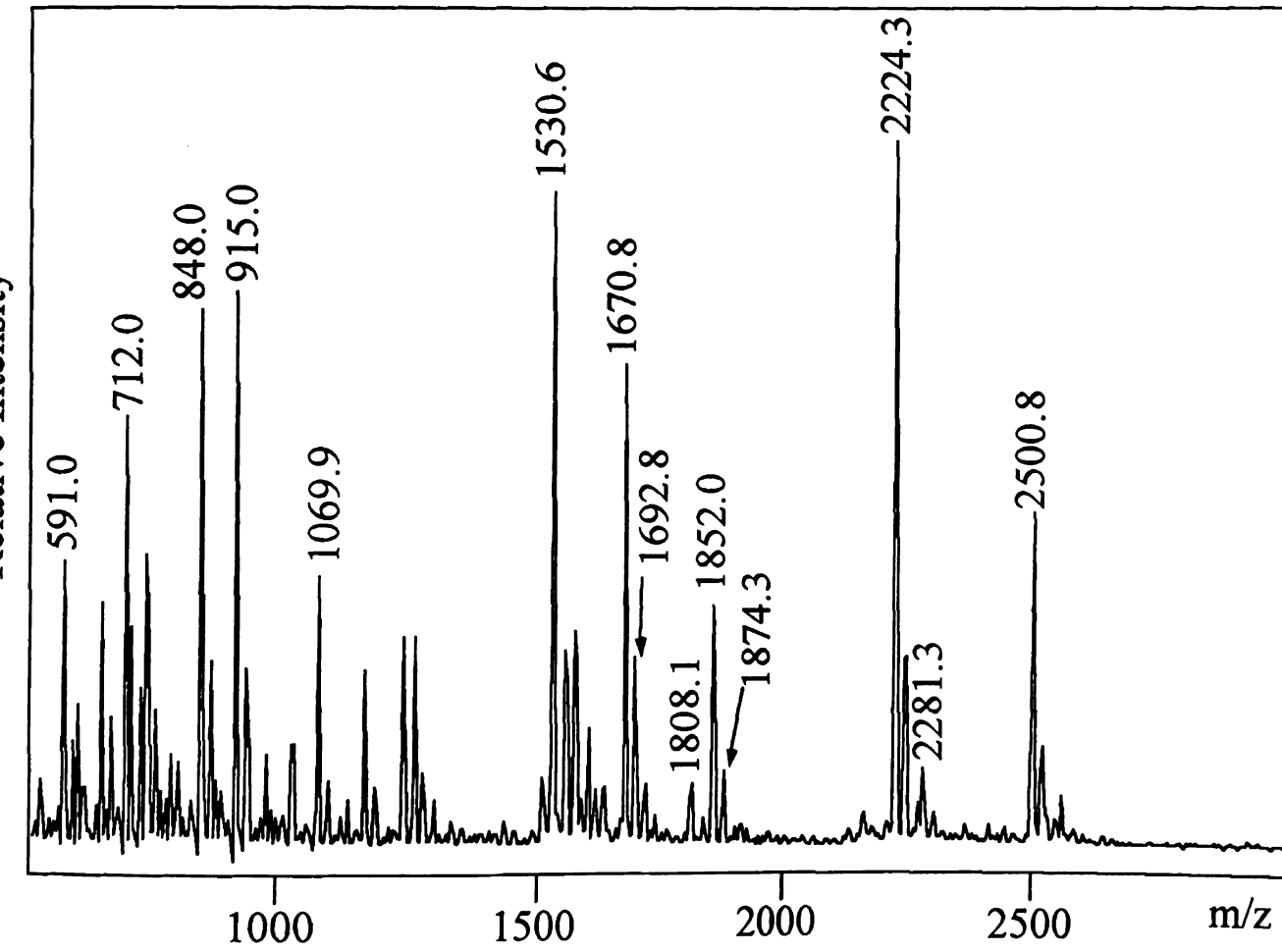


Figure 4.16: Direct MALDI-MS analysis of tryptic digest of ribonuclease B.

the sample. MALDI

the micro-batch

were predominant

the tryptic digest v

ed 2263.2 repres

residues in the oli

1756.7, 1898.7, 20

peptides. Dir

showed the presen

micro-batch lectin

MALDI mass spect

the sample. MALDI-MS analysis of glycopeptides from the tryptic digest of RNase B after micro-batch lectin (Con A) binding is shown in Figure 4.17. Glycopeptide signals were predominant in the MALDI mass spectrum while most other peptide components in the tryptic digest were not detected. The peaks at m/z 1714.8, 1876.9, 2039.1, 2201.6, and 2263.2 represented the $[M+Na]^+$ ions of glycopeptides with 5, 6, 7, 8, 9 mannose residues in the oligosaccharide structure, respectively. Another series of peaks at m/z 1736.7, 1898.7, 2061.8, 2223.9, and 2385.4 represented the $[M+2Na-H]^+$ ions of these glycopeptides. Direct MALDI-MS analysis of the tryptic digest of avidin (Figure 4.18) showed the presence of the glycopeptides in the region from m/z 3000 to m/z 3700. After micro-batch lectin (Con A) binding, these glycopeptides are more clearly visible in the MALDI mass spectrum (Figure 4.19).



Figure 4.17: MA
nuclease B after

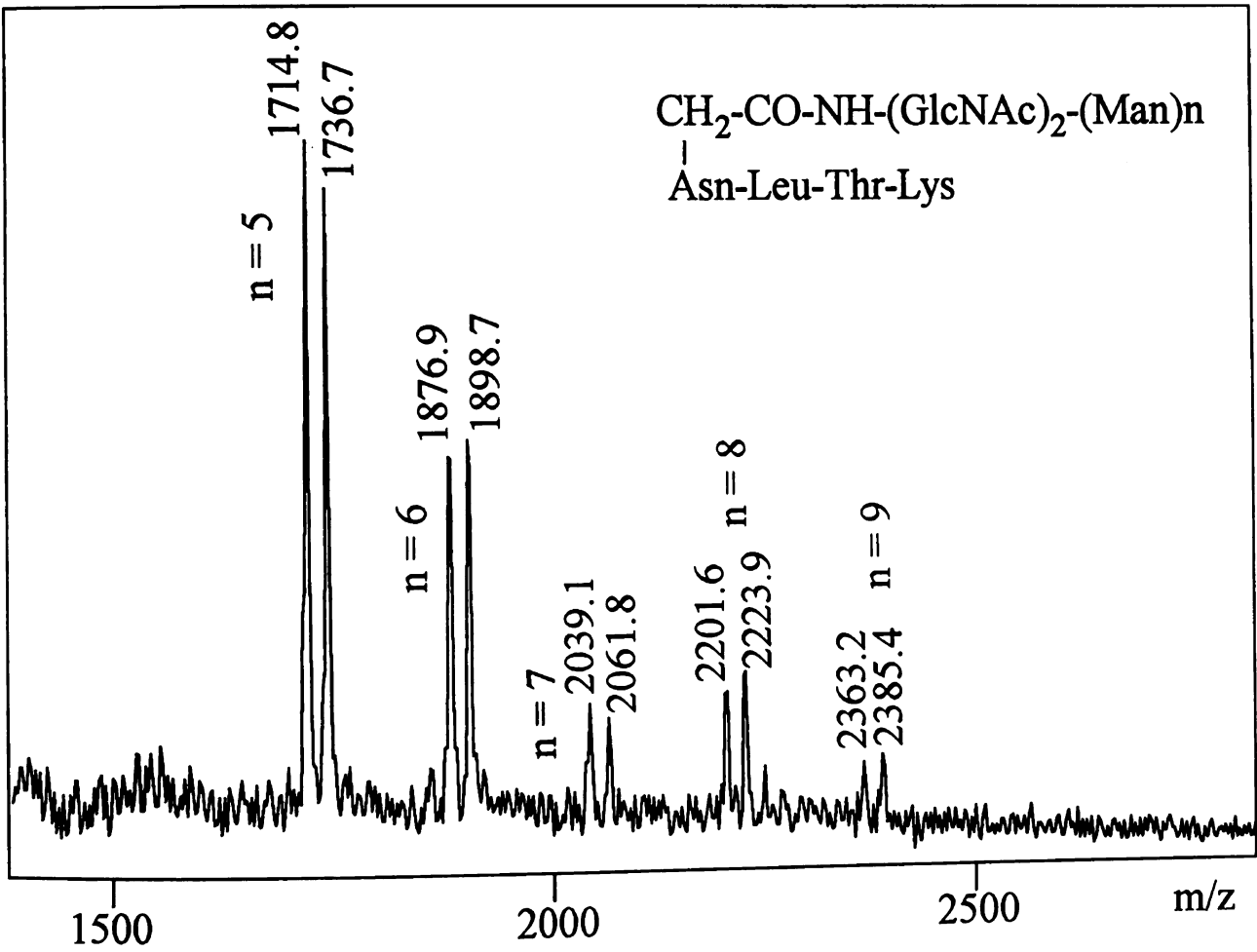
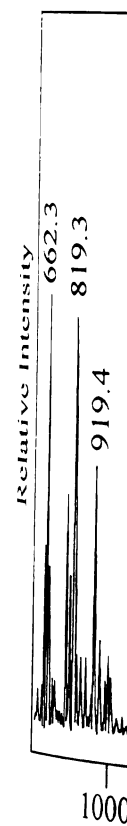


Figure 4.17: MALDI-MS Analysis of Glycopeptides from the Tryptic Digest of α-mannosidase B after Micro-batch Lectin (ConA) Affinity Binding.



Figure

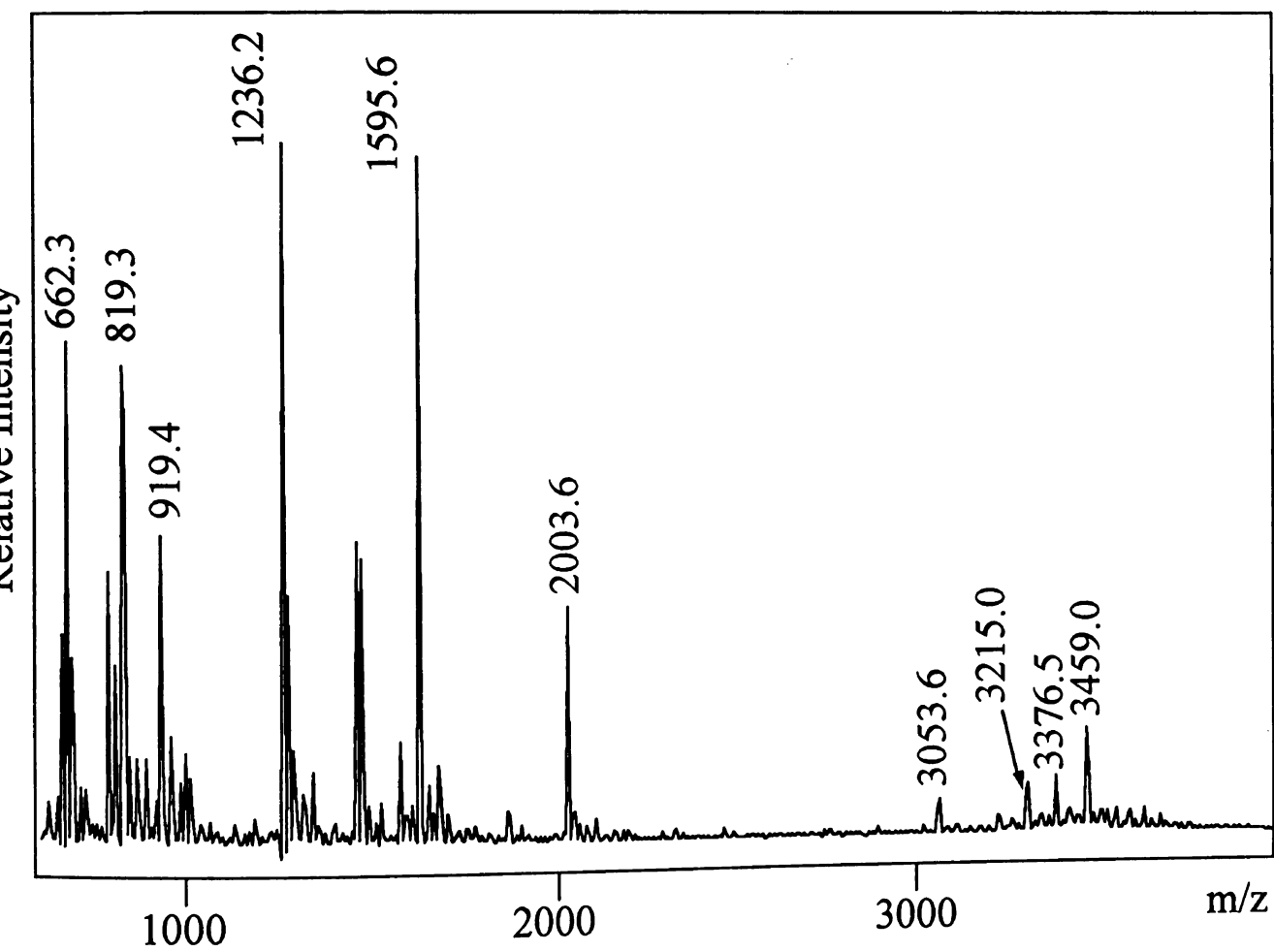


Figure 4.18: Direct MALDI-MS analysis of tryptic digest of Avidin.

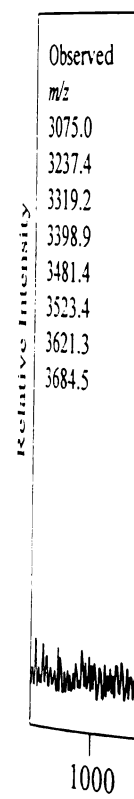


Figure 4.19: MALDI-TOF mass spectrum of the sample after micro-batch leaching. The peak at m/z 3684.5 is assigned to the $[M+H]^+$ ion.

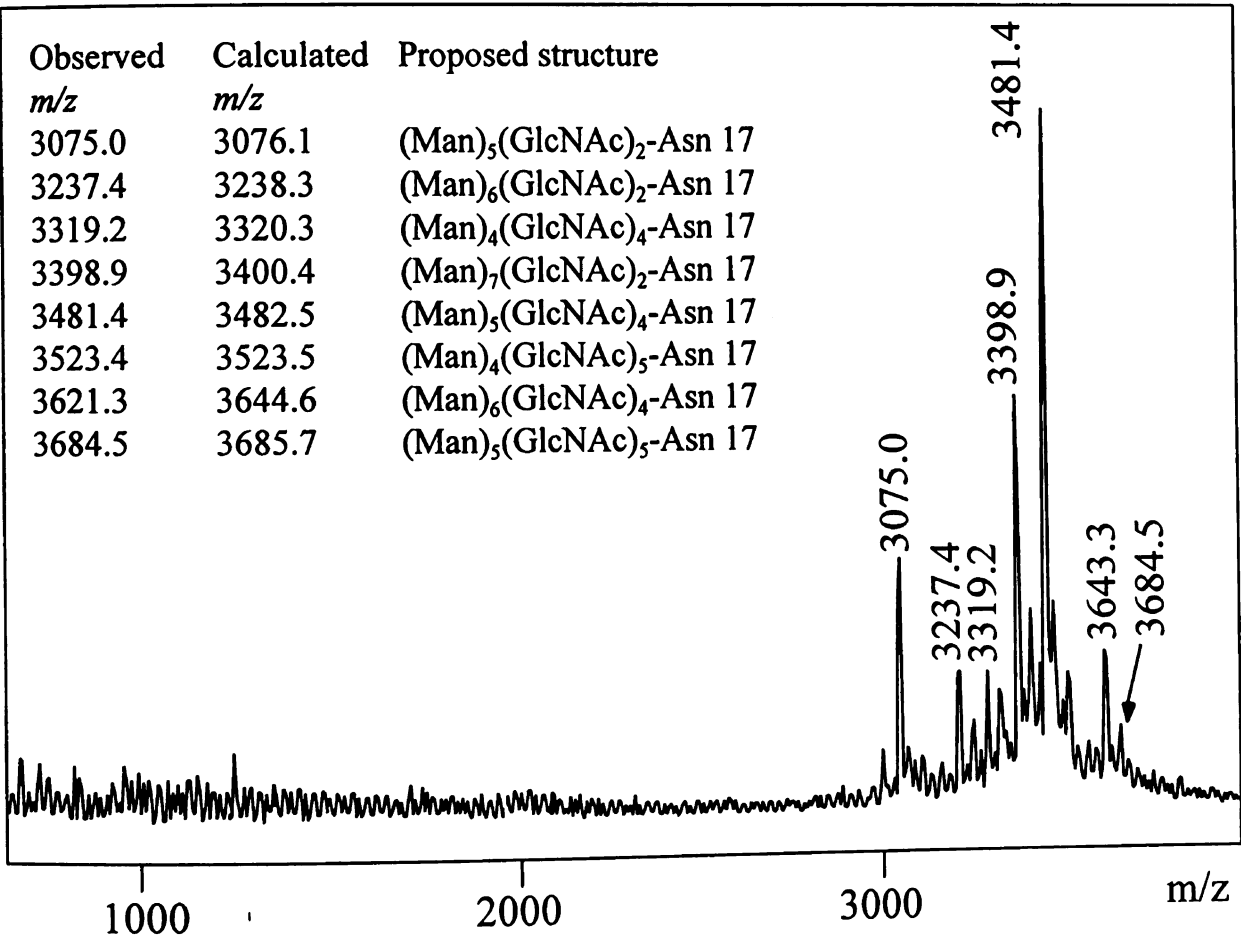


Figure 4.19: MALDI-MS analysis of glycopeptides from the tryptic digest of avidin or micro-batch lectin (ConA) affinity binding. Observed *m/z* and calculated *m/z* were assigned to the [M+Na]⁺ ions of the glycopeptides.

4. Conclusions

The tryptic digest
glycopeptide fraction
MALDI-MS spectra
distinguished by le
spectrum of the gl
that of the glyco
exoglycosidases v
MALDI-MS analy
and unfucosylated
Asn333 is compos
oligosaccharides.
binding and MAL
direct MALDI-MS
identifying glycope
glycosylation site.

4. Conclusions

The tryptic digest of ZP3 α glycoprotein was separated by reverse phase HPLC and four glycopeptide fractions were identified by their microheterogeneity as revealed in their MALDI-MS spectra. *N*-linked and *O*-linked glycopeptides of ZP3 α glycoprotein can be distinguished by lectin affinity chromatography / HPLC / MALDI-MS. The MALDI mass spectrum of the glycopeptide at Asn 203 of ZP3 α showed more microheterogeneity than that of the glycopeptide at Asn 220 or Asn 333. After the treatment with specific exoglycosidases which release most of the outer branches of these glycopeptides, MALDI-MS analysis revealed that the glycan at Asn 203 consists a mixture of fucosylated and unfucosylated bi- and tri-antennary structures, whereas the glycan at Asn 220 or Asn333 is composed of a mixture of fucosylated complex bi-, tri-, and tetra-antennary oligosaccharides. To minimize sample handling between the stages of lectin affinity binding and MALDI-MS, a micro-batch lectin binding process was also developed for direct MALDI-MS analysis for isolated glycopeptides. This approach is capable of identifying glycopeptides derived from ribonuclease B and avidin, proteins having a single glycosylation site.

References:

1. Cumming, D.
2. Dell, A. and R.
3. Sastry, M. V. I.
4. Young, N. M.,
Biophys., 270,
5. Green, E. D.,
Chem., 263, 18
6. Edge, C. J., R
Wing, D. R., and
7. Conboy, J. J. and
8. Hsi, K. L., Che
198, 238-245, 1
9. Mortz, E., Sar
1109-1118, 199
10. Yang, Y. and O
11. Treuheit, M. J.,
12. Huberty, M. C.,
1993.
13. Carr, S. A., Hu
14. Huddleston, M.
1993.

References:

1. Cumming, D. A., *Glycobiology*, 1, 115-130, 1991.
2. Dell, A. and Reason, A. J., *Curr. Opin. Biotechnol.* 4, 52-56, 1993.
3. Sastry, M. V. K. and Surolia, A., *Biosci. Repots*, 6, 853-860, 1986.
4. Young, N. M., Johnston, R. A. Z., Szabo, A. G., and Watson, D. C., *Arch. Biochem. Biophys.*, 270, 596-603, 1989.
5. Green, E. D., Adelt, G., Baenziger, J. U., Wilson, S., and van Halbeek, H., *J. Biol. Chem.*, 263, 18253-18268, 1988.
6. Edge, C. J., Rademacher, T. W., Wormald, M. R., Parekh, R. B., Butters, T. D., Wing, D. R., and Dwek, R. A., *Proc. Natl. Acad. Sci. USA*, 89, 6338-6342, 1992.
7. Conboy, J. J. and Henion, J. D., *J. Am. Soc. Mass Spectrom.* 3, 804-814, 1992.
8. Hsi, K. L., Chen, L., Hawke, D. H., Zieske, L. R., and Yuan, P. M., *Anal. Biochem.*, 198, 238-245, 1991.
9. Mortz, E., Sareneva, T., Julkunen, I., and Roepstorff, P., *J. Mass Spectrom.*, 31, 1109-1118, 1996.
10. Yang, Y. and Orlando, R., *Rapid Commun. Mass Spectrom.*, 10, 932-936, 1996.
11. Treuheit, M. J., Costello, C. E., and Halsall, H. B., *Biochem. J.*, 283, 105-112, 1992.
12. Huberty, M. C., Vath, J. E., Yu, W., and Martin, S. A., *Anal. Chem.*, 65, 2791-2799, 1993.
13. Carr., S. A., Huddleston, M. J., and Bean, M. F., *Protein Science*, 2, 183-196, 1993.
14. Huddleston, M. J., and Bean, M. F., and Carr., S. A., *Anal. Chem.*, 65, 877-884, 1993.

15. Schindler, P.

Protein Science

16. Liu, J., Volk,

Chromatogr.,

17. Fu, D. and Ha

18. Sutton, C. W.,

19. Mizuochi, T.,

7404-7409, 19

20. Yamamoto, T.

21. West, I. And C

22. Endo, T., *J. C*

23. Ogata, S., Mur

24. Ogata, S., Mur

25. Narasimhan, S.

26. Taniguchi, T.,

Chem., 261, 17

27. Yamashita, K.,

260, 4688-4693

28. Kornfeld, S., R

1981.

29. Baenziger, J. U.

30. Yamashita, K., I

15. Schindler, P. A., Settineri, C. A., Collet, X., Fielding, C. J., and Burlingame, A. L., *Protein Science*, 2, 791-803, 1995.
16. Liu, J., Volk, K. J., Kerns, E. H., Klohr, S. E., Lee, M. S., and Rosenberg, I. E., *J. Chromatogr.*, 632, 45-56, 1993.
17. Fu, D. and Halbeek, H. V., *Anal. Biochem.*, 206, 53-63, 1992.
18. Sutton, C. W., O'Neil, J. A., and Cottrell, J. S., *Anal. Biochem.*, 218, 34-46, 1994.
19. Mizuochi, T., Yonemasu, K., Yamashita, K., and Kotaba, A., *J. Biol. Chem.*, 253, 7404-7409, 1978.
20. Yamamoto, T. Tsuji, and T. Osawa, *Mol. Biotech.*, 3, 25-36, 1995.
21. West, I. And Golding, O., *Mol. Biotech.*, 2, 147-155, 1994.
22. Endo, T., *J. Chromatogr.*, 720, 251-261, 1996.
23. Ogata, S., Muramatsu, T., and Kobata, A., *J. Biochem. (Tokyo)*, 78, 687-696, 1975.
24. Ogata, S., Muramatsu, T., and Kobata, A., *Nature*, 259, 580-582, 1976.
25. Narasimhan, S., Freed, J. C., and Schachter, H., *Carbohydr. Res.*, 149, 65-83, 1986.
26. Taniguchi, T., Adler, A. J., Mizuochi, T., Kochibe, N., and Kobata, A., *J. Biol. Chem.*, 261, 1730-1736, 1986.
27. Yamashita, K., Kochibe, N., Ohkura, T., Ueda, I., and Kobata, A., *J. Biol. Chem.*, 260, 4688-4693, 1985.
28. Kornfeld, S., Reitman, M. L., and Kornfeld, R., *J. Biol. Chem.*, 256, 6633-6640, 1981.
29. Baenziger, J. U. and Fiete, D., *J. Biol. Chem.*, 254, 9795-9799, 1979.
30. Yamashita, K., Hitoi, A., and Kobata, A., *J. Biol. Chem.*, 258, 14753-14756, 1983.

31. Irimura, T., Ts
560-566, 1981
32. Cummings, R.
33. Yamashita, K.,
Biol. Chem., 26
34. Yoshima, H., M
Chem., 256, 84
35. Taniguchi, T.,
Kobata, A., *Bio*
36. Parekh, R. B.,
EMBO J., 6, 12
37. Hortin, G. L., A
38. Hortin, G. L. and
39. Wassarman, P.
C., Mortillo, S.
Mammalian Egg
40. Dunbar, B. S., V
41. Greve, J. M., S
759, 1982.
42. Shimizu, S., Tsu
43. Mori, E., Taka
Biochemistry, 30

31. Irimura, T., Tsuji, T., Tagami, S., Yamamoto, K., and Osawa, T., *Biochemistry*, 20, 560-566, 1981.
32. Cummings, R. D. and Kornfeld, S., *J. Biol. Chem.*, 257, 11230-11234, 1982.
33. Yamashita, K., Totani, K., Ohkura, T., Takasaki, S., Goldstein, I. J., Kobata, A., *J. Biol. Chem.*, 262, 1602-1607, 1987.
34. Yoshima, H., Matsumoto, A., Mizuochi, T., Kawasaki, T., and Kobata, A., *J. Biol. Chem.*, 256, 8476-8484, 1981.
35. Taniguchi, T., Mizuochi, T., Beale, M., Dwek, R. A., Rademacher, T. W., and Kobata, A., *Biochemistry*, 24, 5551-5557, 1985.
36. Parekh, R. B., Tse, A. G. D., Dwek, R. A., Williams, A. F., and Rademacher, T. W., *EMBO J.*, 6, 1233-1244, 1987.
37. Hortin, G. L., *Anal. Biochem.*, 191, 262-267, 1990.
38. Hortin, G. L. and Trimpe, B. L., *Anal. Biochem.*, 188, 271-277, 1990.
39. Wassarman, P. M., Bleil, J., Fimiani, C., Florman, H., Greve, J., Kinloch, R., Moller, C., Mortillo, S., Roller, R., Salzmann, G., Vazquez, M., In Dietl, J. (ed): "The Mammalian Egg Coat.", Berlin:, Springer-Verlag, pp 18-37, 1989.
40. Dunbar, B. S., Wardrip, N. J., Hedrick, J. L., *Biochemistry*, 19, 356-365, 1980.
41. Greve, J. M., Salzmann, G. S., Roller, R. J., and Wassarman, P. M., *Cell*, 31, 749-759, 1982.
42. Shimizu, S., Tsuji, M., Dean, J., *J. Biol. Chem.*, 258, 5858-5863, 1983.
43. Mori, E., Takasaki, S., Hedrick, J. L., Wardrip, N. J., Mori, T., Kobata, A., *Biochemistry*, 30, 2078-2087, 1991.

4. Yurewicz, E.

1987.

5. Yurewicz, E.

molecular and

1989

6. Yurewicz, E. C.

7. Yurewicz, E. C.

44. Yurewicz, E. C., Sacco, A. G., Subramanian, M. G., *J. Biol. Chem.*, 262, 564-571, 1987.
45. Yurewicz, E. C., Sacco, A. G., Subramanian, M. G., in Hedrick, J. L. (ed): "The molecular and cellular biology of fertilization.", New York, Plenum Press, pp 407-427, 1989.
46. Yurewicz, E. C., Pack, B. A., Sacco, A. G., *Mol. Reprod. Dev.*, 30, 126-134, 1991.
47. Yurewicz, E. C., Pack, B. A., Sacco, A. G., *Mol. Reprod. Dev.*, 33, 182-188, 1992.

1. Introduction

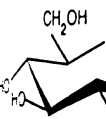
The complex
usually requires
Oligosaccharides c
Linkages between
second sugar unit).
known, most of w
Fuc), xylose (Xyl
and N-acetylneuram
commonly found in
Sequence an
great interest for m
collisionally induce
past for the structur
the oligosaccharide
cleavage of the glyc
the sugar ring have
Figure 5.2, when th
fragments are design

Chapter 5. A New Glycosylamine Derivative for Sequence Analysis of Oligosaccharides by MALDI-PSD-MS and ESI-MS

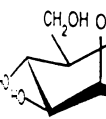
1. Introduction

The complete characterization of glycoproteins, glycolipids, and polysaccharides usually requires the structural analysis of their respective oligosaccharides (1-2). Oligosaccharides consist of monosaccharide units jointed together by glycosidic bonds (linkages between C(1), the anomeric carbon, of one sugar unit and an OH group of a second sugar unit). About 80 different kinds of naturally occurring glycosidic linkages are known, most of which involve mannose (Man), glucose (Glc), galactose (Gal), fucose (Fuc), xylose (Xyl), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), and *N*-acetylneuraminic acid (NeuAc, sialic acid). The structures of monosaccharides commonly found in eukaryotic glycoproteins and glycolipids are illustrated in Figure 5.1.

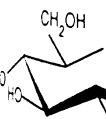
Sequence analysis of oligosaccharides by mass spectrometry has been a topic of great interest for many years. Fast atom bombardment mass spectrometry (FAB-MS) and collisionally induced dissociation (CID) MS/MS have been primarily used in the recent past for the structural analysis of oligosaccharides (3-4). The simplest fragmentation of the oligosaccharides occurring during FAB-MS and FAB-MS/MS results from the cleavage of the glycosidic bonds. More complex processes involving the fragmentation of the sugar ring have been observed, particularly in CID-MS/MS spectra. As illustrated in Figure 5.2, when the charge is retained on the non-reducing end of the oligosaccharides, fragments are designed as A_i , B_i and C_i , where i represents the position of the glycosidic



α -D-Glucose



α -D-Mannose



N-Acetyl- α -D-glucose



β -D-Xylose

Figure 5.1: Structure and glycolipids.

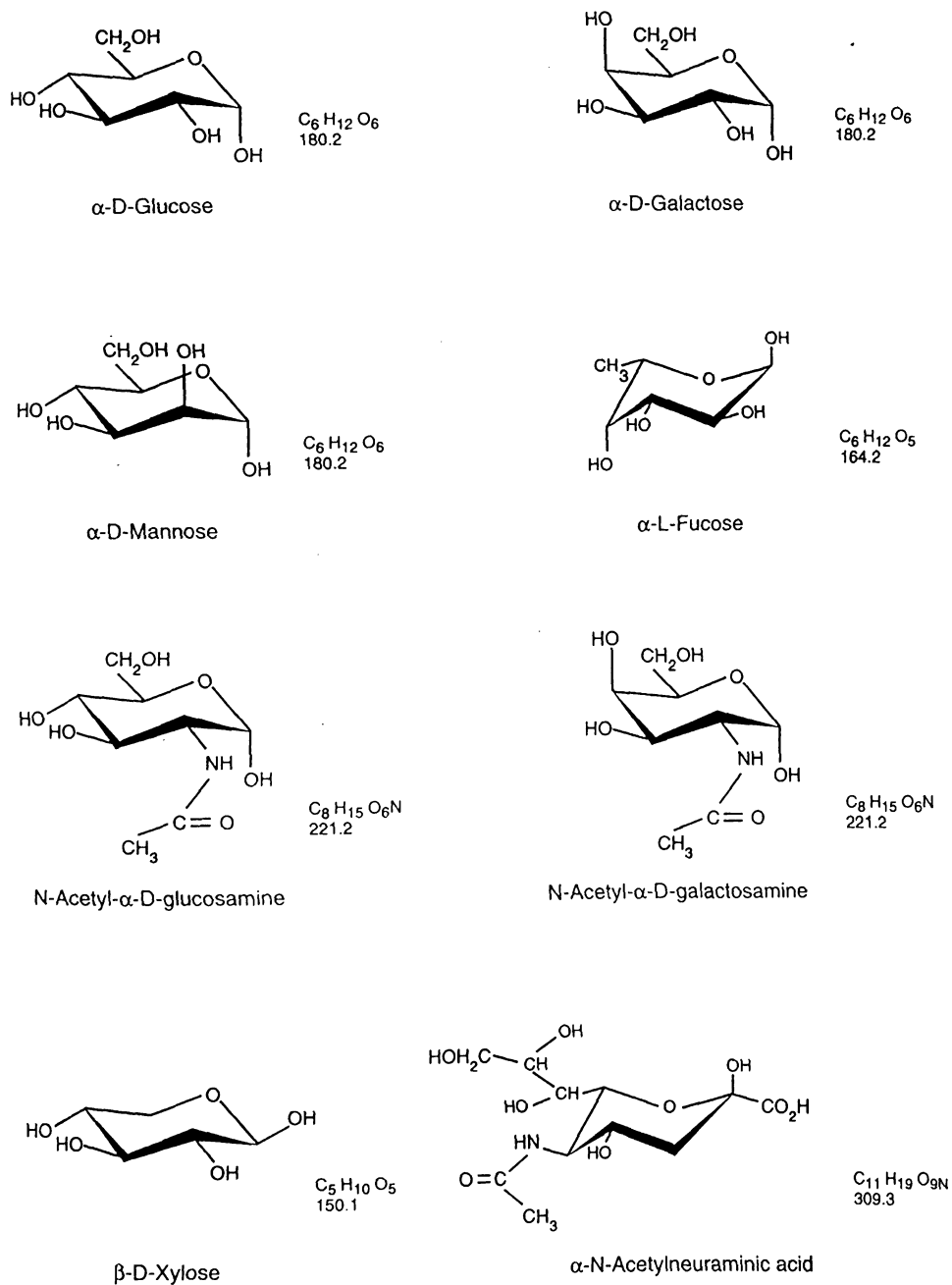
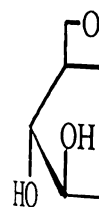


Figure 5.1: Structures of monosaccharides commonly found in eukaryotic glycoproteins and glycolipids.

bond cleaved, cou

reducing sugar un

bond counted from



0,2A

R

As an altern
two powerful ioniza
(6) and electrospra
determination of olig
analyzers, has been c
of oligosaccharides f
9) At an early sta
structural analysis be

bond cleaved, counted from the non-reducing end. On the other hand, ions containing the reducing sugar unit are labeled as X_j , Y_j , and Z_j , where j is the number of the glycosidic bond counted from the reducing end (5).

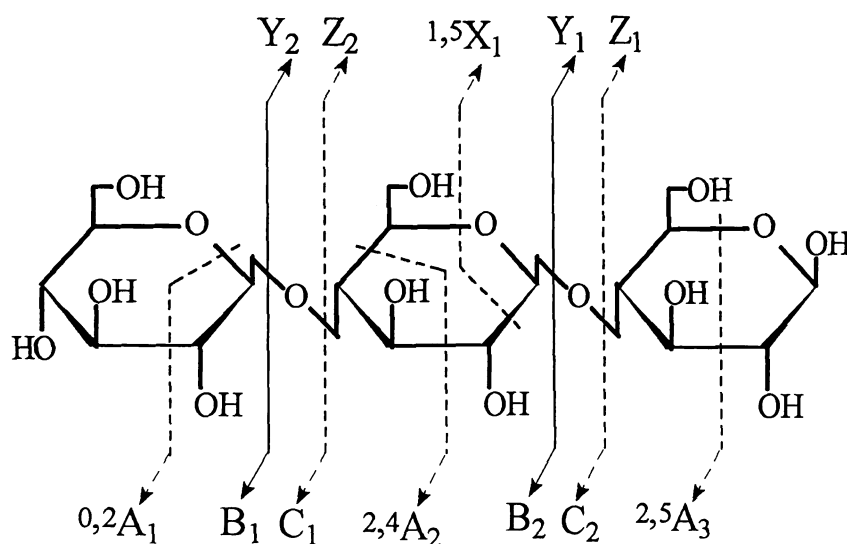


Figure 5.2: Types of oligosaccharide fragmentation.

As an alternative to fast atom bombardment desorption ionization (FAB) (3-4), two powerful ionization techniques, matrix-assisted laser desorption ionization (MALDI) (6) and electrospray ionization (ESI) (7) provide means for more sensitive structural determination of oligosaccharides. MALDI, used typically with time-of-flight (TOF) mass analyzers, has been demonstrated to be a promising tool to determine the molecular mass of oligosaccharides from purified sample or complex mixtures with great sensitivity (6, 8-9). At an early stage of its development, MALDI TOF-MS lacked the capacity for structural analysis because relatively little or no fragmentation was observed in the spectra,

although the com

be valuable in de

extension of MAL

molecular weight

roduced early in

a portion of the

electron-equippe

technique for stru

ionized ions of

conditions, althoug

and intensive inter

may complicate or e

In addition

characterization of n

source fragmentation

(9) It was also re

allow unambiguous

in ESI may be ineffic

A variety of

previously for impro

details prior to mass

reducing-end conjug

glycosylation (25-27)

although the combination of exoglycosidase digestion and MALDI-MS has been shown to be valuable in determining of the primary structure of oligosaccharides (10). A recent extension of MALDI TOF-MS allows structural information to be obtained in addition to molecular weight data, by means of so-called post-source-decay (PSD) analysis (11-13) as introduced early in Chapter 1. With this technique, fragments formed by decomposition of a portion of the intact ionized molecules in the field-free region can be detected in reflectron-equipped instruments. Spengler *et al.* have demonstrated the feasibility of PSD technique for structural analysis of carbohydrates (14). They reported that sodium cationized ions of native oligosaccharides undergo extensive fragmentation under PSD conditions, although reducing and non-reducing terminal fragments, cross-ring fragments and intensive internal fragments were all produced from native oligosaccharides, which may complicate or even compromise sequence assignment (14-15).

In addition to MALDI-MS, ESI-MS has been used successfully for the characterization of macromolecules including carbohydrates (16-19). ESI MS/MS and in-source fragmentation have been extended to the oligosaccharide structural analysis (17-19). It was also reported that fragment ions from native carbohydrates do not always allow unambiguous determination of the sequence, and further that underivatized samples in ESI may be inefficiently ionized (17-19).

A variety of derivatization methods of oligosaccharides have been developed previously for improving sensitivity for HPLC separation and investigation of structural details prior to mass spectrometric experiments (20-27). Most of these methods employ reducing-end conjugation which involves either reductive amination (21-24) or direct glycosylation (25-27) depending on the conjugating reagents. Both types of derivatives

provide a basis
possibly direct
HPLC separa
spectrometry.

In this
new glycosyla
reducing end v
carbohydrate a
useful for facil
MS.

2. Experiment

Materials

Lacto-N
Sigma Chemical
and LS-tetrasac
NY). 3-Amino
from Aldrich C
without further
sample was a gif
The cationic
(μm) was purch

provide a basic moiety which serves as a good protonation site during ionization and may possibly direct fragmentation after protonation. All these derivatization products require HPLC separation, and sometimes, extra desalting procedures before analysis by mass spectrometry.

In this chapter, we describe the preparation and mass spectral characteristics of a new glycosylamine derivative of oligosaccharides formed by direct glycosylation at the reducing end with 3-aminoquinoline (3AQ), a matrix material which has been reported for carbohydrate analysis in MALDI-MS (28). This new type of derivative was found to be useful for facile sequence analysis of oligosaccharides under MALDI-PSD-MS and ESI-MS.

2. Experimental

Materials

Lacto-N-tetraose (LNT), maltohexaose (G6), maltoheptaose were purchased from Sigma Chemical Co. (St. Louis, MO, USA). Lacto-N-fucopentaoses (LNFP-1, LNFP-2) and LS-tetrasaccharide c (LSTc) were purchase from Oxford Glycosystems (Rosedale, NY). 3-Aminoquinoline (3AQ) and 2,5-dihydroxybenzoic acid (DHB) were purchased from Aldrich Chemical Co. (Milwaukee, WI). These commercial materials were used without further purification. Cyclamen seed xyloglucan triantennary oligosaccharide sample was a gift from Dr. C. Hervé du Penhoat (*Département de Chimie, E.N.S., France*).

The cation-exchange resin (i.e., analytical grade 200-400 mesh AG 50W-X8 in H⁺ form) was purchased from Bio-Rad Laboratories (Hercules, CA). Cation-exchange resin

was converted

Biemann (2)

Derivatization

Oligo

with 25 μ l of

of 200 mg/ml

μ l acetic acid

Sample Preparation

One

evaporated to

in 100 mg/

Approximate

delivered to

and used directly

beads left on

"blown off" v

HPLC of 3A

HPLC

PC computer

chromatograph

was converted to the NH_4^+ form according to the procedure outlined by Wang and Biemann (29).

Derivatization of Oligosaccharides with 3-Aminoquinoline (3AQ)

Oligosaccharide samples (500 pmol to 500 nmol) were dissolved in reaction vials with 25 μl of a 8:2 mixture of methanol and water. To these solutions were added 25 μl of 200 mg/ml 3-aminoquinoline (3AQ) solution freshly prepared in the same solvent and 1 μl acetic acid. The reaction mixture was heated at 80 °C for 3-4 h.

Sample Preparation for MALDI-MS

One μl of the reaction mixture was deposited on a sample plate well and evaporated under vacuum. The NH_4^+ form cation-exchange resin beads were suspended in 100 mg/ml DHB matrix solution which was dissolved in 8:2 methanol/water. Approximately 1 μl of such matrix solution which typically contains 10-20 resin beads was delivered to the target by a pipette tip. The sample plate was then dried under vacuum and used directly for the MALDI-MS analysis and MALDI-PSD-MS analysis with resin beads left on the target. Resin beads can also be loosened with a microspatula and then "blown off" with a pressurized air stream as described by Rouse *et al.* (30).

HPLC of 3AQ-treated Oligosaccharides

HPLC analyses were carried out using Waters model 6000 pumps controlled by a PC computer with the detector operated at 284 nm. 3AQ-treated samples were chromatographed on a Phase Sep Spherisorb S5 amino column (25 cm \times 4.6 mm). The

solvent sys

linear gradi

for ESI-MS

Instrument

MA

Elite time-e

MA, USA)

delay time

PSD mode

selection w

mass select

were seque

reflectron v

depending

reach ± 0.2 -

ESI

spectromete

the capillar

Scanning co

pump (Hart

solvent system was run from 20% to 60% water-acetonitrile in 1 ml/min with a 30-minute linear gradient program. The derivatized oligosaccharide fraction was collected and used for ESI-MS or MALDI-MS.

Instrumentation

MALDI-MS experiments were carried out on the PerSeptive Biosystems Voyager-Elite time-of-flight (TOF) mass spectrometer (PerSeptive Biosystems Inc., Framingham, MA, USA). The ion source accelerating voltage was 20 kV in all TOF experiments. A delay time of 50 ns was used in the linear mode and a delay time of 100 ns was used in the PSD mode before the ion extraction in the source. In PSD experiments, precursor ion selection was performed by means of an electrostatically switched ion gate, which allows a mass selectivity of about 90 M/dM (1 out of 90 mass units). PSD fragment ion spectra were sequentially recorded over 8-10 mass windows by incremental reduction of the reflectron voltages. In each spectral window 30-100 single-shot spectra were averaged depending on the sample conditions. Accuracy and precision of PSD mass assignments reach ± 0.2 - 0.3 u in the mass range of m/z 300-1600.

ESI-MS experiments were performed on a Fisons VG *Platform* electrospray mass spectrometer with a single-quadrupole analyzer. The ESI voltage was set to 2.5 KV and the capillary temperature was 100 °C. The pressure of the sheath gas was 30 psi. Scanning covered the range m/z 300-1400 with a scan rate of 2 s/scan. A Harvard syringe pump (Hartford, CT, USA) was used in the infusion mode at a flow rate of 10 μ l/min.

Res
be labeled
according t

3. Results

Pre
reagent we
mixture, af
section, wa
spectrum (h
the derivati
[M+Na]⁺ i
1013.3 wa
current rat
than 90%
weight take

Lik
(25), the
aromatic ar
support the
structure ar
2) MALD

Resolution was sufficient so that all peaks in ESI and MALDI mass spectra could be labeled with their monoisotopic masses. Oligosaccharide fragment ions were labeled according to the nomenclature of Domon and Costello (31).

3. Results and Discussion

Preliminary conditions for derivatization with 3-aminoquinoline (3AQ) as the reagent were initialized using maltohexaose (G6) as a model compounds. The reaction mixture, after on-plate desalting with ion-exchange beads as described in the experimental section, was directly subjected to linear MALDI-MS analysis. The resulting MALDI mass spectrum (Figure 5.3) revealed a major peak at m/z 1117.4 representing the $[M+H]^+$ ion of the derivatized maltohexaose (3AQ-G6) and a minor peak at m/z 1139.4 representing the $[M+Na]^+$ ion of the product. The $[M+Na]^+$ ion of underivatized maltohexaose at m/z 1013.3 was also observed in Figure 5.3 with much lower intensity. Based on the ion current ratio of derivatized and unreacted maltohexaose the reaction yield was greater than 90% after 3 hours at 80 °C. It has been noted that samples of larger molecular weight take longer for complete reaction.

Like the N-(2-pyridinyl)-glycosylamines of carbohydrates studied by Her et al. (25), the derivative formed with 3-aminoquinoline is the dehydration product of an aromatic amine and the glycosyl hemiacetal (Figure 5.4). Observations in this work which support the glycosylamine product instead of Schiff base product or Amadori product structure are: 1.) 3AQ labeled glucose showed anomeric protons when studied by NMR; 2.) MALDI-MS molecular weight analysis of the peracetyl 3AQ-glucose derivative

indicated

acyclic S

hydrolyz

U

MALDI-

usually ic

molecule

reported

reductive

liquid-liq

the requi

The prote

advantag

The other

oligosacc

desalting

Na⁺, in di

resin lead

With the

species in

manuscrip

seed xylo

indicated the incorporation of four acetyl groups (five expected if the product was the acyclic Schiff base); 3.) the amino group of the 3AQ derivative of maltohexaose can be hydrolyzed by acetic acid. An Amadori product, if formed, is stable to acid hydrolysis.

Unlike peptides or proteins which are usually ionized as a protonated ion in MALDI-MS and ESI-MS, native neutral and acidic oligosaccharides lack basic groups and usually ionized by alkali metal cationization, especially in MALDI-MS. In fact, protonated molecules of native oligosaccharides are rarely detected in MALDI-MS. Lemoine *et al.* reported for the first time the protonated species of oligosaccharides in MALDI-MS after reductive amination with benzylamine (15). Nevertheless, such a protocol requires several liquid-liquid phase extractions and further reversed phase HPLC purification because of the requirement for excess reagents, especially sodium borohydride, used in the reaction. The protocol reported in this chapter allows direct analysis of glycosylamines by taking the advantage of the fact that excess reagent, 3AQ, can be used as part of the MALDI matrix. The other half of the matrix material, DHB, was used to improve the signal intensity of the oligosaccharide derivative. Cation-exchange resin was used in our protocol for on plate desalting of the derivatized products to control the variable amounts of metal ions, mainly, Na^+ , in different sample sources. In favorable cases, sample preparation without using the resin lead to the formation the protonated molecule of the 3AQ glycosylamine derivative. With the resin protocol, protonated molecules could always be generated as the dominant species in the MALDI mass spectrum of the other oligosaccharides investigated in the manuscript including LNT, maltoheptaose, LSTc, LNFP-1, LNFP-2, and the Cyclamen seed xyloglycan triantennary oligosaccharide.

Figure 5
maltohex

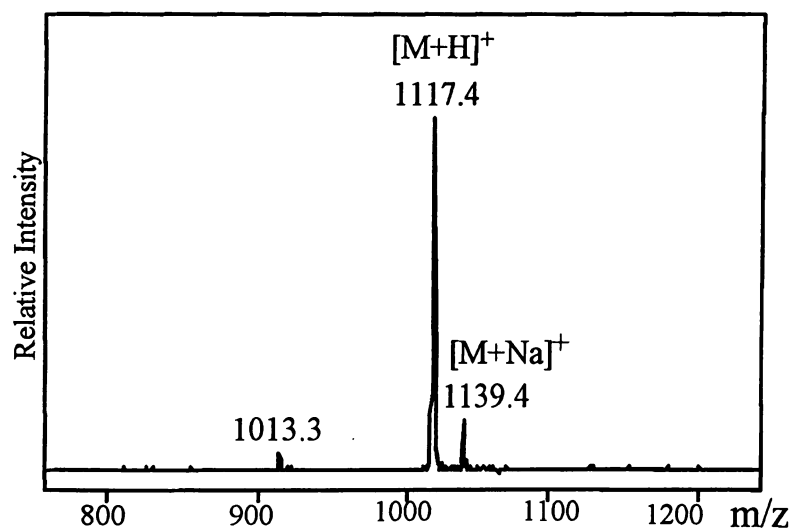
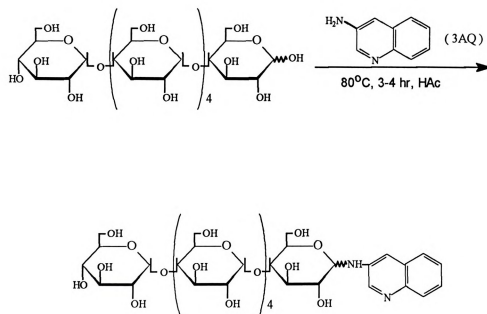


Figure 5.3: Linear positive MALDI-MS spectrum of 3-aminoquinoline (3AQ) derivatized maltohexaose (G6). The reaction mixture was used for direct MALDI-MS analysis.

H

Scheme

the mod



Scheme 5.4: 3-Aminoquinoline glycosylamine derivative formation using maltotriose as the model compound.

analys

non-re

intern

seque

from

which

ions o

likely

fragme

suitabl

MALL

1125.4

m/z 11

[M+H]

mass

ganglic

The ab

HexNA

the red

and m

bonds)

As reported before by Lemoine *et al.* (15) and Spengler *et al.* (14), MALDI-PSD analysis on $[M+Na]^+$ of native or derivatized oligosaccharide produced both reducing and non-reducing fragments. For native oligosaccharides, cross-ring fragments and intensive internal fragments were also present in the MALDI-PSD spectrum, which can complicate sequence assignment. Figure 5.5 shows the positive PSD spectrum of the $[M+H]^+$ ion from 3AQ-G6. A series of abundant, evenly spaced Y ions were observed in Figure 5.5 which allowed the sequence to be easily assigned. No significant fragments for B type ions or cross-ring cleavage were observed. In this case, the charge for $[M+H]^+$ ion is most likely to reside on the quinoline ring due to its high proton affinity and observed the fragmentation pattern.

The reaction conditions outlined in the experimental section were found to be suitable for acidic oligosaccharides like LSTc. Figure 5.6(a) presents the linear positive MALDI mass spectrum obtained from 3AQ derivatized LSTc. The peak of $[M+H]^+$ at m/z 1125.4 again dominated the spectrum accompanied by a relatively small $[M+Na]^+$ peak at m/z 1147.4. A peak resulting from the loss of *N*-acetylneuraminic acid (NeuAc) from the $[M+H]^+$ ion was observed at m/z 834.3. This type of neutral loss in the positive MALDI mass spectrum was also reported previously for NeuAc containing compounds like gangliosides (32-33). The PSD spectrum of the $[M+H]^+$ ion is shown in Figure 5.6(b). The abundant Y ion series determined the sequence of the molecule to be NeuAc-Hex-HexNAc-Hex-Hex. The cleavage of the HexNAc glycosidic bond with charge retention on the reducing end led to the formation of the B_3 ion at m/z 657.2. The ions at m/z 204.1 and m/z 366.1 corresponded to two internal fragments (i.e., double cleavage of glycosidic bonds) because of the presence of HexNAc moiety.

Figure

m/z

carbo

▲, Fu

accor

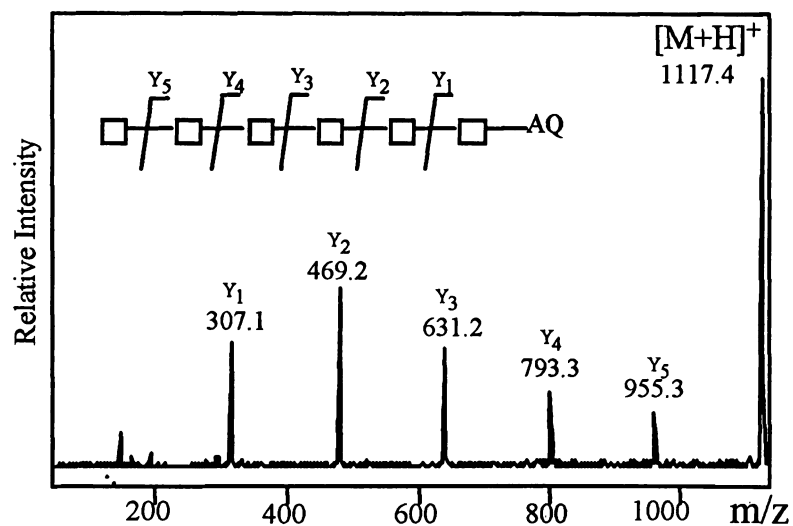


Figure 5.5: Positive ion MALDI-PSD mass spectrum of 3AQ-G6. The $[M+H]^+$ ion at m/z 1117.4 was selected as the precursor ion. The monosaccharide symbols in the carbohydrate structures in all figures are defined as follows: \square , Glc; \blacksquare , GlcNAc; \triangle , Xyl; \blacktriangle , Fuc; \circ , Gal; \bullet , NeuAc. Oligosaccharide fragment ions in all figures were labeled according to the nomenclature of Domon and Costello.

Figure

tetrasa

underi

[M+H]

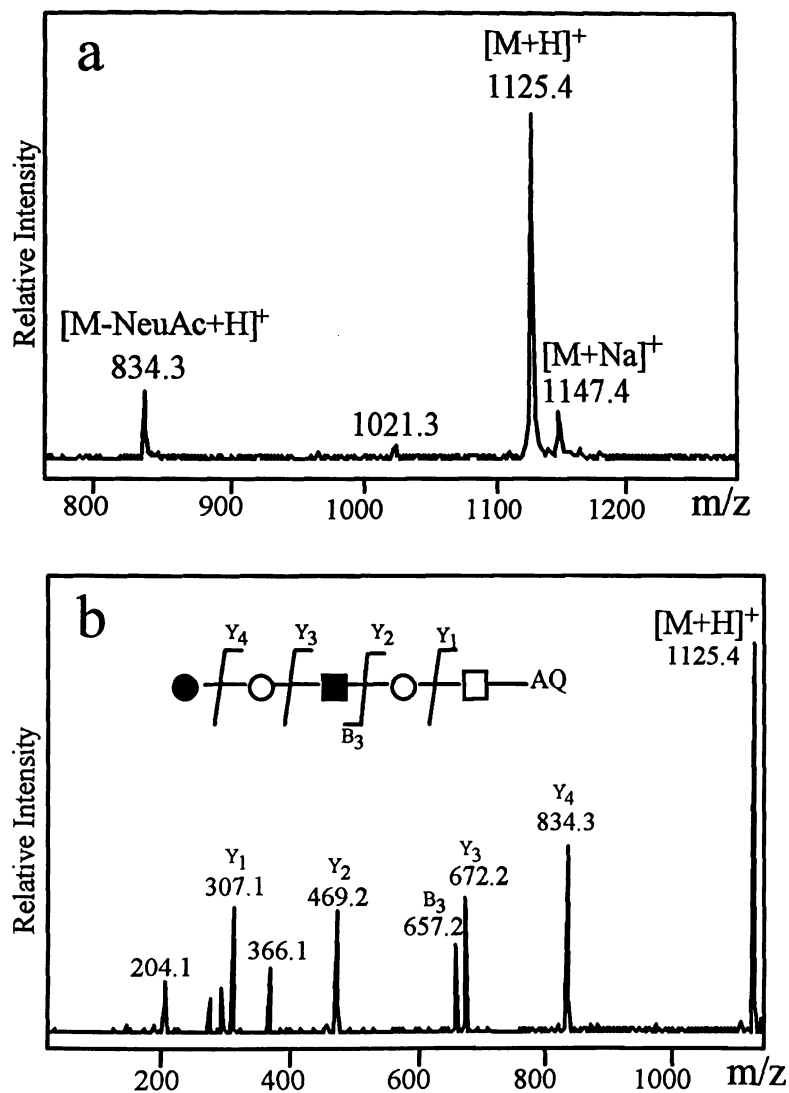


Figure 5.6: (a) Linear positive MALDI-MS spectrum of 3AQ-derivatized LS-tetrasaccharide c (LSTc). The small peak at m/z 1021.3 represented the $[M+\text{Na}]^+$ ion of underivatized LSTc. (b) Positive ion MALDI-PSD mass spectrum of 3AQ-LSTc with $[M+\text{H}]^+$ ion at m/z 1125.4 selected as the precursor ion. For symbols, see Figure 5.5.

2, al

[M+

isom

MAL

LNF

either

Figur

the lo

ion o

5.7(a

LSTc

analy

derivz

ion se

showr

were

on the

helped

m/z 13

residu

the pro

Differentiation of two isomeric fucosylated pentasaccharides, LNFP-1 and LNFP-2, also could be easily achieved by MALDI-PSD analysis after 3AQ glycosylation. The $[M+H]^+$ ion of either derivatized products had the same m/z value, 980.4. LNFP-1, the isomer with linear structure, was characterized by Y ion series from Y_0 to Y_4 in the MALDI-PSD spectrum as shown in Figure 5.7(a). The branching pattern of its isomer, LNFP-2, was clearly indicated by two adjacent Y_3 fragments corresponding to the loss of either one or the other non-reducing terminal residue at the branching site as shown in Figure 5.7(b). A small peak at m/z 672.3 was observed in Figure 5.7(b) corresponding to the loss of both non-reducing terminal residues. Because of the presence of HexNAc, the ion of m/z 204.1, the ion of m/z 366.2, and a B type ion were produced in both Figure 5.7(a) and Figure 5.7(b), which agrees with the fragmentation pattern observed for 3AQ-LSTc.

To test our protocol for larger and more complicated oligosaccharide structural analysis, a triantennary oligosaccharide derived from Cyclamen seed xyloglucan (27) was derivatized with 3-aminoquinoline and subjected to direct MALDI-PSD-MS analysis. Y ion series dominated the PSD spectrum of the $[M+H]^+$ ion of the resulting derivative as shown in Figure 5.8. Some Y ions such as $Y_{2\alpha/2\beta}$ at m/z 469.2 and $Y_{2\alpha/3\beta}$ at m/z 601.2 were produced from double cleavages of glycosidic bonds with the charge retention still on the reducing-end AQ group. These two Y ions at m/z 469.2 and m/z 601.2 actually helped to determine the position of an internal Xyl residue. Other two adjacent Y ions at m/z 1351.4 and m/z 1381.4 illustrated that there were two types of non-reducing terminal residues, Xyl and Hex, at different antenna. Other fragments are in good agreement with the proposed structure.

Figur

LNFP

were s

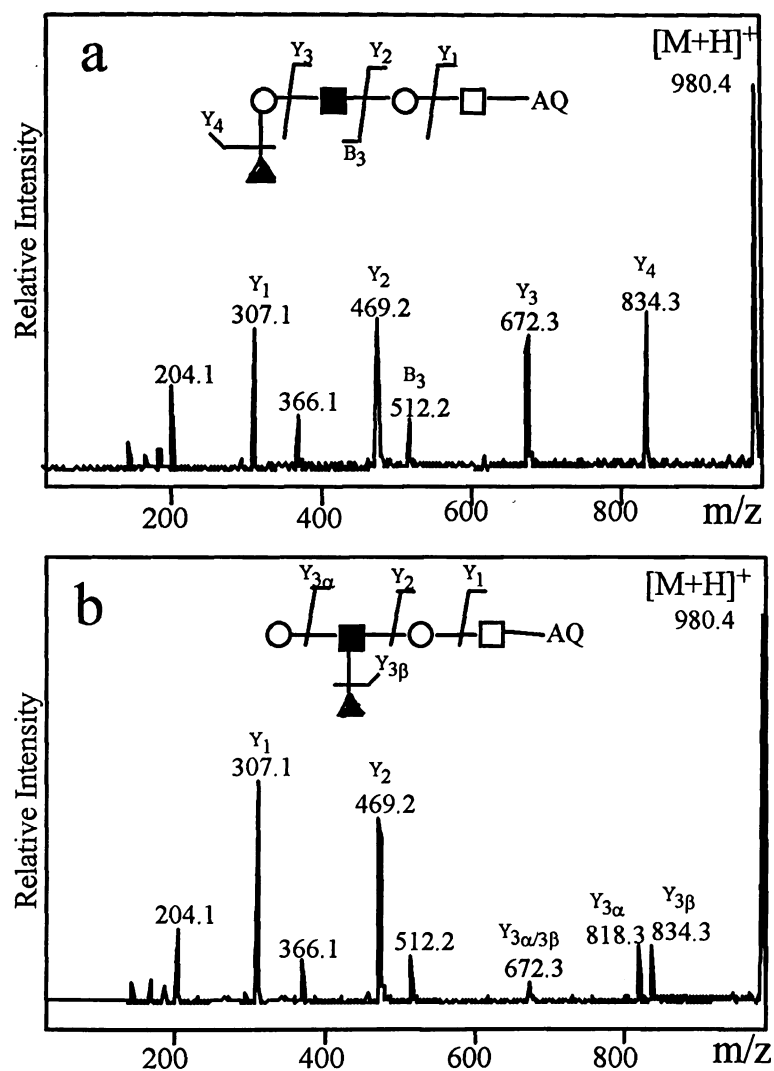


Figure 5.7: Positive ion MALDI-PSD mass spectrum of 3AQ-derivatized (a) LNFP-1 (b) LNFP-2. The $[M+H]^+$ ion of either derivatized isomer had same m/z value, 980.4, and were selected as the precursor ions for both analyses. For symbols, see Figure 5.5.

Relative Intensity

Figure

oligos

m/z 15

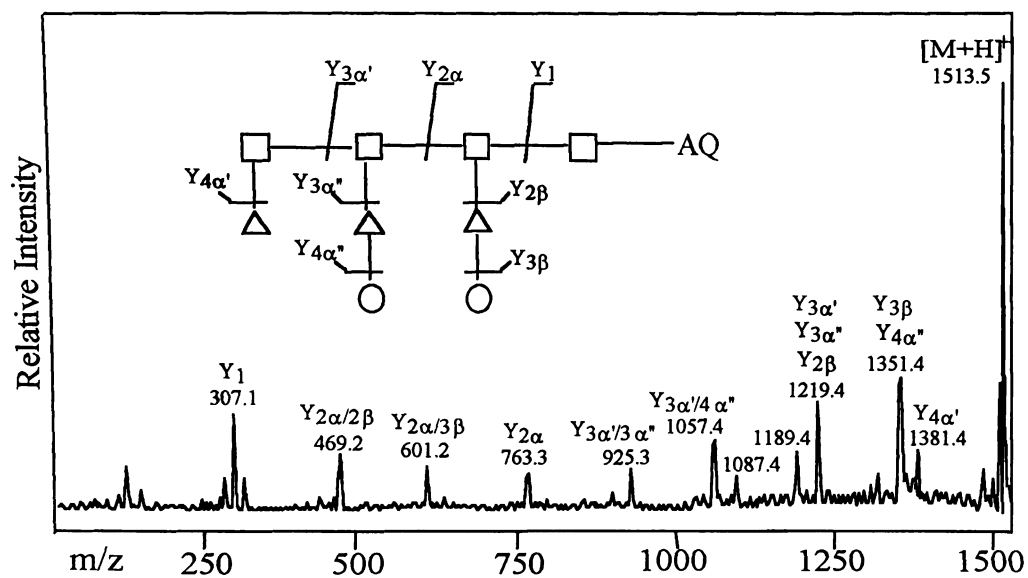


Figure 5.8: Positive ion MALDI-PSD mass spectrum of a 3AQ-derivatized triantennary oligosaccharide which was derived from Cyclamen seed xyloglucan. The $[M+H]^+$ ion at m/z 1513.5 was selected as the precursor ion. For symbols, see Figure 5.5.

of the
HPLC
Chrom
exists
molec
integr
found
and th
produ
studie
found
signs
acidic
solutio
tempe
after t
analys
that o
cone v
[M+H]
about

After the oligosaccharide has been labeled with the chromophore, 3AQ, separation of the resulting glycosylamine product from the reaction mixture can be achieved by HPLC as shown in Figure 5.9 in which derivatized maltohexaose was used as a model. Chromatographic separation revealed two isomeric forms of 3AQ-G6 because of the existence of two different positions (i.e., equatorial or axial position) of AQ group in the molecule. The yield of the total 3AQ-G6 obtained from this procedure was measured by integrating the peaks observed during the HPLC against an external standard and was found to be ~70%. The difference between the reaction yield from HPLC peak integration and that from MALDI-MS ion current integration may be due to the fact that derivatized product (3AQ-G6) had higher unit response than underivatized maltohexaose when studied by MALDI-MS. The AQ group of the HPLC-separated collected products was found to be resistant to hydrolysis under mildly basic or neutral conditions and showed no signs of hydrolysis after ten days in 50% aqueous acetonitrile solution. However, under acidic conditions the AQ group is gradually hydrolyzed. A 2% aqueous acetic acid solution brought about complete hydrolysis of the AQ group in one day at room temperature. This actually provides a way to regenerate the underivatized carbohydrate after the HPLC separation.

The fractions of 3AQ-G6 from HPLC were collected and subjected to ESI-MS analysis. The intensity of the $[M+H]^+$ ion was also found to be much more intense than that of $[M+Na]^+$ ion or $[M+K]^+$ ion. No significant fragmentation was observed when cone voltage (i.e., the capillary-skimmer voltage) was set at 20 V. Detection of the $[M+H]^+$ ion was achievable at the level of 1 pmol of total sample infused, which provides about a 50-fold increase in sensitivity compared to that for underivatized maltohexaose on

Fig

on a

water

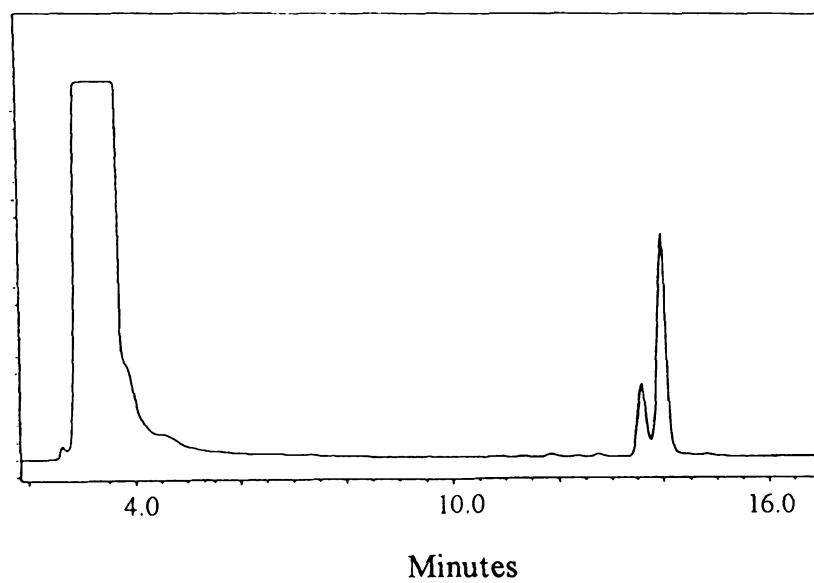


Figure 5.9: HPLC separation of 3-aminoquinoline (3AQ) derivatized maltohexaose (G6). on a Phase Sep Spherisorb S5 amino column. The system was run from 20% to 60% water-acetonitrile in 1 ml/min with a 30 minute linear gradient program.

the

occu

those

When

Y_0 to

the same ESI instrument. It was further noted that intensive in-source fragmentation occurred at elevated cone voltage for infused 3AQ-G6. The fragment ions resembled those observed in Figure 5.2, which were generated during analysis by MALDI-PSD-MS. When the cone voltage was raised to 100 V, the complete series of Y ions ranging from Y_0 to Y_5 were all produced as shown in Figure 5.10.

Fig

(cap

min

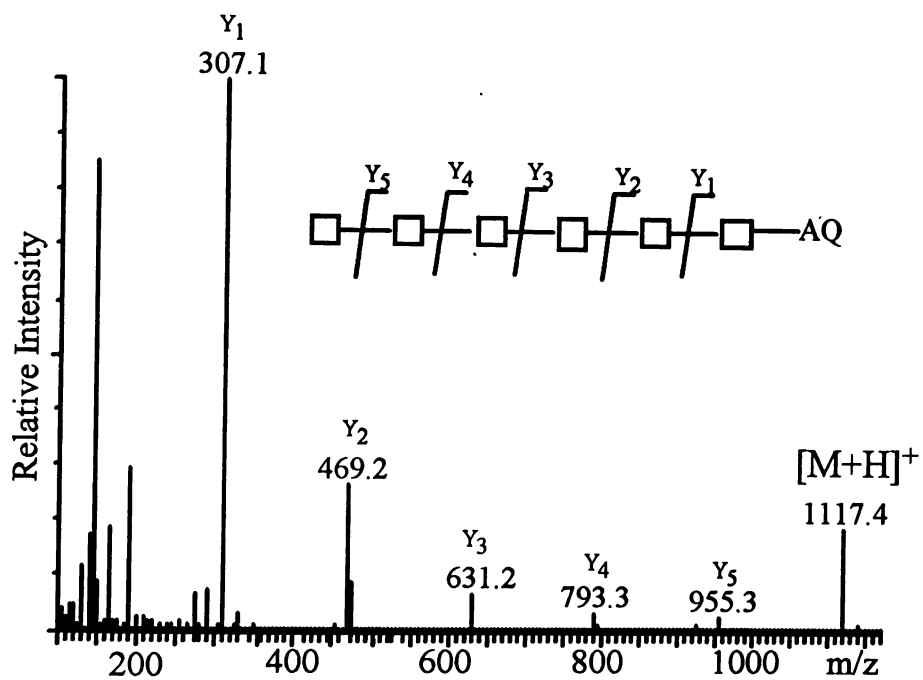


Figure 5.10: Positive ion ESI-MS spectrum of HPLC separated 3AQ-G6. Cone voltage (capillary-skimmer voltage): 100 V. The spectrum shown represented an average over 1-min recording time. For symbols, see Figure 5.5.

4.6

ana

spe

MS

deri

pro

form

ions

dete

reac

matr

HPL

the E

skim

analy

4. Conclusion

In this study a new glycosylamine derivative was developed for facile sequence analysis of oligosaccharides by matrix-assisted laser desorption ionization mass spectrometry (MALDI-PSD-MS) and electrospray ionization mass spectrometry (ESI-MS). 3-Aminoquinoline (3AQ) was used as the reagent for making glycosylamine derivatives at the reducing terminus. This type of derivative produces primarily protonated molecules for neutral and acidic oligosaccharides in MALDI-MS without the formation of significant $[M+Na]^+$ ions. Post-source decay (PSD) analysis of such $[M+H]^+$ ions yields a spectrum dominated by Y ions, allowing the sequence to be readily determined. Another attractive feature of this derivative in MALDI-MS analysis is that reaction cleanup is not necessary and excess reagent can be used as part of the MALDI matrix. Derivatized oligosaccharides also can be separated from the reaction mixtures by HPLC and subject to ESI-MS study. Protonated molecules are also dominant species in the ESI spectrum. Intensive in-source fragmentation was observed at elevated capillary-skimmer voltage. The fragmentation pattern was very similar to that observed during analysis by MALDI-PSD-MS.

Re

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

11.

12.

13.

References:

1. Dell, A. and Reason, A. J., *Curr. Opin. Biotechnol.* 4, 52, 1993.
2. Cumming, D. A., *Glycobiology*, 1, 115-130, 1991.
3. Barber, M., Bordoli, R. S., Elliot, G. J., Sedgwick, R. D., and Tyler, A. N., *Anal. Chem.*, 54, 645A-657A, 1982.
4. Carr, S. A., Reinhold, V. N., Green, B. N., and Hass, J. R., *Biomed. Mass Spectrom.*, 12, 288, 1985.
5. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T., *Anal. Chem.*, 63, 1193A-1203A, 1991.
6. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M., *Science*, 246, 64-71, 1989.
7. Stahl, B., Steup, M., Karas, M., and Hillenkamp, F., *Anal. Chem.*, 63, 1463-1466, 1991.
8. Stahl, B., Thurl, S., Zeng, J., Karas, M., Hillenkamp, F., Steup, M., and Sawatzki, G., *Anal. Biochem.*, 223, 218-226, 1994.
9. Harvey, D. J., Rudd, P. M., Bateman, R. H., Bordoli, R. S., Hayes, K., and Vickers, R. G., *Org. Mass Spectrom.*, 29, 753, 1994.
10. Sutton, W. C., O'Neil, J. A., and Cottrell, J. S., *Anal. Biochem.*, 218, 34, 1994.
11. Huberty, M. C., Vath, J. E., Yu, W., and Martin, S. A., *Anal. Chem.* 65, 2791-2799, 1993.
12. Kaufmann, R., Spengler, B., and Lützenkirchen, F., *Rapid Commun. Mass Spectrom.* 7, 902-910, 1993.
13. Kaufmann, R., Kirsch, D., and Spengler, B., *Int. J. Mass Spectrom. Ion Processes* 131, 355-385, 1994.

14.

15.

16.

17.

18.

19.

20.

21.

22.

23.

24.

25.

26.

14. Spengler, B., Kirsch, D., Kaufmann, R., and Lemoine, J., *J. Mass Spectrom.*, 30, 782-787, 1995.
15. Lemonie, J., Chirat, F., and Domon, B., *J. Mass Spectrom.*, 31, 908-912, 1996.
16. Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J., and Udseth, H. R., *Anal. Chem.*, 62, 882, 1990.
17. Yoshino, K., Takao, T., Murata, H., and Shimonishi, Y., *Anal. Chem.*, 67, 4028-4031, 1995.
18. Gu, J., Hiraga, T., and Wada, Y., *Biomed. Mass Spectrom.*, 23, 212-217, 1994.
19. Okamoto, M., Takahashi, K-I, and Doi, T., *Rapid Commun. Mass Spectrom.* 9, 641-643, 1995.
20. Domon, B., Mueller, D. R., and Richter, W. J., *Biomed. Mass Spectrom.*, 19, 390, 1990.
21. Hase, T., Ibuki, and T. Ikennaka, *J. Biochem.*, 95, 197, 1984.
22. Dell, A., Carman, N. H., Tiller, P. H., and Thomas-Oates, J. E., *Biomed. Mass Spectrom.*, 16, 19-24, 1988.
23. Wang, W. T., LeDonne, N. C., Jr., Ackerman, B., and Sweeley, C. C., *Anal. Biochem.*, 141, 366-381, 1984.
24. Poulter, L. and Burlingame, A. L., in *Methods Enzymol.* (J. A. McCloskey, Eds.), Academic Press, San Diego, 193, 661-689, 1990.
25. Her, G. R., Santikarn, S., Reinhold, V. N., and William, J. C., *J. Carbohydr. Chem.*, 6, 129-139, 1987.
26. Caesar, J. P. Jr., Sheeley, D. M., and Reinhold, V. N., *Anal. Biochem.*, 191, 247-252, 1990.

27.

28.

29.

30.

31.

32.

33.

27. Braccini, I., Hervé du Penhoat, C., Michon, V., Goldberg, R., Clochard, M., Jarvis, M. C., Huang, Z.-H., and Gage, D. A., *Carbohydr. Res.*, 276, 167-181, 1995.
28. Metzger, J. O., Woisch, R., Tuszynski, W., Angermann, R., *Fresenius J. Anal. Chem.*, 349, 473-475, 1994.
29. Wang, B. H. and Biemann, K., *Anal. Chem.*, 66, 1918, 1994.
30. Rouse, J. C., and Vath, J. E., *Anal. Biochem.*, 238, 82-92, 1996.
31. Domon, B., and Costello, C. E., *Glycoconj. J.*, 5, 397-409, 1988.
32. Juhasz, P. and Costello, C. E., *J. Am. Soc. Mass Spectrom.*, 3, 785-796, 1992.
33. Xu, N., Huang, Z.-H., Watson, J. T., and Gage, D. A., *J. Am. Soc. Mass Spectrom.*, 8, 116-124, 1997.

(N

ter

(1

a

de

(P

mo

TC

cha

app

disc

dist

des

indi

exce

mer

the

conv

Chapter 6. Summary

The development of matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) coincides with an increasing demand for more accurate and sensitive techniques in fields such as biochemistry, molecular biology, genetics and biotechnology (1). The MALDI time-of-flight (TOF) mass spectrometer has been quickly recognized as a "molecular weight machine" because of its superior capability in molecular mass determination of a wide range of biopolymers. The introduction of the post-source decay (PSD) technique allowed further structural information to be obtained, in addition to molecular weight data, on reflectron-equipped TOF instrument. In this research, MALDI TOF-MS, as well as the PSD technique, has been used as the primary tool for structural characterization of glycoconjugates including glycoproteins.

The success of analysis by MALDI-MS, in part, depends on the choice of appropriate matrix material. 2-Mercaptobenzothiazole (MBT) and its analogs have been discovered as a class of new matrices for MALDI-MS. These compounds are structurally distinct from the conventional matrices. MBT has been used successfully for the desorption of proteins up to 100,000 daltons. A comparison with conventional matrices indicates that MBT provides the same level of sensitivity and resolution and offers excellent tolerance to sample contaminants such as ionic detergents. 5-Chloro-2-mercaptobenzothiazole (CMBT), an analog of MBT, has been found not only effective for the analysis of peptides, low-mass proteins, and glycolipids, but also superior to conventional matrices for the analysis of glycopeptides and oligosaccharides.

at

w

s:

m

M

c

d

d

ir

a

c

a

M

S

w

w

ly

p

p

re

th

The information gained from the matrix studies has been applied to the structural analysis of peptidoglycans, a type of complex glycoconjugates consisting bacterial cell wall with important immunological significance. In analysis by MALDI-MS, peptidoglycan samples were subjected to enzymatic digestion with muramidase and the resulting muropeptides were purified by HPLC. CMBT was employed for routine analysis by MALDI-MS of muropeptides, which are problematic analytes in MALDI-MS by using conventional matrices. The results have demonstrated that sub-picomole to femtomole detection can be achieved in both positive mode and negative mode allowing unambiguous determination of molecular masses of muropeptides of peptidoglycans. Structural information of muropeptide monomers was obtained by further post-source decay (PSD) analysis. Fragmentation patterns in positive mode and negative mode PSD are found to be complementary to each other for the elucidation of the composition of the peptide chains and oligosaccharide moieties. Enzymatic digestion was also incorporated with MALDI-MS in solving structural problems of muropeptide oligomers from pathogenic *Staphylococcus aureus* strains. Specific cleavage between pentaglycine bridge residues was performed by lysostaphin digestion for these oligomers and the resulting hydrolysates were subjected to direct MALDI-MS. The combination of an analysis by PSD and the lysostaphin digestion approach led to the determination of structural modification of peptidoglycans from *Staphylococcus aureus* strains.

MALDI-MS also played an important role in the study of porcine oocyte zona pellucida 3 α glycoprotein (ZP3 α). Tryptic digest of ZP3 α glycoprotein was separated by reverse phase HPLC and four heterogeneous glycopeptide fractions were identified by their MALDI-MS spectra. *N*-linked and *O*-linked glycopeptides of ZP3 α glycoprotein

w

M

m

tr

gl

of

A

an

af

fo

id

gl

re

re

co

cr

oli

of

Ar

the

pre

reg

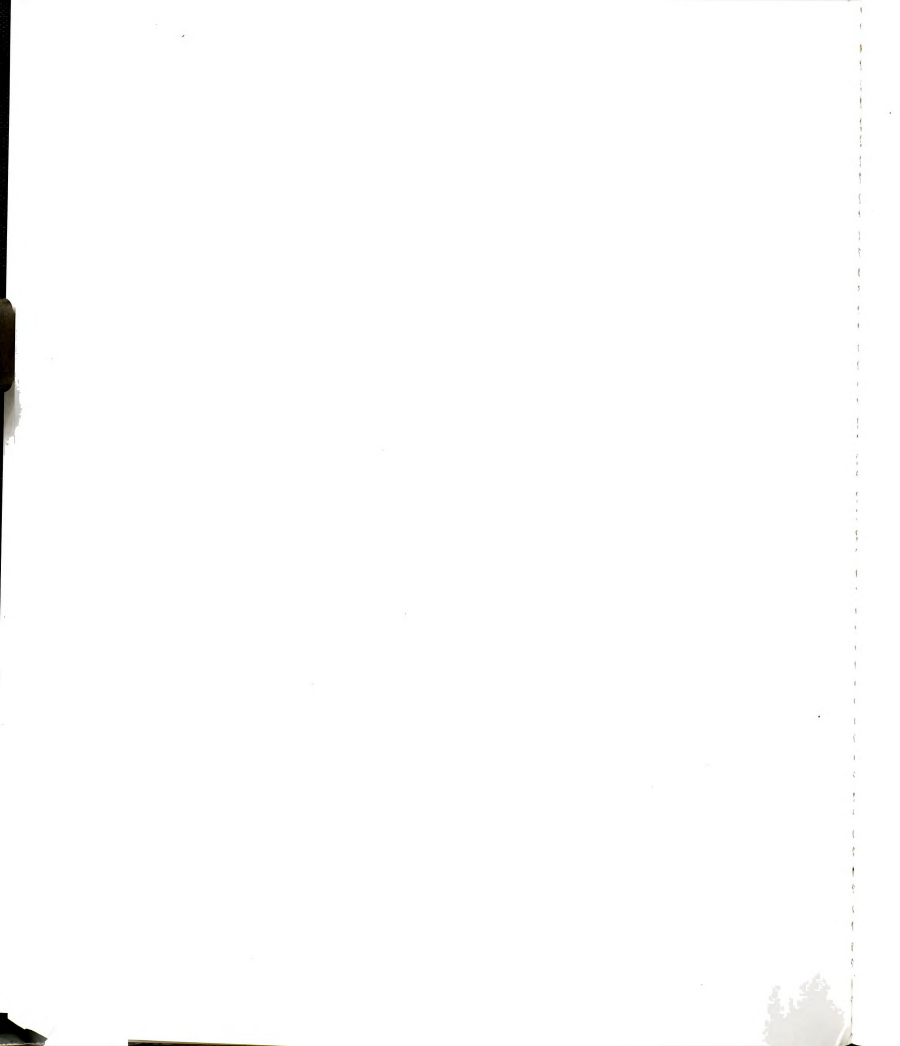
were distinguished by lectin (jacalin) affinity chromatography / HPLC / MALDI-MS. The MALDI mass spectrum of the glycopeptide at Asn 203 of ZP3 α showed more microheterogeneity than that of the glycopeptide at Asn 220 or Asn 333. After the treatment with specific exoglycosidases which released most of the outer branches of these glycopeptides, MALDI-MS analysis revealed that the glycan at Asn 203 consists a mixture of fucosylated and unfucosylated bi- and tri-antennary structures, whereas the glycan at Asn 220 or Asn333 is composed of a mixture of fucosylated complex bi-, tri-, and tetra-antennary oligosaccharides. To minimize sample handling between the stages of lectin affinity binding and MALDI-MS, a micro-batch lectin binding process was also developed for direct MALDI-MS analysis for isolated glycopeptides. This approach is capable of identifying glycopeptides derived from ribonuclease B and avidin, proteins having a single glycosylation site.

The complete characterization of glycoproteins and other types of glycoconjugates requires the structural analysis of their respective oligosaccharides. However, the resulting spectra from MALDI-PSD analysis on native oligosaccharides are rather complex and difficult to interpret. Reducing and non-reducing terminal fragmentation, cross-ring fragmentation and internal fragmentation all occur to the native oligosaccharides, complicating sequence determination. A new glycosylamine derivative of oligosaccharides has been developed for facile sequence analysis in MALDI-MS. 3-Aminoquinoline was used as the derivatizing reagent in making glycosylamine derivatives at the reducing terminus. This type of glycosylamine derivative produces primarily protonated molecules in MALDI-MS. Reaction cleanup is not necessary because excess reagent can be used as the MALDI matrix without interfering the analysis by MALDI-MS.

PSD analysis of the protonated molecule yields a spectrum dominated by reducing terminal fragments allowing easy determination of the oligosaccharide sequence.

References:

1. Siuzdak, G., *Proc. Natl. Acad. Sci. U.S.A.*, 91, 11290-11297, 1994.



MICHIGAN STATE UNIV. LIBRARIES



31293015630951