This is to certify that the

dissertation entitled

I DERIVATIVES FOR ISOLATION, PURIFICATION AND MASS SPECTROMETRI CHARACTERIZATION OF HOMO AND HETERO OLIGOSACCHARIDES

II STRUCTURE ELUCIDATION AND SYNTHETIC TRANSFORMATION RELATED TO THE INNER AND OUTER CORE OF THE LIPOPOLYSACCHARIDES

presented by

YUANDA ZHANG

has been accepted towards fulfillment
of the requirements for
Ph.D CHEMISTRY
degree in

Major professor

Date 01/30/93

PART I: DERIVATIVES FOR THE ISOLATION, PURIFICATION AND MASS SPECTROMETRIC CHARACTERIZATION OF HOMO AND HETERO OLIGOSACCHARIDES

PART II: STRUCTURE ELUCIDATION AND SYNTHETIC TRANSFORMATIONS RELATED TO THE INNER AND OUTER CORE OF THE LIPOPOLYSACCHARIDES OF *RHIZOBIUM MELILOTI* 41 AND *LEGUMINOSARUM* BIOVAR *VICIAE* VF-39 MUTANTS

By

Yuanda Zhang

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Chemistry

1992

ABSTRACT

PART I: DERIVATIVES FOR THE ISOLATION, PURIFICATION AND MASS SPECTROMETRIC CHARACTERIZATION OF HOMO AND HETERO OLIGOSACCHARIDES

By

Yuanda Zhang

A simple, sensitive method for the structural characterization of oligosaccharides by fast atom bombardment mass spectrometry (FAB-MS) has been designed. Oligosaccharides are labeled with a UV chromophore (which also serves as a charge-stabilizing group) and with a hydrophobic alkyl tail. The chromophore, 2,4-dinitrophenyl or dansyl group, aids UV detection during high performance liquid chromatography (HPLC) and stabilizes ion species formed during analysis by FAB-MS. The hydrophobic tail, provided by an octyl group, enhances the surface activity of the analytes and makes them amenable to separation by reverse-phase chromatography using a C18-bonded phase. This method was applied to the structural analysis of homooligosaccharides, namely a mixture of starch maltodextrins with a degree of polymerization 1-16, potato starch, pure maltohexaose, isomaltohexaose, and N,N',N"-triacetylchitotriose. The method was also applied to the structure of heterooligosaccharides, namely lacto-N-fucopentaose, lacto-N-difucohexaose (these two oligosaccharides were from human milk), glycoprotein fetuin, and a previously

characterized oligosaccharide from a Rhizobial capsular polysaccharide (1). The method gave a good yield of ions for the derivatized compounds, which in the best cases, were detectable at a level of about 1 picomole. In the case of maltohexaose, four series of sequence anions corresponding to sequential loss of glycosyl residues from the reducing and non-reducing ends by different mechanisms were observed. The mixture of derivatized malto-oligosaccharides was easily separated by HPLC. Based on the relative proportions of the individual oligomers in the mixture calculated from HPLC analysis, even though the higher oligomers were present in amounts of about 0.1%, they were still easily detected in mass spectra of the entire mixture. This represents an improvement in sensitivity of at least 100-fold over that reported using the aminobenzoic acid alkyl ester method. In the analyses reported here, single scans were taken and no attempts were made to improve signal to noise by signal averaging.

I would like to dedicate this thesis to my family,
for their support and patient understanding, for their constant encouragement and love.

ACKNOWLEDGMENTS

I would like to express my most sincere appreciation to my research advisor Dr. Rawle I. Hollingsworth for his support, encouragement and guidance in conducting my graduate research, for sharing his knowledge and experience with me. I would also like to thank Dr. C. C. Sweeley and Dr. J. Preiss for their interest in my research, collaborative work and help. I am also grateful to Dr. David G. McConnell for his useful advice, especially for suggesting that Dr. Rawle I. Hollingsworth be my Ph.D. research advisor, an important step in my scientific career. I am very grateful for the support and encouragement received from Dr. C. K. Chang, Dr. G. Babcock, and Dr. D. Nocera.

I wish to thank the members of Dr. Rawle I. Hollingsworth's group: Mr. Kim K. I, Jung S. H., Debbie L. E., Maria B. B., Chuck Cory Campbell, Tim J. N., Baojen S., Lars P., Robert A. C., Luc B., and Ben H., for many informative discussions. The group meetings, group picnic, group tennis and group pizza lunches will make 116 biochemistry building unforgettable.

I also wish to thank the members of the Mass Spectrometry Facility: Dr. J. T. Watson, Dr. D. A. Gage, Bev, Mel, Melinda and Mike, for their hard work with instrumental maintenance and for the free facility lessons, which made the JEOL HX110 more enjoyable.

Finally I wish to thank the Departments of Chemistry and Biochemistry at Michigan State University for using their facilities. I would like to express my gratitude to the US Department of Energy for their support through a research assistantship from Dr. Rawle I. Hollingsworth's research grants.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABEE	*p*-aminobenzoic ethyl ester
CI	chemical impact
Dansyl	5-N,N-dimethylnaphthalene sulfonyl
DASO	N,N-dansyloctylamine
Dimsyl	dimethylsulfinyl
DME	dimethoxyethane
DMSO	dimethylsulfoxide
DNPO	N,N-(2,4-dinitrophenyl)octylamine
EDC	1-ethyl-3(3-dimethylamino-propyl)carbodiimide
EI	electron impact
FAB-MS	fast atom bombardment mass spectrometry
GC-MS	gas chromatography-mass spectrometry
HMPA	hexamethylphosphoramide
HPLC	high-performance liquid chromatography
KDO	3-deoxy-D-manno-oct-2-ulosonic acid
LC-MS	high pressure liquid chromatography interfaced with mass spectrometry
LPS's	lipopolysacchrides
MDO	membrane-derived oligosaccharides
NMR	nuclear magnetic resonance
n.O.e.	nuclear Overhauser effect
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TFA	trifluoroacetic acid
TIC	total ion chromatogram

TLC	thin layer chromatography
UV	ultraviolet

PART I: DERIVATIVES FOR THE ISOLATION, PURIFICATION AND MASS SPECTROMETRIC CHARACTERIZATION OF HOMO AND HETERO OLIGOSACCHARIDES

CHAPTER 1

BACKGROUND RESEARCH

1.1 Introduction

The carbohydrate-rich glycocomplexes on biosurfaces are among the most important current topics of research in biological function and structure. To understand the normal biological and physiological process generally involves the complete characterization of polysaccharides, glycolipids, and glycoproteins, and requires detailed information with regard to the carbohydrate moieties such as molecular weight, sequence, and linkage. The structural elucidation of polysaccharides and larger oligosaccharides generally involves specific hydrolysis to small oligosaccharides followed by the separation and identification of the fragments. Normally, the isolated carbohydrate is available only in very small amounts and has multiple functional groups. This makes the structural analysis of carbohydrates one of the most difficult areas in analytical chemistry.

1.2 The Classical Methods of Sequencing Carbohydrates

Valent *et al.* reported a "general and sensitive chemical method for sequencing the glycosyl residues of complex carbohydrates" in 1980 (2). The total procedure is described in Figure 1. Using this method, successful structural characterization of a complex carbohydrate requires a very complicated procedure. The carbohydrate to be characterized must be relatively pure. The purification of large oligosaccharides is achieved by gel

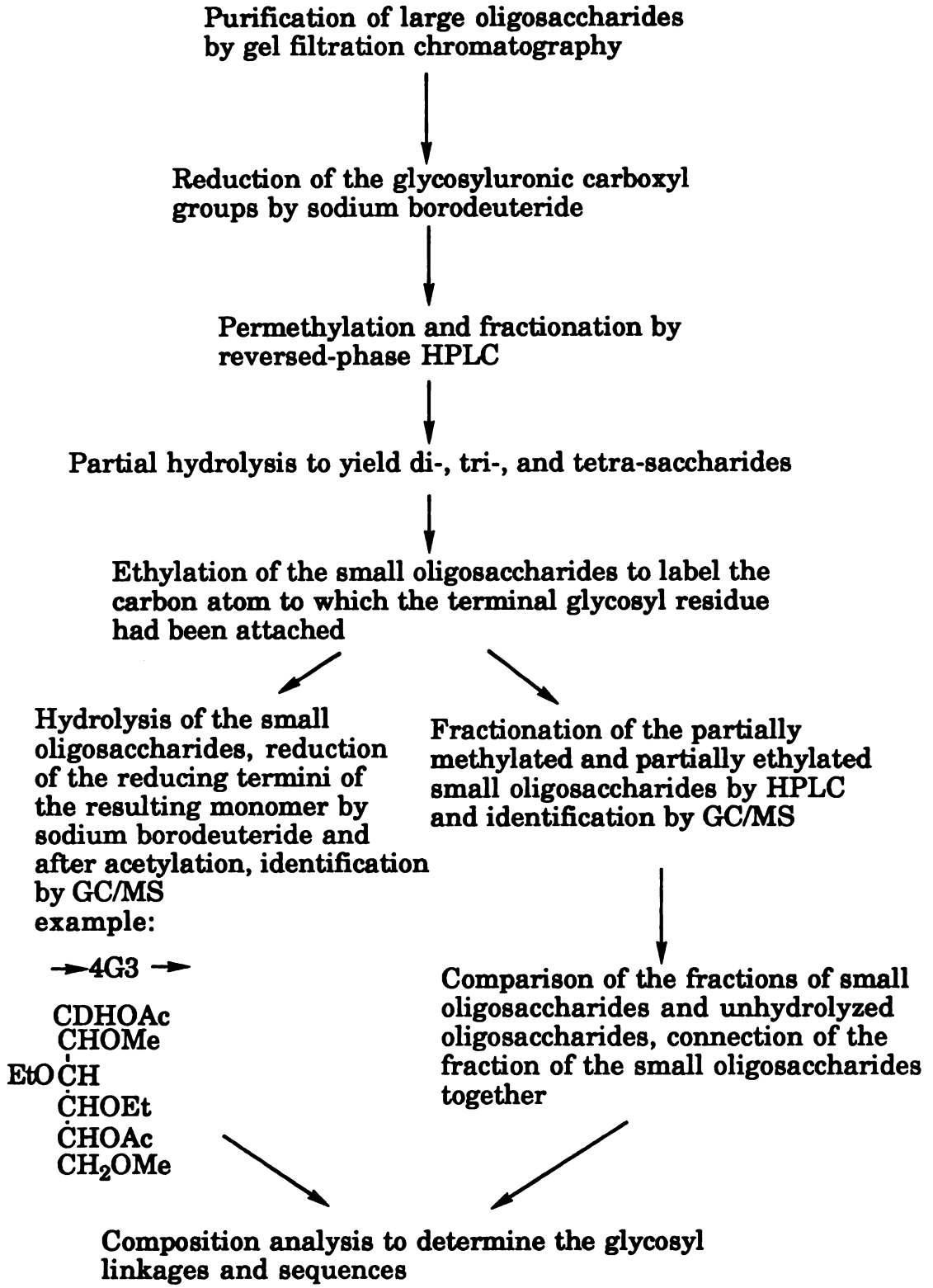


Figure 1: The classical method of sequencing oligosaccharides.

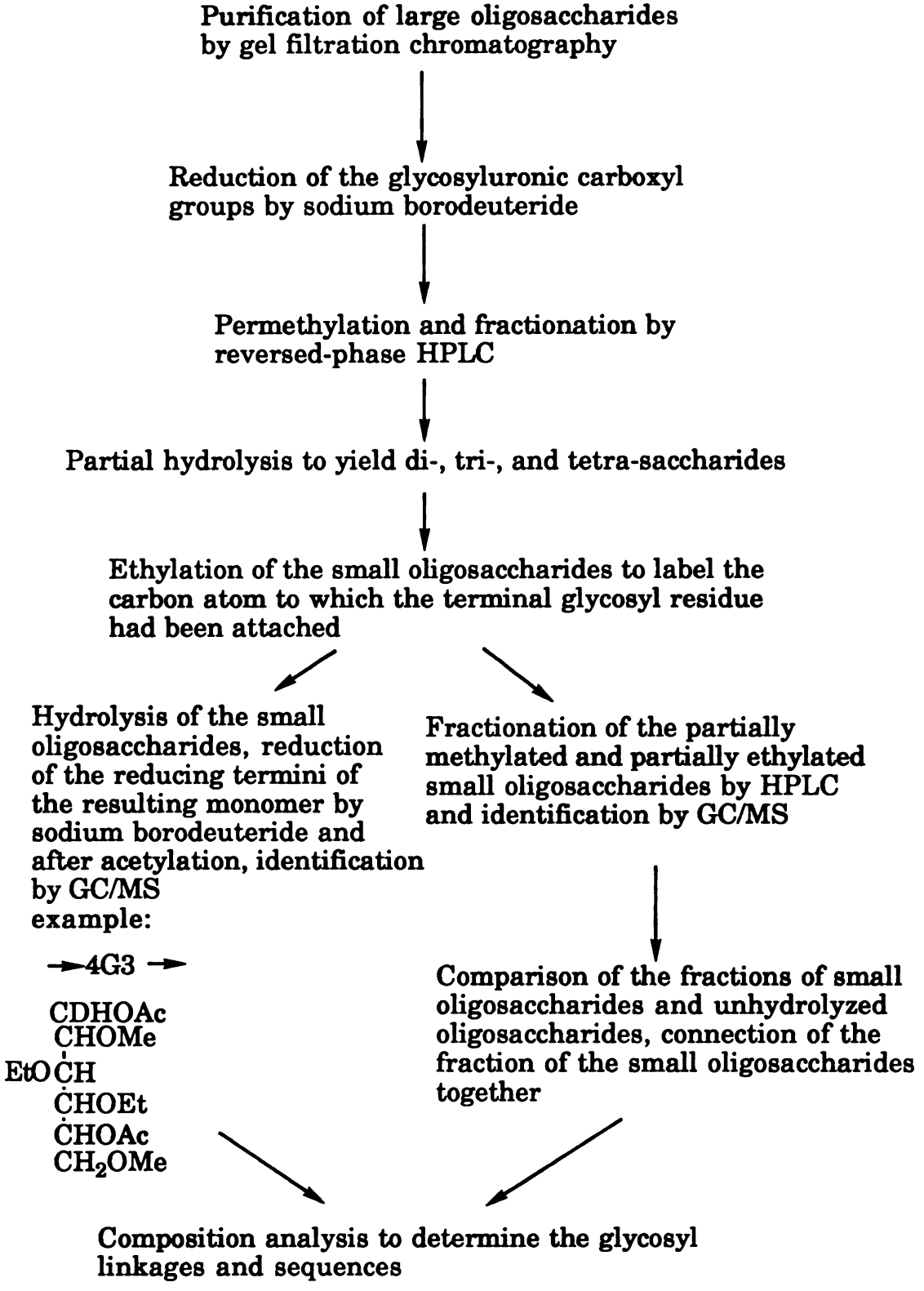


Figure 1: The classical method of sequencing oligosaccharides.

filtration chromatography, which generally uses Biogel P-2 or an ion-exchange resin. The reduction is performed with sodium borodeuteride to identify carboxyl group of glucosyluroic residues present in the complex carbohydrate. The carboxyl group is reduced to the corresponding dideuterio-labeled primary alcohol, which then can be structurally characterized by mass spectrometry. After permethylation, the latter oligosaccharides are fractionated by HPLC. The larger oligosaccharides are partially hydrolyzed. The resulting disaccharides, trisaccharides and tetrasaccharides are ethylated in order to label the carbon atom to which the terminal glycosyl residue had been attached. The partially methylated and partially ethylated small oligosacchrides are then fractionated by chromatography on a reversed-phase HPLC column and identified by GC/MS. According to the overlap of the small oligosaccharides with the unhydrolyzed complex carbohydrate, the oligosaccharide sequences may be pieced together. The fractions from the reversed phase HPLC column are hydrolyzed to monomers, sequentially reduced by sodium borodeuteride, and acetylated to the corresponding partially methylated, partially ethylated alditol acetate. After acetylation, the deuterated alditol acetates are uniquely labeled as the hydroxyl groups at C-1 and C-5 (or C-4 for furanosyl residues) and the positions of linkage are acetylated. The resulting monomers are again identified by GC/MS. The glycosyl-linkage composition of each oligomer is then used to determine the glycosyl sequence.

The typical limitations of this classical method are that it is complicated, time consuming and often requires in excess of a year to complete the whole procedure. Another limitation is the quantity of material required. Valent *et al.* said that this method requires only

milligrams of material but the quantities of many isolated biological samples are far below milligram amounts and therefore, cannot be analyzed by this method.

1.3 Electron Ionization Mass Spectrometry

The mechanism of electron ionization (EI) mass spectrometry is electron loss from the parent ion thus putting a charge on the molecule so its path can be directed. Sufficient energy is also put into the molecule to break some of its chemical bonds. When the analyte collides with the fast electron, the analyte loses an electron and becomes a particle with an odd number of electrons and a positive charge:

$$M + e^- = M^{+\cdot} + 2\ e^-$$

Most EI mass spectra are recorded at an electron energy of 70 eV. Usually only a small fraction (less than 0.1%) of the analyte molecules are ionized in EI (3), and the rest of the sample is pumped away into the vacuum system. Because the pressure is so low in the ion source, the ionized molecules generally do not collide with the walls and other molecules before their fragmentation. The EI mass spectrum usually includes signals from the molecular ion and fragments.

The limitations of EI mass spectrometry is that generally EI requires that the sample be in the gas phase. Thus non-volatile, high-molecular-weight carbohydrate materials have remained beyond the scope of the EI measurement. Because high-molecular-weight carbohydrates are not thermally stable and have to be degraded to the simplest species and chemically derivatized to provide thermally stable condition. The

peralkylation methods add mass to the analyte, increasing the complexity of the analysis. The additional energy required for volatilization of the molecule causes unavoidable thermolytic processes, producing high-abundance low-mass fragments, and decreasing the intensity of the molecular ion. Sometime too many fragmentations make the EI spectra difficult to interpret. However, if the energy of electron beam is lowered to decrease the extra fragments, the entire sensitivity of the spectrum is simultaneously decreased. To obtain complete structural information for a high-molecular-weight carbohydrate by EI generally requireds 100-200 μg sample (4). These large sample requirements are impossible to meet in the case of scarce, isolated biological materials.

1.4 Chemical Impact Mass Spectrometry

EI mass spectrometry works by colliding molecules with 70 eV electrons but also results in a lot of fragmentation in the process. Because of this fact, there often are not enough surviving molecular ions to indicate the molecular weight of the compound. This is one of the most important pieces of information. The chemical impact (CI) method is a less energetic, softer ionization mode. In the CI process, ionization of the analyte is a result of a gas-phase chemical reaction, not a collision by energetic electrons. Chemical impact uses methane or ammonia as reagent gas. At low pressure in a mass-spectrometer ion source, the reagent gas is bombarded with energetic electrons to produce the expected primary ions.

$$CH_4 + e^- = CH_4^{+\cdot} + 2\,e^-$$

$$CH_4^{+\cdot} = CH_3^+ + H\cdot$$

peralkylation methods add mass to the analyte, increasing the complexity of the analysis. The additional energy required for volatilization of the molecule causes unavoidable thermolytic processes, producing high-abundance low-mass fragments, and decreasing the intensity of the molecular ion. Sometime too many fragmentations make the EI spectra difficult to interpret. However, if the energy of electron beam is lowered to decrease the extra fragments, the entire sensitivity of the spectrum is simultaneously decreased. To obtain complete structural information for a high-molecular-weight carbohydrate by EI generally requireds 100-200 μg sample (4). These large sample requirements are impossible to meet in the case of scarce, isolated biological materials.

1.4 Chemical Impact Mass Spectrometry

EI mass spectrometry works by colliding molecules with 70 eV electrons but also results in a lot of fragmentation in the process. Because of this fact, there often are not enough surviving molecular ions to indicate the molecular weight of the compound. This is one of the most important pieces of information. The chemical impact (CI) method is a less energetic, softer ionization mode. In the CI process, ionization of the analyte is a result of a gas-phase chemical reaction, not a collision by energetic electrons. Chemical impact uses methane or ammonia as reagent gas. At low pressure in a mass-spectrometer ion source, the reagent gas is bombarded with energetic electrons to produce the expected primary ions.

$$CH_4 + e^- = CH_4^{+\cdot} + 2\,e^-$$

$$CH_4^{+\cdot} = CH_3^+ + H\cdot$$

These primary ions will collide with neutral reagent gas, and produce a series of secondary ions:

$$CH_4^{+\cdot} + CH_4 = CH_5^+ + CH_3\cdot$$
$$CH_3^+ + CH_4 = C_2H_5^+ + H_2$$

If a small amount of analyte gas is present, the collisions of the secondary ions and the analyte molecules produce gas-phase acid-base chemical reactions. CH_5^+ is a very strong gas-phase acid, and can readily give up a proton. Most organic molecules are proton acceptors:

$$CH_5^+ + M = CH_4 + (M + H)^+$$

In the CI process, the collision energy of the secondary ions and the analyte is lower than the 70 eV in the EI process, so in CI spectra, a large proportion of the molecular ions remain and fewer fragment ions are formed than in EI spectra.

Although CI is a soft ionization technique, it still requires that the analyte be in the gas phase. There are a lot of reports of CI studies for peralkylated carbohydrates, but there were still restrictions for the high-molecular-weight non-volatile carbohydrate to be analyzed.

1.5 Fast Atom Bombardment Mass Spectrometry (FAB-MS)

The need for effective determination of the structure of large and complex biomolecules has led to the development of several other soft ionization methods, such as secondary ion, fast atom bombardment, electrospray, plasma desorption and laser desorption mass spectrometry.

The fast atom bombardment mass spectrometry method was introduced in the last decade (5).

1.5.1 Operating Process

In the FAB-MS process, non-volatile and thermally labile compounds can be analyzed. The samples are first dissolved in a solvent and then mixed with a polar, viscous, low-vapor-pressure liquid matrix, which is generally glycerol, or thioglycerol often doped with trifluoroacetic acid (for positive FAB) or triethanolamine (for negative FAB). In the FAB-MS, a neutral accelerated beam of inert xenon, argon, or other suitable gas atoms, with 6 to 10 keV of translational energy, is fired from an atom gun. The stream of neutral gas molecules is generated by ionizing, accelerating and then charge-neutralizing noble gas atoms. The noble gas atoms undergo a collision with a high-energy ion beam to form fast ions which are accelerated. The fast ions capture electrons from an electron-rich saddle-field to become fast atoms and move towards a small metal target which has previously been loaded with the liquid matrix containing the sample to be analyzed. When the atom beam collides with the sample and matrix, kinetic energy is transferred to the surface molecules, causing many molecules to be sputtered out of the liquid into the high vacuum region of the ion source. A large number of the molecules are ionized during the sputtering process. The ionization is achieved by interaction with matrix: addition of a proton or cation (positive ion mode) or loss of a proton (negative ion mode) giving pseudo-molecular ions as major signals in the spectrum. During ionization, some internal energy is imparted to the molecule which leads to fragmentation.

1.5.2 The Sample Preparation

Generally, the sample is dissolved in a volatile solvent and 1-1.5 μl of solution is mixed into 1-2 μl of matrix which was previously placed on the target of the probe tip. The probe carrying the loaded target is put into the ion source through a vacuum lock system. The solvent evaporates from the probe in the ion source chamber prior to exposure to the source.

1.5.3 Matrix

The viscous liquid matrix plays an important role in the diffusion of sample from the surface of the matrix. It also promotes a stable, reproducible ion current lasting for longer periods of time for the diffusion process (6). A closer view of this process is shown in Figure 1.2.

The figure describes the surface of the matrix during bombardment by the neutral fast atom beam and how the desorbed species are produced. When the bombarding beam approaches a 90° angle of incidence, the majority of the energy deposition is limited to the surface layer of the solution, causing intramolecular and intermolecular vibrations deeper in the matrix. The viscous liquid matrix replaces the destroyed area and provides additional analyte to the matrix surface. With this process, using a glycerol matrix, the positive or negative ion signals can be constantly supplied for periods of 20 to 30 min to the detector.

Because of the requirement for high viscosity, most of the matrices are naturally hydrophilic. This causes several notable problems. One is poor sensitivity, causing high detection limits. Other problems are intense background ions, intense cluster ions from the matrix, and ion suppression effects. These problems inhibit the formation of certain ions from the sample.

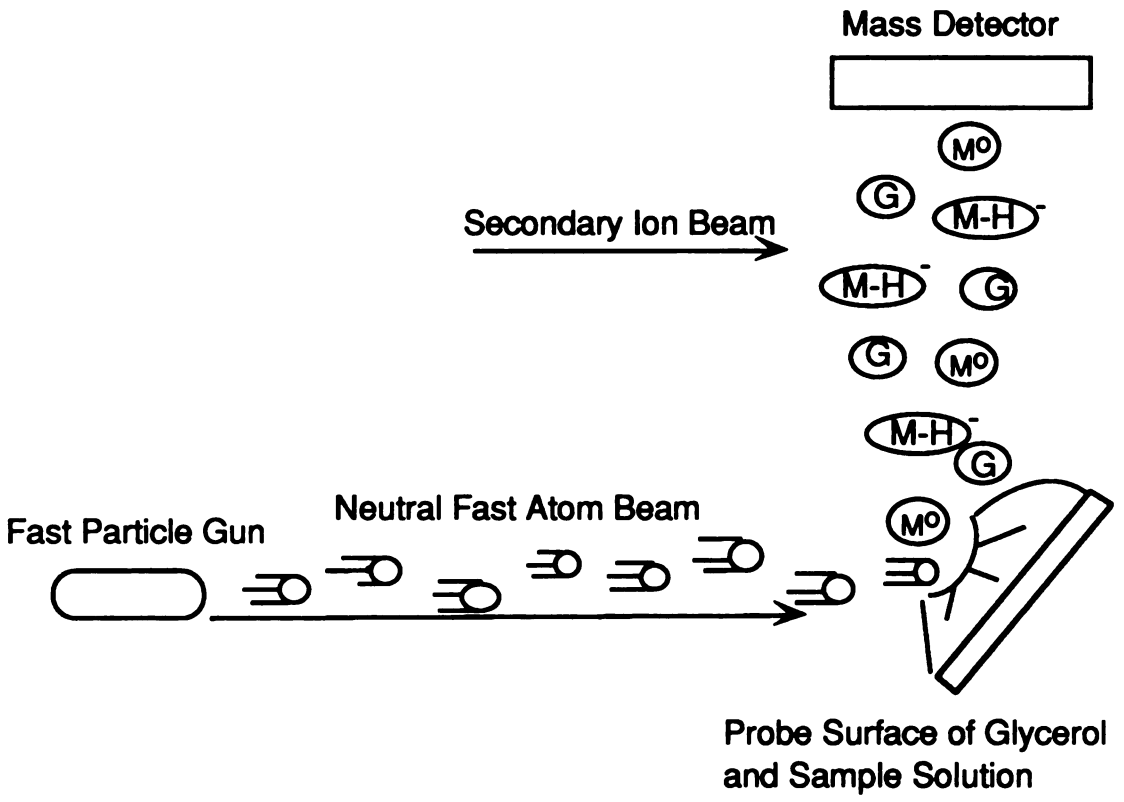


Figure 1.2 View of the matrix surface during negative FAB-MS process. $(M-H)^-$ is a negative charged ion, (G) is a matrix molecular such as glycerol, and $(M)^o$ is a neutral molecule.

The interaction between analyte and matrix depends on hydrophobicity, and hydrophilicity. To improve the quality and intensity of a spectrum, different matrices or matrix additives can be used. For example, in this research, a mixture of glycerol and thioglycerol (1:1) was employed. If glycerol alone is used, it gives strong cluster ions at multiples of 92 + H^+ (in positive mode spectra), starting at 93 amu and continuing into the region of m/z 1000. Addition of thioglycerol suppresses cluster ions up to m/z 500. Addition of 0.1 M trifluoroacetic acid (TFA) solution directly to the matrix helps to displace the cations and is employed in positive mode

spectra. The addition of triethanolamine to the matrix helps to displace anions. Sensitivity can also be improved by modifying the analyte. Hydrophobic groups have less of an attraction for polar matrix molecules. When analyte is modified with a hydrophobic group, which reduce hydrogen-bonding interactions with matrix, the analyte tends to occupy the surface layer of the polar sample solution. The FAB-MS signal recorded is largely determined by the interactions between hydrophilicity and hydrophobicity of the matrices and analytes.

Some requirements for the liquid matrix are (7, 8):

1. The solubility of sample in the matrix must be very good.
2. The vapor pressure should be as low as possible.
3. The matrix must be chemically inert, and not reactive with the sample.
4. To ensure the diffusion process, the viscosity of the matrix must be reasonably low.

1.6 Methods of Carbohydrate Analysis

Methods for determination of the structures of carbohydrates are quite laborious, insensitive, and time-consuming compared to established sequencing methods for the two other major biopolymers: proteins and nucleic acids.

Gas chromatography has been used to separate small oligosaccharides and derivatized oligosaccharide fragments. The problem is that the large oligosaccharides never elute, or at best, do so at a very slow rate because of their high polarity and high molecular weight. In addition, the analyzed material can not be recovered.

Gel filtration can be used to isolate both large and small oligosaccharides, and the material can be recovered quantitatively. However, several days are needed to complete the analysis, and detailed structural information still can not be obtained.

Although information can be obtained by ^{1}H NMR spectroscopy in some instances using pulse schemes which reveal connectivities involving the anomeric proton or carbon and nuclei across the interglycosidic linkage, it requires a major commitment of personnel and facility resources as well as large amounts of pure materials. These methods also require a fairly high degree of sophistication in experimental set-up and interpretation and are far from routine.

The high-performance liquid chromatography (HPLC) technique has the capability of separating, non-destructively, oligosaccharides with high yields. The lack of a fluorophore makes carbohydrate very poorly detectable by UV, which restricts using a conventional UV detector, is a limiting feature for analysis of carbohydrates by HPLC.

Mass spectrometry has contributed significantly to carbohydrate structure elucidation and holds much promise for sequencing carbohydrates because the quantity of material required is typically orders of magnitude lower than that required for NMR spectroscopy. In principle, the technique can be applied directly to mixtures if the instrumentation exists for precursor and product ion analysis as is the case with triple quadrupole and four-sector instruments. One of the most popular of current ionization/desorption methods for analyzing non-volatile analytes by mass spectrometry is fast atom bombardment (FAB-MS) (9). FAB-MS data can be obtained from native carbohydrates as long as they are free of salts and other contaminants (10). However, it is very difficult to obtain

analytes of such purity. Trace amounts of salts will bear the greater proportion of the total ion current, and surface active components will occupy the matrix surface, displacing the sample and preventing desorption. Currently, several nanomoles of underivatized carbohydrates are required for analysis by FAB-MS (11). The negative ion FAB-MS of γ -cyclodextrin was reported in 1981 (12). The spectrum contained a peak for a molecular ion (M-H = 1295), and one for another ion due to loss of one sugar unit. The preliminary data indicate that FAB-MS will be useful in obtaining the molecular weight of complex oligosaccharides.

1.7 Classical Methods for Carbohydrate Derivatives

Strategies for solving the problems in carbohydrate analysis usually involve the classical derivatization methods, such as permethylation and peracetylation. Systematic and reliable procedures have been developed in the early years for these transformations which also allow sample vaporization without thermal decomposition. The permethylated polysaccharides are first degraded by acid-catalyzed hydrolysis (13, 14). The linkage information for the resulting monosaccharides is based on the mass-spectral identification of the reduced acetylated compounds (15, 16). These procedures have been widely used, but mostly for obtaining EI and CI spectra usually by GC-MS (17). The disadvantage of this method is that permethylation and peracetylation are sometimes incomplete for large oligosaccharides. Another drawback is that the molecular weight increases significantly after derivatization, resulting in mass window limitations.

Often methods such as the radio labeling of carbohydrates at their reducing termini by reduction of the oligosaccharides with NaB^3H_4 (18)

have also been used. The problems have been also addressed by the use of hydrophobic chromophores and charge localization groups attached to the carbohydrate molecule. This method leads to increased UV sensitivity and better fragmentation in mass spectrometric analyses, as recently reviewed (19). The general strategy for making such derivatives involves the covalent attachment of the chromophore group to the reducing end of oligosacchride by reductive amination.

1.8 Derivatization by Chromophore Groups

The sensitivity of FAB analysis is strongly influenced by the solubility of the analyte in the matrix. In general for carbohydrates, as the hydrophilicity of the aglycone decreases, solubility and sensitivity also decrease. The poor surface activity of carbohydrates in polar matrices is another one of the features that limits the utility of ion sputtering/ionization techniques such as FAB-MS in the study of carbohydrates.

Using reductive amination, the preparation of several derivatives of reducing oligosaccharides has been reported. In 1978, Matsushim et al. (20) reductively aminated oligosaccharides with 2-aminopyridine by means of reduction with sodium cyanoborohydride giving a fluorescent 2-aminopyridine derivative. They used this derivative to determine degree of polymerization by chromatographic properties, sequence of the sugar units, and the linkage points of the sugar units. For the linkage information, they reported the permethylated derivative. The molecule they characterized was a disaccharide. In 1983, Prahash et al. (21) coupled 7-amino-4-methylcoumarin to the aldehyde group of oligosaccharides. One picomole of the derivative can be detected after thin-layer chromatography by

illumination with a UV light. The mass spectrometry technique was not utilized.

One of the most popular and successful groups for sequencing carbohydrates is the aminobenzoic acid alkyl ester group (22-24). This method has led to enhancement of the ion intensity observed for carbohydrate analytes during FAB-MS, and has allowed the sequence for the carbohydrate chains of several glycoproteins to be determined. In 1984, Sweeley et al. (22) described a method for structural characterization of oligosaccharides by preparing the UV-absorbing oligoglycoaminoalditols from reductive amination with *p*-aminobenzoic acid ethyl ester to form the so-called ABEE derivatives. They used this method for HPLC and FAB-MS analysis. The mass range they studied had an upper limit of 900 amu which corresponded to derivatized maltotetraose. In 1988, Burlingame et al. (23) applied the ABEE method to the study of the structure of oligosaccharides released from glycoproteins. They reported that the ABEE derivatives provided a deprotonation center enhancing the mass spectra. Positive and negative ion spectra contained molecular species of similar abundance, but the fragment ion peaks giving the sequence information were significantly more prominent in the negative ion mass spectra. They did not observed cleavages suggesting or confirming branched oligosaccharides in the pattern of fragmentation. Later, Burlingame et al. (24-27), using the heptasaccharide maltoheptaose as a model oligosaccharide, recognized that lengthening the alkyl chain from methyl to n-tetradecyl resulted in a concomitant increase in the molecular ion abundance. They reported that very high mass spectral sensitivities were achieved with *n*-tetradecyl and *n*-decyl *p*-aminobenzoates. However, the yields of derivative obtained were significantly lower than those obtained for

n-octyl, n-hexyl, n-butyl, ethyl, and methyl *p*-aminobenzoates. n-Octyl and n-hexyl *p*-aminobenzoates were found to be optimal considering both yields of derivatives and mass spectral sensitivity. Using maltoheptaose as a test compound, the ABEE derivative analyte could be detected at levels on the order of 10 pmole with a signal to noise ratio of 8:1. Higher sensitivity could be realized with longer alkyl chain derivatives, but the use of such derivatives was compromised by solubility problems when very long chains were employed. The use of an octyl chain (ABOE derivative) gave significantly higher sensitivity (about 20:1) in comparison to underivatized carbohydrate without severe solubility problems. Limitations of the aminobenzoic acid alkyl ester method are that the derivatizing reagent must be prepared beforehand, the actual derivatization yield is not quantitative for longer chain alkyl esters, and significant sample workup must be performed prior to analysis.

Another modification to the ABEE method was reported by Reinhold *et al.* (28). They attached the aminobenzoic acid pentafluorobenzyl ester group to the reducing end of glycan residues. This imparted fluorescent and UV properties for chromatographic detection, and also allowed the molecule to behave as an electron trap under negative ion chemical ionization. They employed this method using a liquid chromatography interface with mass spectrometry (LC-MS). The advantage of using this derivative is that the sensitivity is very high. Detection levels of 800 fmol were realized with signal-to-noise of 100:1. Even with a 16 fmol injection, a signal-to-noise ration of 3:1 was still observed. The disadvantage of this technique is that only a strong molecular signal can be observed; no sequence or other structural information could be obtained.

Recently a new method was developed by Dr. Rawle I. Hollingsworth for the sequencing of peptides by FAB-MS. In this method, the sensitivity of detection and amount of sequence-specific fragmentation of these molecules was increased by attaching a positively charged 2-(triphenylphosphonio) ethyl group to the N- terminus of the molecule (29). The strategy we adopt here is one that modifies the carbohydrate with a N,N-(2,4-dinitro)phenyloctylamine (DNPO) group for negative mode and N,N-Dansyloctylamine (DASO) derivatives for positive mode detection. These groups serve as the site for charge localization. The dinitrophenyl group was used since it can readily form a negatively charged species. The negative charge is readily delocalized by the two nitro groups on the phenyl ring. In a previous study, it was determined that dinitrophenyl derivatives of secondary amines have molar absorption maxima which were very insensitive to the nature of the alkyl group, the only difference being in the rate of derivatization (30). This makes this derivative ideal for the direct quantitation of molar proportions of components in a mixture by direct integration of peak areas. Also, the positive charge on the DASO derivative (formed by protonation of the demethylamino group) is readily delocalized and stabilized. Both the dinitrophenyl and dansyl groups are also excellent chromophores for UV detection and can be attached in a facile manner. The octyl group increases the surface activity of the carbohydrate analytes which have poor surface activity in the hydrophilic matrix. The resulting DNPO and DASO derivatives are therefore designed to overcome problems of low UV visibility, low ionization efficiency, and low surface activity typically observed with carbohydrate compounds during analysis by HPLC and FAB-MS. A large amount of fragmentation with charge retention on the reducing end is also observed.

1.9 Nomenclature

The fragmentation patterns of the carbohydrate investigated by FAB-MS are generally more complicated than those of other bimolecules, such as peptides and nucleic acids. A systematic nomenclature for assigning and describing the different types of fragmentations of the carbohydrate moieties has been introduced by Domon and Costello (31).

1.9.1 Nomenclature of Carbohydrate Fragments

There are two major modes of cleavage normally observed in FAB-MS of carbohydrates. The simplest one is the fragmentation which results from the cleavage of the glycosidic bond giving information about the oligosaccharide sequence. The other mode of fragmentation is more complex and involves the sugar ring cleavage. Sequences from the latter mode are more difficult to assign, but they contain very important structural information.

In Figure 1.3, when the charge is retained on the non-reducing end, A_i, B_i, and C_i labels are used to designate the fragments, where the i indicates the number of the glycosidic bond cleaved counting from the non-reducing end. In the case of fragmentation containing the aglycone or reducing unit, X_j, Y_j, and Z_j labels are used to designate the fragments. Here, the subscript j indicates the number of the glycosidic bond cleaved counting from the reducing end or the aglycone unit. In Figure 1.3, only A and X sequences indicate the ring cleavage. The glycosidic bond linked to the aglycone is numbered 0. The fragments formed from the reducing end

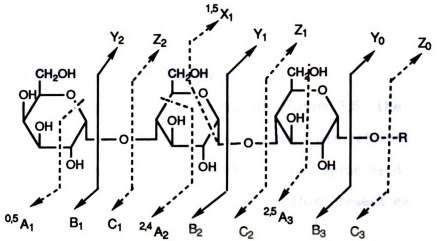


Figure 1.3. Nomenclature of carbohydrate fragments.

with retention of the glycosidic oxygen is the Y sequence. The fragments formed from the non-reducing end with retention of the glycosidic oxygen is the C sequence. The fragmentation without retention of the glycosidic oxygen on the species formed, from the reducing end is the Z sequence. The fragmentation without retention of the glycosidic oxygen, from the non-reducing end, is the B sequence.

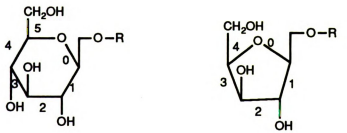


Figure 1.4 The numbering of the ring bonds.

The superscripts represent the cleavage involved in the carbohydrate rings. Two superscripts k and l (*i.e.* $^{k,l}A_i$, $^{k,l}X_j$) are used to represent the

sugar ring bonds, the ring bonds are numbered according to Figure 1.4.

1.9.2. Branched Oligosaccharide Moieties

In the case of the branched oligosaccharide (Figure 1.5), the main body is the longest moiety, the "core", and the branches are referred to as the "antennae." The labeling of the core portion is the same as described in section 1.9.1. The Greek letters α, β, γ are used to represent each branch which are designated in the order of decreasing molecular weight. The fragmentation of the core portion is described without Greek letters. Their

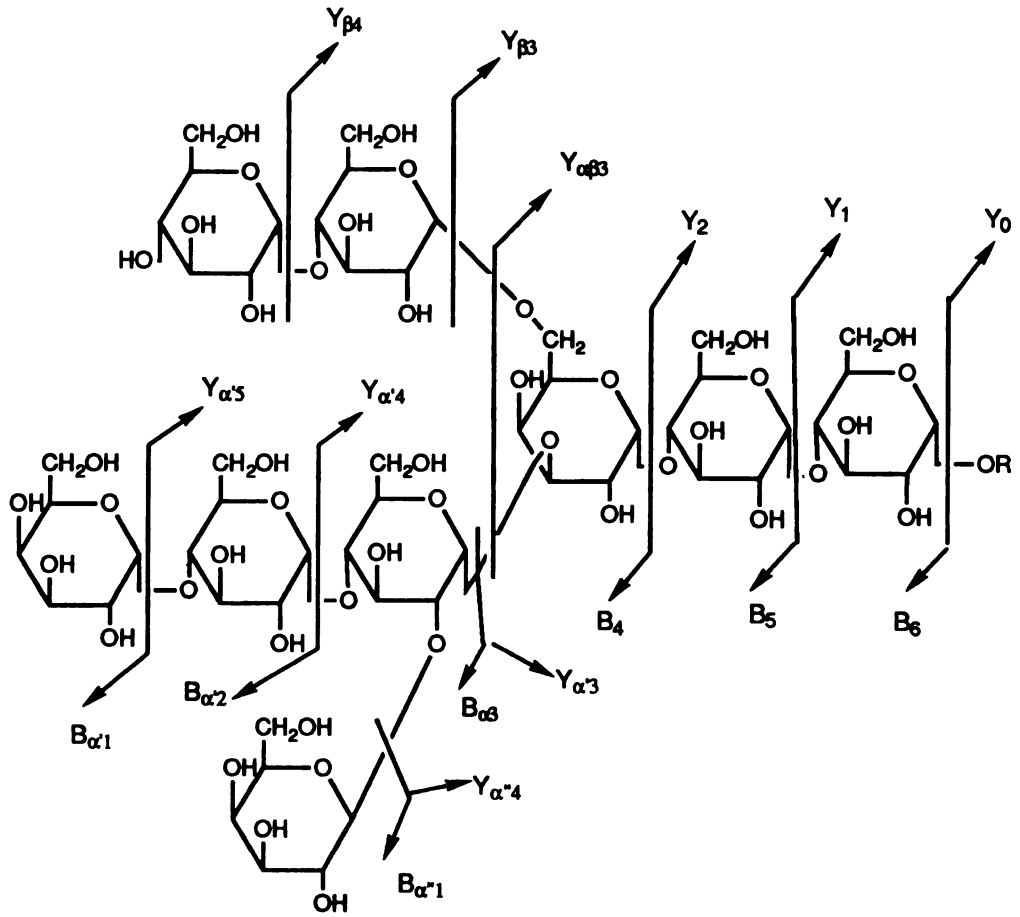


Figure 1.5 The nomenclature for branched oligosaccharide.

use begins from the point of connection of the antennae to the core. In the case where there are two branches cleaved on a same ring, the Greek letters representing each cleaved branch can be used together, such as αβ. The Greek letter α' and α" are used, the α', α" are designated in the order of decreasing molecular weight, in the case of sub-branch observed.

1. 10 Sequence from Oligosaccharide

The fragment ions generated from the ion source reflect the structure of the oligosaccharide. The general sequence information was summarized by Anne Dell (32) in 4 type of cleavage: those are depicted in Figure 1.6. In series A, the charge is retained on the non-reducing end and occurs in both positive and negative modes. In series B, the charge is retained on the reducing end, and the fragmentation yields an unsaturated hexose. This kind of fragmentation occurs in the positive mode only. In series C, the charge is retained on the reducing end, and the ions are 28 amu heavier than those formed in pathway A. This sequence can be observed in both positive and negative modes. Series D gives fragmentation with charge retained on the non-reducing end. Fragments are 42 amu heavier than those from pathway A. This sequence is most commonly observed in negative mode spectra and is rarely observed in the positive mode (32).

Understanding the mechanism of fragment formation is very important in identification of the sequence from the derivatized oligosaccharides. In cases using different derivatives, the sequence from internal glucosidic bond cleavage with charge retained on the non-reducing end always gives the same ions. In sequence with charge retained on the reducing end, ions varies depending on the specific type of derivatives. In

very complicated FAB-MS of oligosaccharides, this character can help to identify the sequence as coming from the reducing end or non-reducing end.

The reactions may proceed by radical mechanisms (33), or by concerted mechanisms (34).

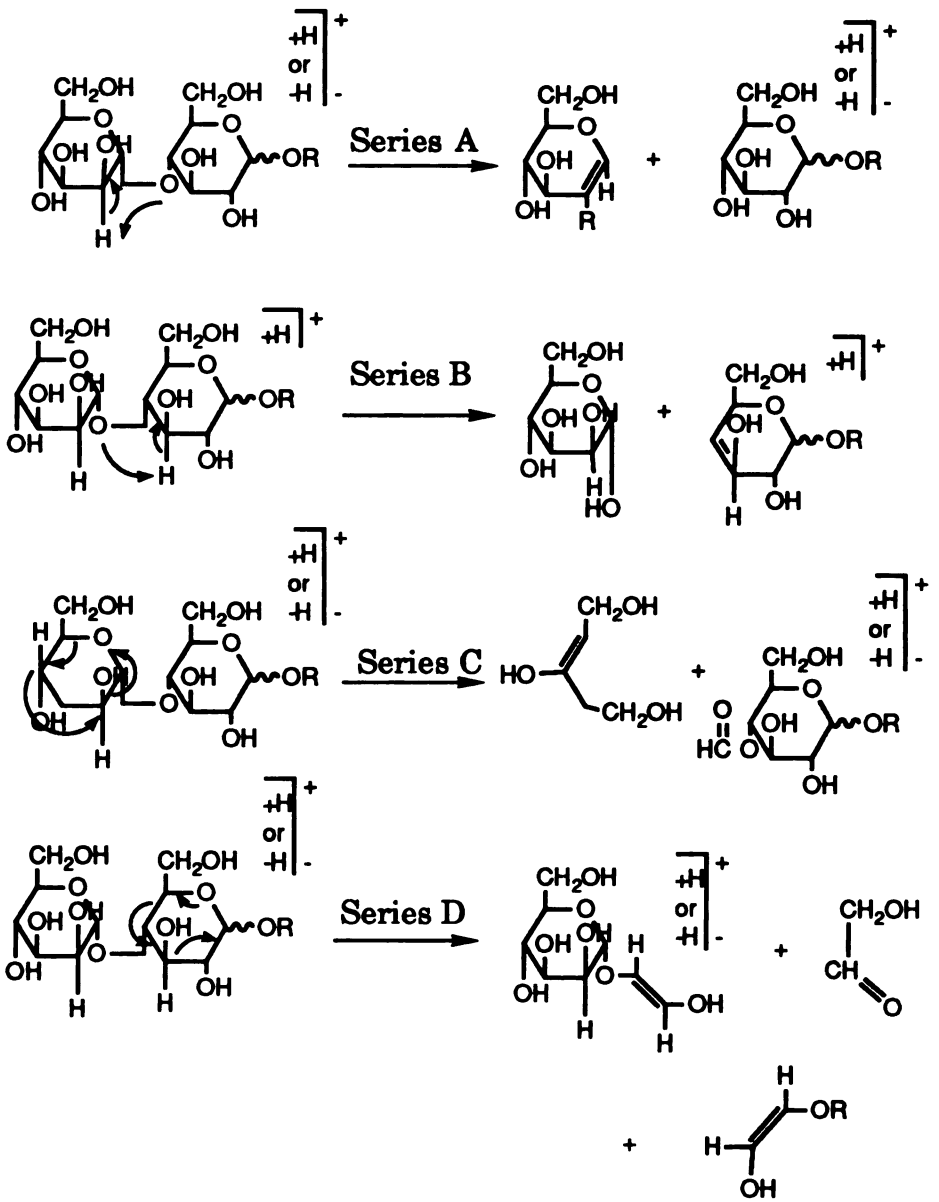


Figure 1.6 Oligosaccharide fragmentation pathways.

1. 11 Reference

1. Hollingsworth, R. I., and Dazzo, F. B.,*Carbohydr. Res.*, (1988) 172 97-112.
2. Valent, B. S., Darvill, A. G., McNeil, M., Robertsen, B. K., and Albershem, P., *Carbohydr. Res.* (1980) 79 165-192.
3. Suelter, C. H., and Watson, J. T., *"Biomedical applications of mass spectrometry"* (1989) 5, An Interscience Publication, John Wiley & Sons, New York.
4. Reinhold, N. V., and Carr, S. A., *Mass Spectrom. Rev.* (1983) 2 153-221.
5. Barber, M., Borodoli, R. S., Sedgwick, R. D., and Tyler, A. N., *J. Chem. Soc. Comm.* (1981) 7 325.
6. Barber, M., Bordoli, R. S., Sedgwick, R. D., and Tyler, A. N., *Nature* (1981) 293 270-275.
7. Gower, J. L. , *Biomed. Mass Spectrom.* (1985) 12 191.
8. DePauw, E., *Mass Spectrom. Rev.* (1986) 5 191.
9. Beth, L. Gillece-Castro and Burlingame, A. L. *"Methods in Enzymology"* (1990) 193, 689-712, (Abeison, J. N., and Simon, M. I., Eds,) Academic Press, San Diego.
10. Egge, H., and Peter-Katalinic, J. *Mass Spectrom. Rev.* (1987) 6, 331-393.
11. Kamerling, J. P., Heerma, W., Vliegenthart, J.F.G., Green, B. N., Lewis, I. A. S., Strecker, and Spik, G. *Biomed. Mass Spectrom.*, (1983) 10, 420-425.
12. Fenselau, C., Cotter, R., Hansen, G., Chen, T., Heller, D., *J. Chromatog.* (1981) 218 21-30.

13. Lindberg, B., Lonngren, J., Svensson, S., *Adv. Carbhydr. Chem. Biochem.* (1975) 31, 185-240.

14. Hakamori, S.-I., *J. Biochem.* (Tokyo) (1964) 55, 205-208.

15. Lindberg, B., Lonngren, J., *"Methods in Enzymology"*, (1978) 50 1 Ginsburg, V., Ed., Academic Press, New York.

16. Biermann, C. J., *"Analysis of Carbohydrates by GLC and MS,"* (1989) 27-41, Biermann, C. J., McGinnis, G. D., Eds; CRC Press, Boca Raton, FL.

17. Carpita, N. C., Shea, E. M., *"Analysis of Carbohydrates by GLC and MS,"* (1989) 157-216, Biermann, C. J., McGinnis, G. D., Eds; CRC Press, Boca Raton, FL.

18. Turco, S. J. *Anal. Biochem.* (1981) 118 278-283.

19. Poulter, L., and Burlingame, A. L. *"Methods in Enzymology"* (1990) 193, 661-689, Abeison, J. N., and Simon, M. I., Eds, Academic Press, San Diego.

20. Hase, S., Ikenaka, T., and Matsushim, Y., *Biochem. Biophys. Res. Comm.* (1978) 85 257-263.

21. Prahash, C., and Vijay, I. K. *Anal. Biochem.* (1983) 128 41-46.

22. Wang, W. T, LeDonne, N. C. JR., Ackerman, B., and Sweeley, C.C. *Anal. Biochem.,* (1984) 141, 366-381.

23. Webb, J. W., Jiang, K., Gillece-Castro, B. L., Tarentino, A. L., Plummer, T. H., Byrd, J. C., Fisher, S. J., and Burlingame, A. L. *Anal. Biochem.,* (1988) 169 337-349.

24. Poulter, L., Kerre, R., and Burlingame, A. L. *Anal. Biochem.* (1991) 195 1-13.

25. Hernandez, L. M., Ballou, L., Alvarado, E., Gillece-Castro, B. L., Burlingame, A. L., and Ballou, C. E. *J. Biol. Chem.* (**1989**) 264 11849-11856.

26. Poulter, L., Earnest, J. P., Stroud, R. M., and Burlingame, A. L. *Proc. Natl. Acad. Sci. USA* (**1989**) 86 6645-6649.

27. Poulter, L. J., Earnest, J. P., Strond, R. M., and Burlingame, A. L. *Biomed. Environ. Mass Spectrom.*, (**1988**) 16, 25-30.

28. Caesar, J. P. Jr., Sheeley D. M., and Reinhold V. N. *Anal. Biochem.* (**1990**) 191 247-252.

29. Wagner, D. S., Salari, A., Gage, D. A., Leykam, J., Fetter, J., Hollingsworth, R. I., and Watson, J. T., *Biol. Mass Spectrom.*, (**1991**) 20 419-425.

30. McIntire, F. C., Clements, L. M. and Sproull, M. *Anal. Chem.* (**1953**) 25 1757-1758.

31. Domon, B., and Costello, C. *J. Glycocojugate* (**1988**) 5 397-409.

32. Dell, A., *Adv. Carbohydr. Chem. Biochem.*, (**1987**) 45, 19.

33. Dell, A., and Oates, J. E. T., *"Analysis of Carbohydrate by GLC and MS"* Biermann, C. J., McGinnis, G. D., Eds, CRC Press, Boca Raton, FL, (**1988**) 217-235.

34. Hendrickson, J. B. *Angew. Chem. Int. Ed. Engl.* (**1974**) 13, 47-76.

CHAPTER 2

2.1. General Methods

2.1.1. Chemicals

Glycerol, maltohexaose, Fetuin, N,N',N"-triacetylchitotriose, and triethanolamine were obtained from Sigma Chemical Co. (St. Louis, MO). Thioglycerol, sodium cyanoborohydride, octylamine, cesium iodide, and all HPLC grade solvents were obtained from Aldrich Chemical Co. (Milwaukee, WI). Ultramark 443 and 6121 were obtained from PCR Inc. (Gainsville, FL). The 2,4-dinitrofluorobenzene was obtained from Eastman Organic Chemicals (Rochester, NY). Glacial acetic acid was obtained from Mallinckrodt Specialty Chemicals Co. (Paris, KY). Methanol was obtained from Baker (Phillipsburg, NJ). Triethylamine was obtained from Fisher Scientific, Inc. (Fair Lawn, NJ). Maltodextrin (degree of polymerization = 1-16) was obtained from Staley Manufacturing, and the *Rhizobium* capsular oligosaccharide was isolated in another study (1). Potato starch was obtained from Dr. Mirta Sivak (MSU, East Lansing, MI).

2.1.2. The General Procedure for Making Derivatives

2.1.2.1. Preparation of N,N- (2,4-dinitrophenyl)octylamine Derivatives (DNPO)

Method I: Maltohexaose (~ 0.7 mg) was added to a 1 dram vial, ~3 equivalents octylamine and 4 equivalents sodium cyanoborohydride, then 0.2 mL 5% acetic acid in a solution of water/methanol (3:1) were added. The Teflon-lined vial was capped and heated at 70° C for 1 hour. Subsequently, the solution was blown to dryness. Approximately 3 equivalents triethylamine, 2 equivalents 2,4-dinitrofluorobenzene, and 0.2 mL

methanol/water (4:1) were added consecutively, and the mixture heated at 45° C overnight. Samples were analyzed directly without purification.

Method II: Maltohexaose (~ 0.7 mg) was placed to a Teflon-lined screw-cap vial, 3 equivalents octylamine, and 4 equivalents of sodium cyanoborohydride. Acetic acid (5%) 0.2 mL in a solution of water/methanol (3:1) was then added and the vial capped and heated at 70° C for 1 hour, and blown to dryness. Afterward, 0.1 mL of 0.2 *M* sodium bicarbonate and 2 equivalents of 2,4-dinitrofluorobenzene in 0.5 mL methanol were added (2). The mixture was heated for 20 minutes at 60° C, and subsequently 0.2 mL of 0.2 *M* sodium hydroxide in dioxane containing trace amounts of water to dissolve the base was added. The mixture was then heated for 1 hour at 60° C. After cooling to room temperature, 0.5 mL water and 0.5 mL ethyl acetate were added. The contents of the aqueous layer were adsorbed on a C18 reversed-phase cartridge column. The column was eluted with 1 mL water, 1 mL water/methanol (1:1), and 1 mL methanol. All three fractions of elutant were analyzed by FAB-MS. Most of the carbohydrate was contained in the water fraction.

2.1.2.2. Preparation of N,N-Dansyloctylamine Derivatives

Maltohexaose (~ 0.5 mg) was placed in a 1 dram vial, ~3 equivalents octylamine and 4 equivalents sodium cyanoborohydride, then 0.2 mL 5% acetic acid in a solution of water/methanol (3:1) wrer added; the Teflon-lined vial was capped and heated at 70° C for 1 hour. Subsequently, the solution was reduced to dryness with a stream of nitrogen. Approximately 3 equivalents triethylamine, 2 equivalents dansyl chloride (5-N,N-dimethyl naphthalene sulfuric chloride), and 0.2 mL methanol/water (4:1) were

added consecutively, and the mixture heated at 45° C overnight. Samples were analyzed directly without purification.

2.1.3. Instrumentation

2.1.3.1. High Performance Liquid Chromatography

Reverse-phase HPLC analysis was accomplished using a Waters (Milford, MA) 600 multisolvent delivery system equipped with a Kratos (Ramsey, NJ) Spectroflow 783 programmable absorbance detector set at 392 nm (for DNPO). The separation of derivatized oligosaccharides was achieved in approximately 100 minutes with a Beckman (San Ramon, CA) Ultrasphere ODS column (4.6 mm x 25 cm) with a 5 μm particle size. Elution was performed at a flow rate of 1.0 mL/min at 50° C using an isocratic gradient, with a mobile phase consisting of water and acetonitrile (7:3).

2.1.3.2. Mass Spectrometry

FAB-MS was performed on a JEOL HX110 (Peabody, MA) double-focusing mass spectrometer (EB configuration) equipped with a high-field magnet operated in the negative and positive ion mode. Ions were produced by bombardment with a beam of Xe atoms (6 kV). The accelerating voltage was 10 kV, and the resolution was set at 1000 or 3000 according to the mass range of interest. The samples were dissolved in a 1:1 (v/v) H_2O-MeOH solution; generally 1-1.5 μL sample was mixed with 1 μL of glycerol/thioglycerol (1:1) or triethanolamine on the FAB-MS stainless steel probe tip. The solvent was removed by pumping before analysis. Calibration was performed using Ultramark (443 and 6121) or $(CsI)_nI^-$ cluster ions, depending on the mass range of interest. A JEOL DA-5000

data system recorded the spectra. The sample was scanned at a scan slope of 2 min from m/z 0-3000. Data presented were acquired in a single scan and found to be reproducible.

2.2. New Approaches to Sequencing Oligosaccharides by Negative Mass Spectrometry

2.2.1. N, N-(2,4-dinitrophenyl)octylamine Derivatives for Negative Mass Spectrometric Characterization of Maltohexaose

There are two methods for preparing the DNPO derivatives of oligosaccharides. The process described in Method II for introducing the dinitrophenyl group is a classical one (2). In this method, however, the sodium salts complicate the mass spectra with many salt adduct peaks. In addition, the salt can totally depress the peaks in the mass spectra. This was also the case when too large an excess of sodium cyanoborohydride was employed. Instead of sodium hydroxide and sodium bicarbonate, we used triethylamine (TEA) as the base (Figure 2.1). When an adequate amount of this base was used, very clean sequence peaks were obtained with the elimination of all salt adducts. Stoichiometric control of the chemical reaction made the purification of the sample before analysis unnecessary, thus avoiding loss of material. The experiment was performed on 0.12 mg (~1.9 x 10-7 mole) of maltohexaose; the product was detected at a level of 3 pmole. In order to test for completeness of derivatization, ^{1}H NMR spectra of the derivatized products were obtained. Completeness of derivatization was indicated by the complete disappearance of the anomeric proton signals of the reducing residue of the oligosaccharides. In the case of maltose, these signals appeared at $\delta 5.3$ ($J = 3$ Hz) and $\delta 4.8$ ($J = 7$ Hz) for α and β anomers respectively. Similar results were obtained with maltohexaose. In the case of maltodextrins complete disappearance of the most downfield protons was also observed. The sensitivity and dynamic range of the NMR spectrum could easily accommodate quantitation of

anomeric resonances due to the low abundance, high molecular weight oligomers.

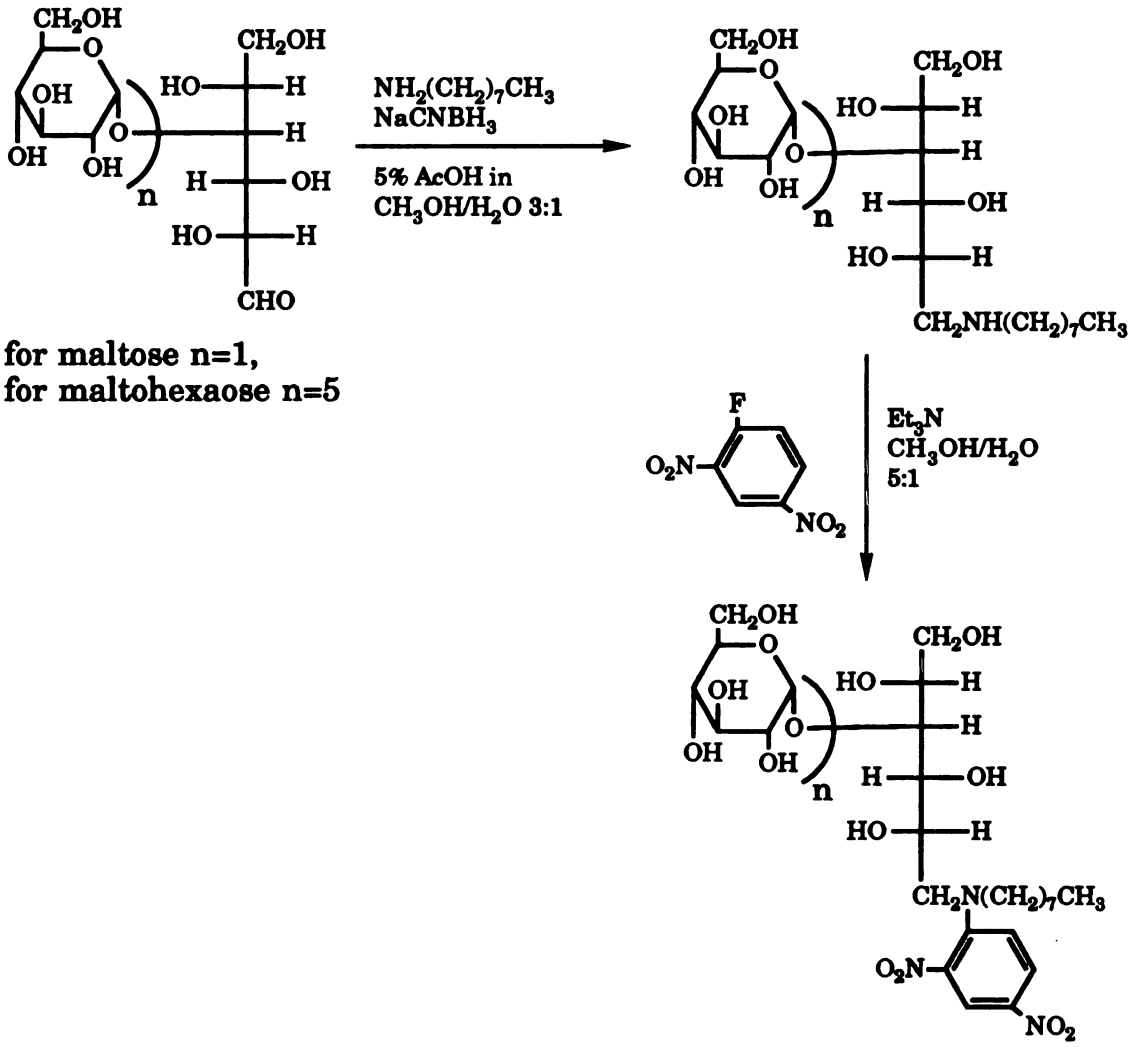


Figure 2.1: The procedure for making DNPO derivatives.

The negative ion mode FAB mass spectrum of the N,N-(2,4-dinitro)phenyloctylamine (DNPO) derivative of maltohexaose (molecular weight 1269, Fig. 2.2) showed a major ion at m/z 1268 corresponding to the $(M-H)^-$ ion which arises from the loss of a proton from the parent ion. This

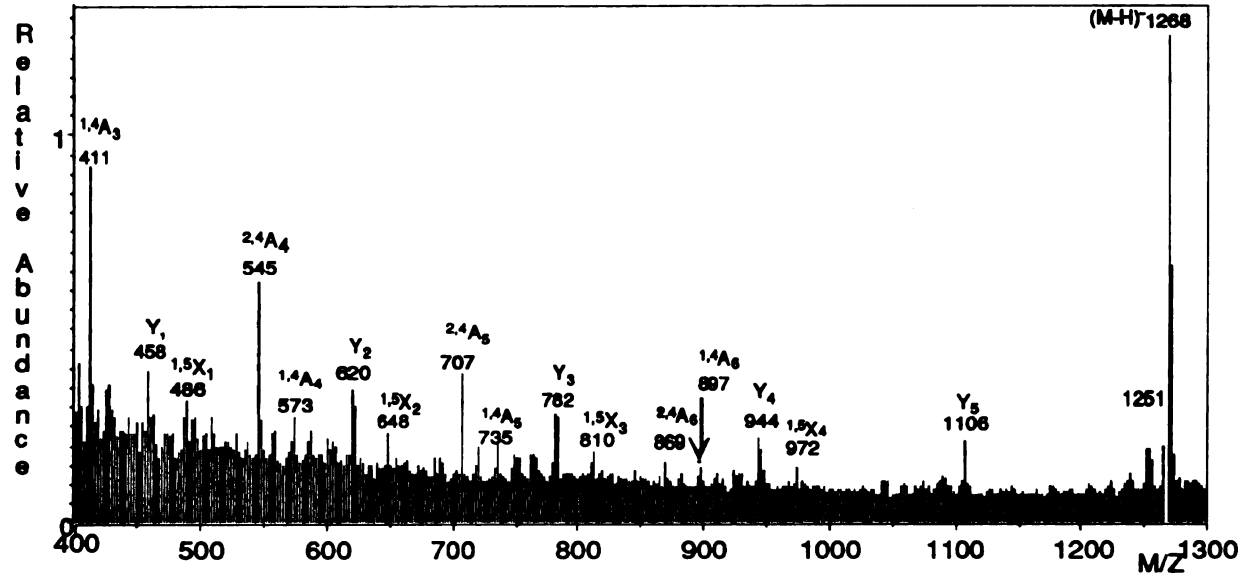


Figure 2.2: Negative mass spectrum of N,N-(2,4-dinitrophenyl)octylamine (DNPO)-derivatized maltohexaose.

mass spectrum displayed four different types of sequence ions. The mechanisms of formation for these ions are shown in Figure 2.3 a-d. Some of the sequences produced a fragmentation pattern involving two-bond ring cleavages within the cyclic sugar units (3, 4). The predominance of two-bond ring fragmentation over glycosidic bond cleavage following ionization from an infrared laser desorption was observed by Cotter *et al.* as well (5, 6). The nomenclature for the sequence fragments is based largely on that introduced by Domon and Costello (7). The Y_n sequence corresponds to the loss of unhydrohexose residues by an elimination process involving loss of the hydrogen atom at C-2 and cleavage between the glycosidic oxygen and the anomeric carbon of the leaving residue at each glycosidic bond. Each loss corresponds to 162 mass units (Figure 2.3 a). The second sequence, $^{2,4}A_n$, originates from cleavage at the non-reducing end by a concerted six

centered scission, which can be explained as pericyclic fragmentation of the penultimate ring (8) to give an initial loss of 222 followed by subsequent losses of 162 mass units, the charge being retained on the fragment

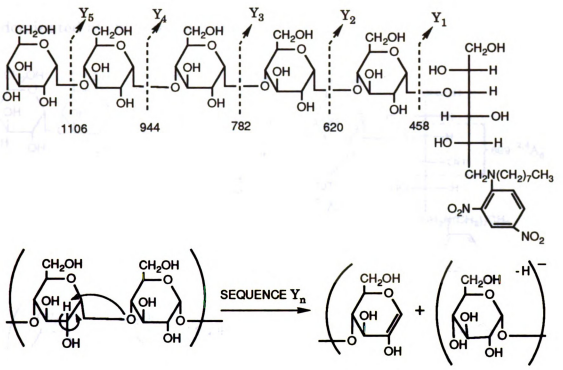


Figure 2.3a: The Y_n sequence from DNPO-derivatized maltohexaose.

containing the non-reducing end (Figure 2.3 b). The sequence $^{1,5}X_n$ was in relatively low abundance. Formation of this series of fragments is facilitated by the proton at the 4 position undergoing a 1,3 hydride transfer accompanied by scission of the C1-C2 bond and the C5-O bond of the terminal residue (Figure 2.3c). The peaks representing ions of this series were also separated by 162 mass units, with the charge retained on the fragment containing the reducing end, the ions appeared at a mass-to-charge ratio 28 amu greater than Y_n sequence ions. The sequence ions, $^{1,4}A_n$, were minor and could be rationalized by formation of an anion at the

C-2 hydroxyl group accompanied by a 3, 5 hydrogen transfer (Figure 2.3 d). Again, members of this sequence were characterized by a difference of 162 mass units. In the maltohexaose spectrum, the Y_n sequence was the most abundant in the high mass end, but the $^{2,5}A_n$ and $^{1,4}A_n$ sequences predominated at the low mass end.

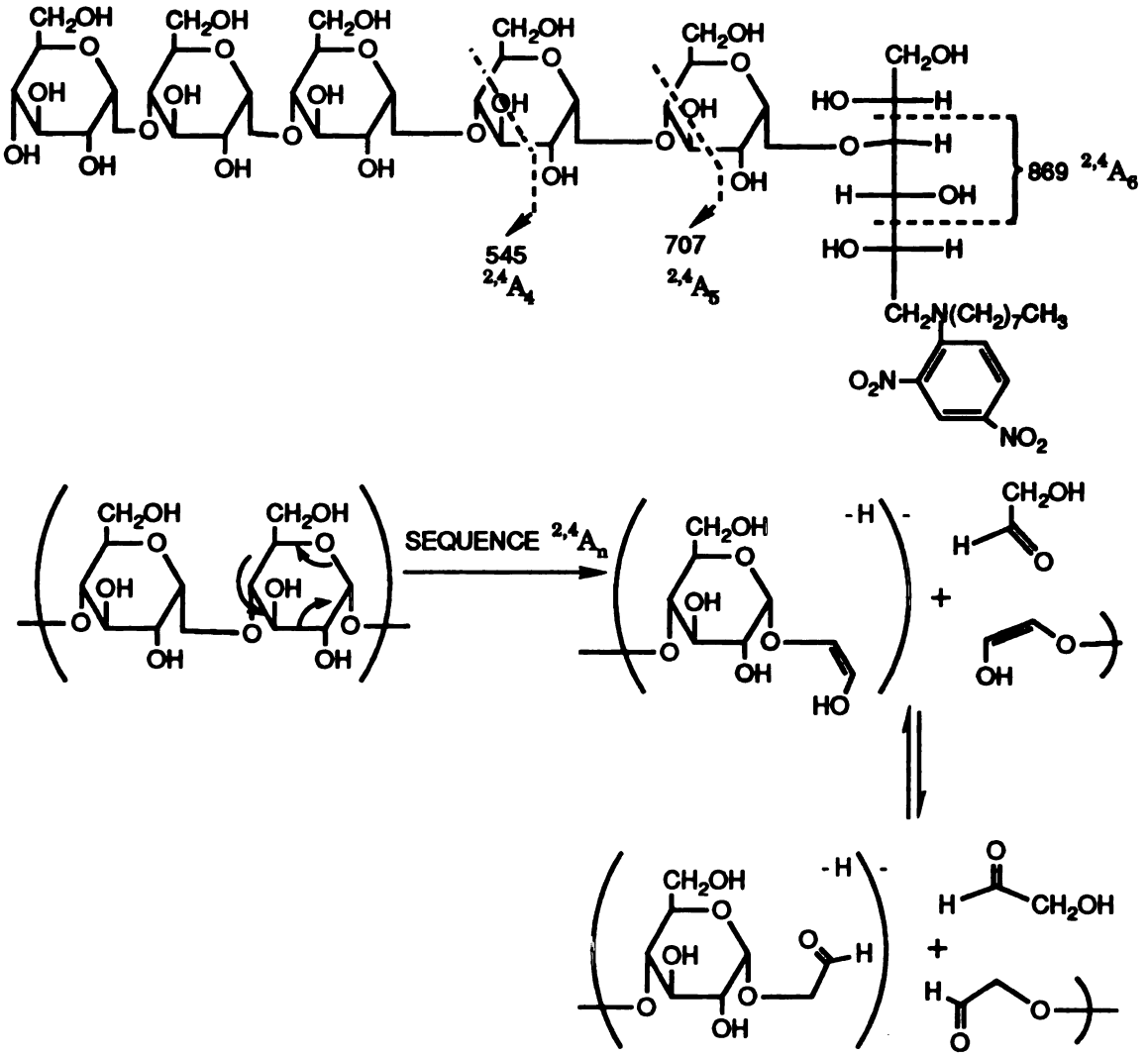


Figure 2.3b: The sequence $^{2,4}A_n$ of DNPO-derivatized maltohexaose.

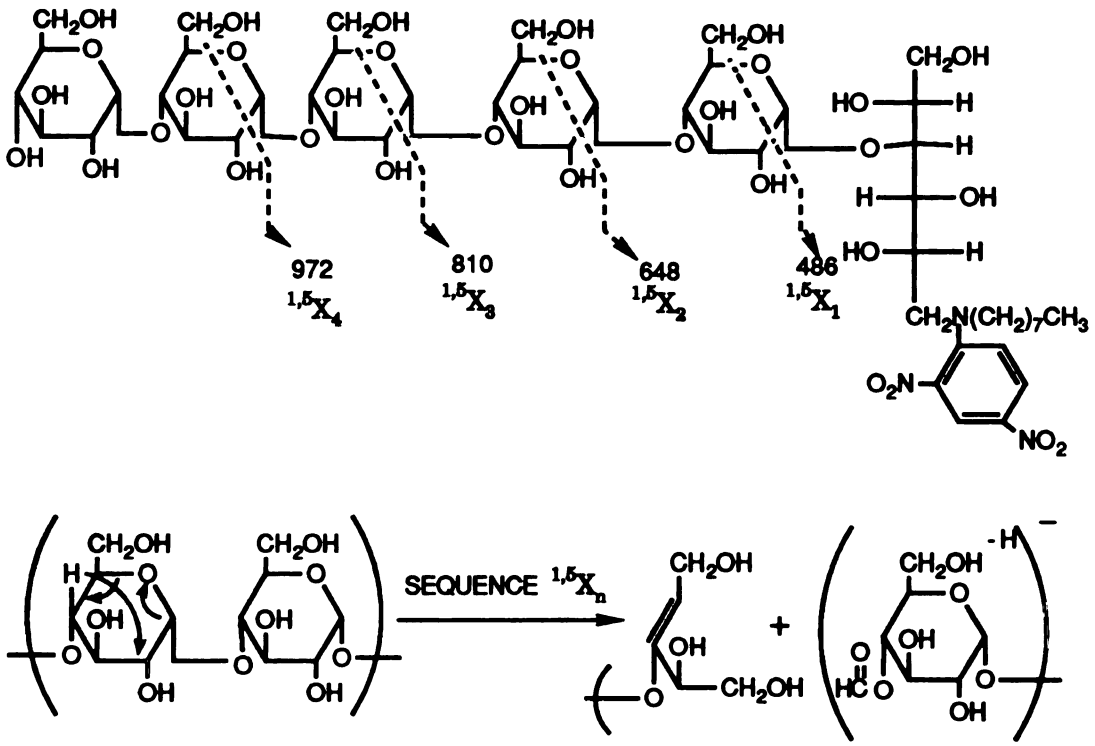


Figure 2.3c: The sequence $^{1,5}X_n$ of DNPO-derivatized maltohexaose.

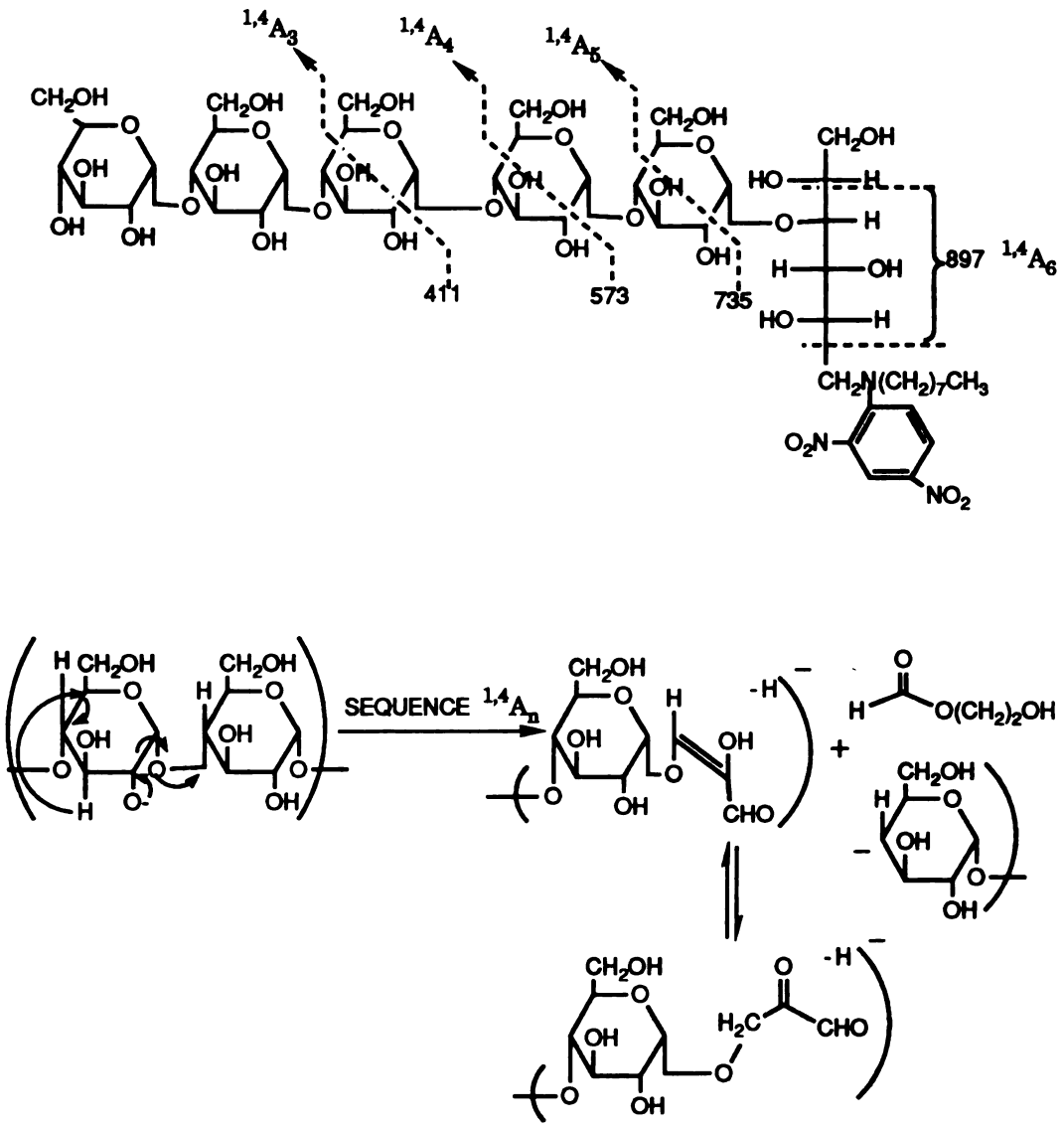


Figure 2.3d: The sequence $^{1,4}A_n$ of DNPO-derivatized maltohexaose.

2.2.2. N, N-(2,4-dinitrophenyl)octylamine Derivatives for the Isolation, Purification and Mass Spectrometric Characterization of Oligosaccharide Mixture Maltodextrin

A mixture of starch-derived maltodextrin of a degree of polymerization equal to 1-16 was derivatized using essentially the same method as that used for maltohexaose. The experiment was started at 0.14 mg (~7.9 x 10^{-8} mole).

The ease of preparation of this derivative and its visibility by UV detection in HPLC makes it an excellent candidate for routine HPLC isolation of trace amounts of carbohydrate materials. The mixture was analyzed by reverse-phase chromatography using an acetonitrile/water (3:7) solvent system. This allowed separation of the mixture of derivatized oligomers into 16 well-resolved peaks with a degree of polymerization in descending order from 16 to 1 with increasing retention time. The assignments were verified by co-injecting the mixture with derivatized maltohexaose which co-eluted with the eleventh peak from the point of injection. The peak of degree of polymerization 16 was determined to represent less than 0.1% of the total mixture based on peak integration.

The mixture was analyzed directly by negative ion FAB-MS (Figure 2.4 a-b). The sensitivity was as low as 0.38 nmole. The most prominent sequence ions observed between 400 and 900 amu correspond to the $^{2,4}A_n$ series of ions. For any given component of the mixture, Y_n series of ions were isobaric with $(M-H)^-$ molecular ions of lower homologues which were also quite prominent. The composite molecular ion Y_n series predominated between 900 and 3000 amu, and a $(M-H)^-$ ion was observed for the highest degree of polymerization component containing 16 glucose units (m/z 2888). This component represents less than 0.1% of the total mixture,

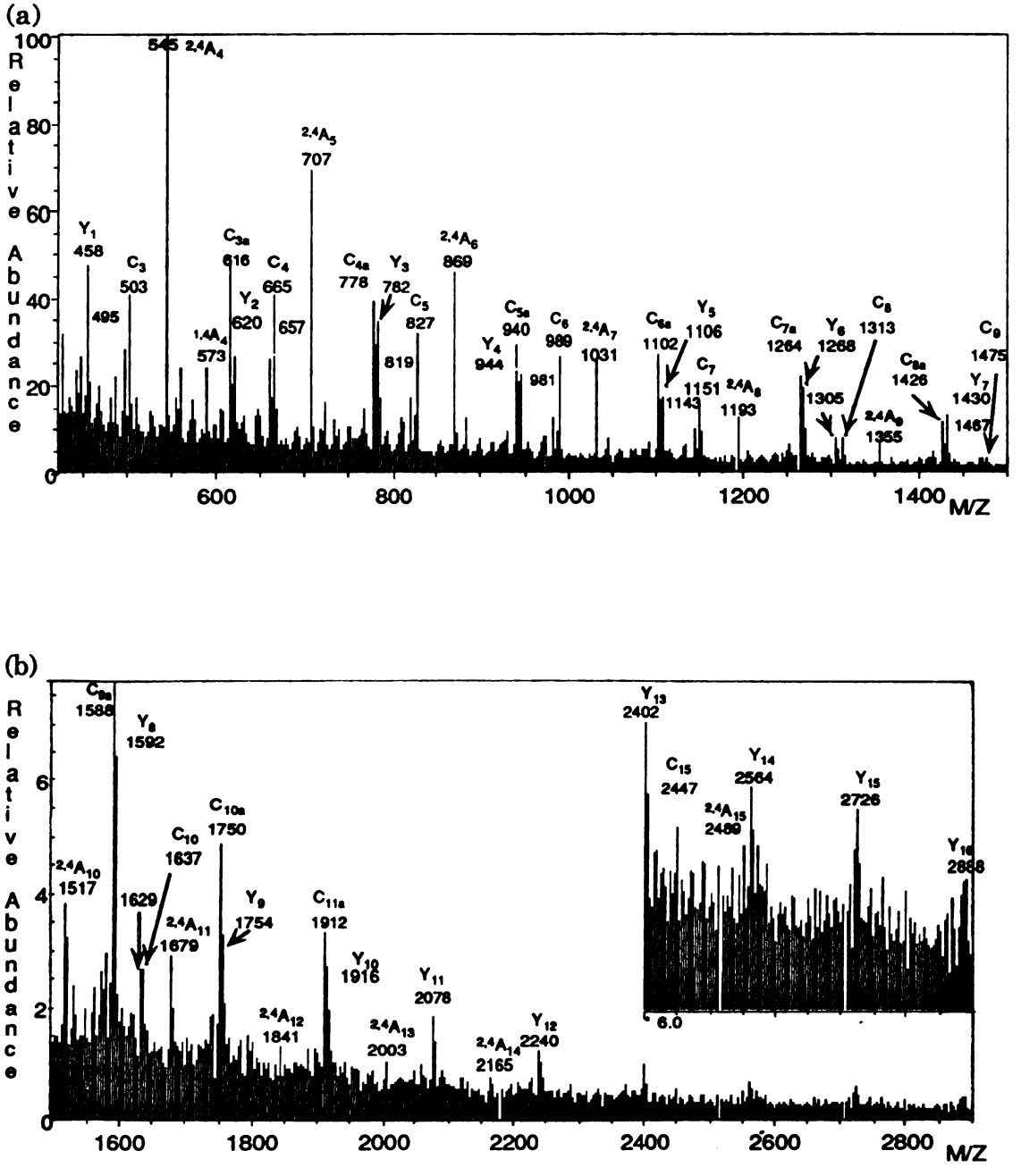


Figure 2.4 a-b: Negative mass spectra of DNPO-derivatized maltodextrin.

which means that it could be detected at levels as low as 0.1 picomole. The limits of sensitivity are far superior to those of the amino benzoic acid ethyl or octyl ester method (9).

In Figure 2.4, a cluster of ions was observed at positions corresponding to the isobaric Y_n series from the (M-H) ion and the (M-H) ion of lower homologues. These clusters were attributed to the fact that a significant proportion of M^- ions are formed by electron capture of parent species. These ions can then break down to give another species two mass units lower by loss of hydrogen, probably in the order of H^- followed by H^+. In fact, with some analytes, using this method we have observed only M species and no $(M-H)^-$ species. This is not surprising, given the high electron capture cross section of the dinitrophenyl group.

A sequence of ions 37 amu greater than the Y_n sequence was observed and is due to a cyanoborohydride adduct from the parent species of the Y_n sequence. This was not observed in samples purified by ion-exchange chromatography. In the low mass region of the spectrum, a C_n series of ions was quite evident. This series involves cleavage of the glycosidic bond, with retention of the glycosidic oxygen atom by the species formed from the non-reducing end (Figure 2.5). The C_{na} series is probably due to an adduct of the analyte with matrix compound triethanolamine followed by the loss of two H_2O molecules from the C_n series. These kinds of adducts are observed very often in large molecules. The spectrum was reproducible, and a similar series was also found during analysis of other starch oligomers. These will be discussed later.

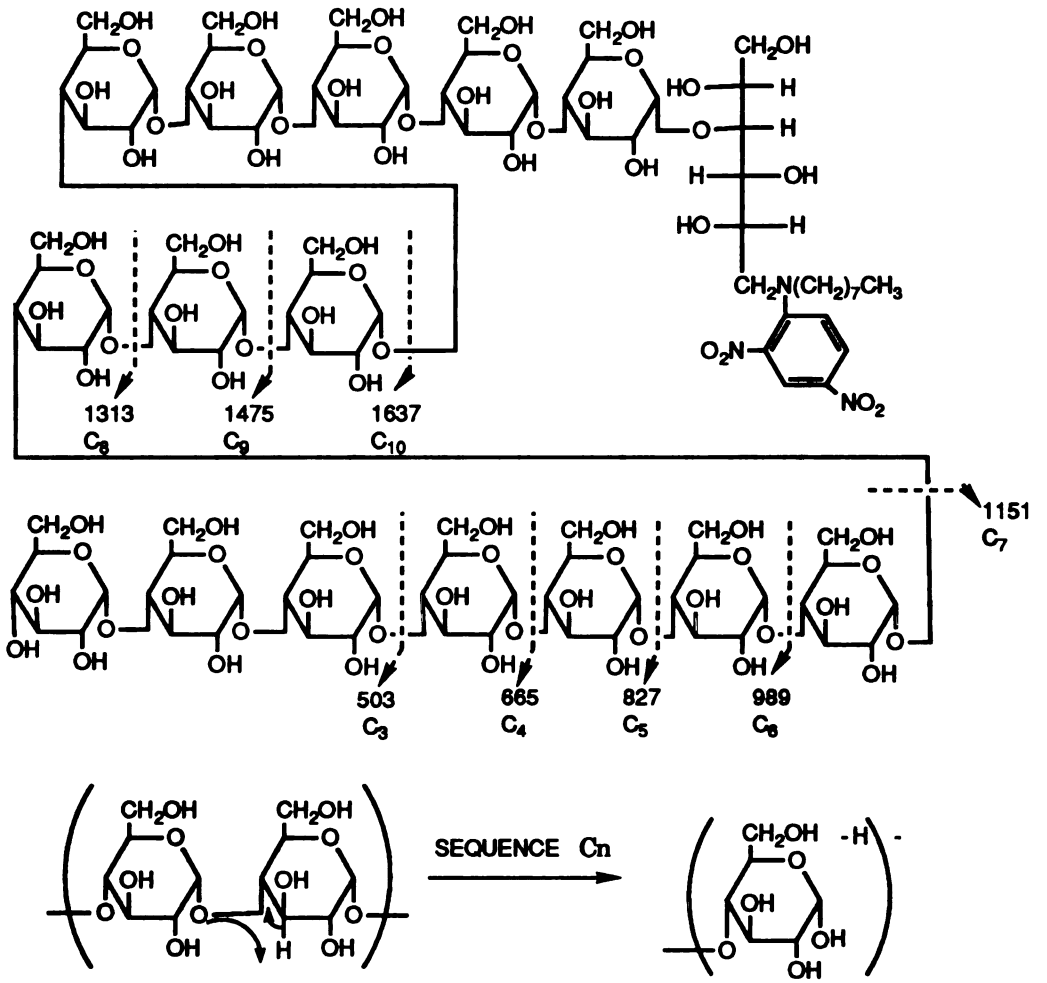


Figure 2.5: Sequence C_n of DNPO-derivatized maltodextrin.

2.2.3 N,N-(2,4-dinitro)phenyloctylamine Derivative for the Negative Mass Spectrometric Characterization of an Oligosaccharide from a Bacterial Cell Surface

As a further test of the efficiency of this derivatization method for enhancing signal intensity, 1 nmol of a previously characterized (1) oligosaccharide from *Rhizobium trifolii* 843 capsular polysaccharide (Figure 2.6) was derivatized and analyzed.

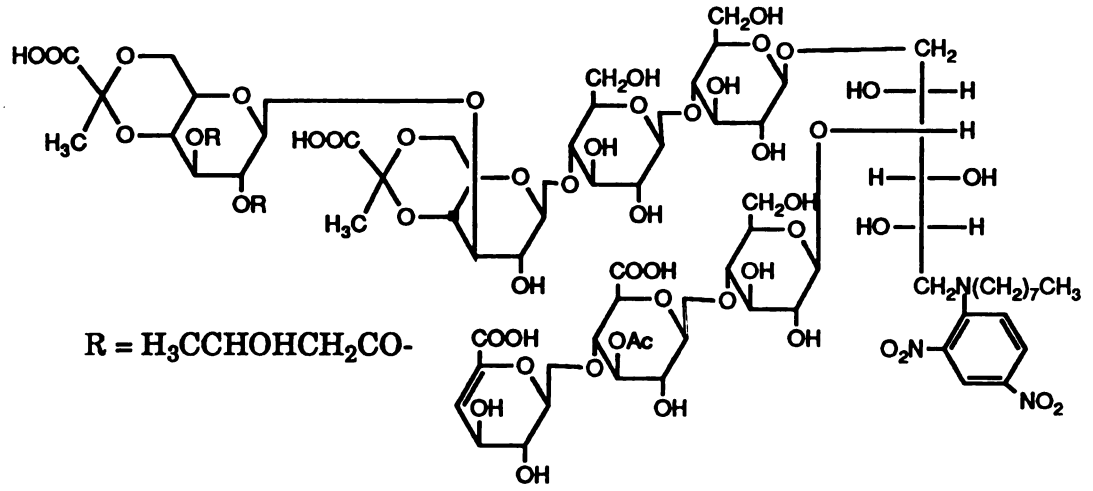


Figure 2.6: Structure of previously characterized and derivatized oligosaccharide from *Rhizobium trifolii* 843 capsular polysaccharide.

The derivatization was performed on 290 μg (~ 1.7 x 10^{-7} mole) of oligosaccharide, and 3 nmole of derivatized material was placed on the probe. The derivatization was expected to result in loss of the acetyl and 3-hydroxybutyryl groups because of the basic conditions employed. A sequence of ions at m/z 1742, 1764, 1786, 1808, 1830, and 1852 corresponding to $(M-H)^-$, $(M+Na-2H)^-$, $(M+2Na-3H)^-$, $(M+3Na-4H)^-$, $(M+4Na-5H)^-$ and (M+5Na-

6H)$^-$, respectively, was observed (Figure 2.7). Another series of ions 152 amu lower corresponded to the loss of a carbohydrate group by a pericyclic mechanism, most likely the one illustrated in Figure 2.8, and yielding pyruvate, formaldehyde, and an unsaturated glycosyl residue with a shift of 152 amu in the sequence of ions .

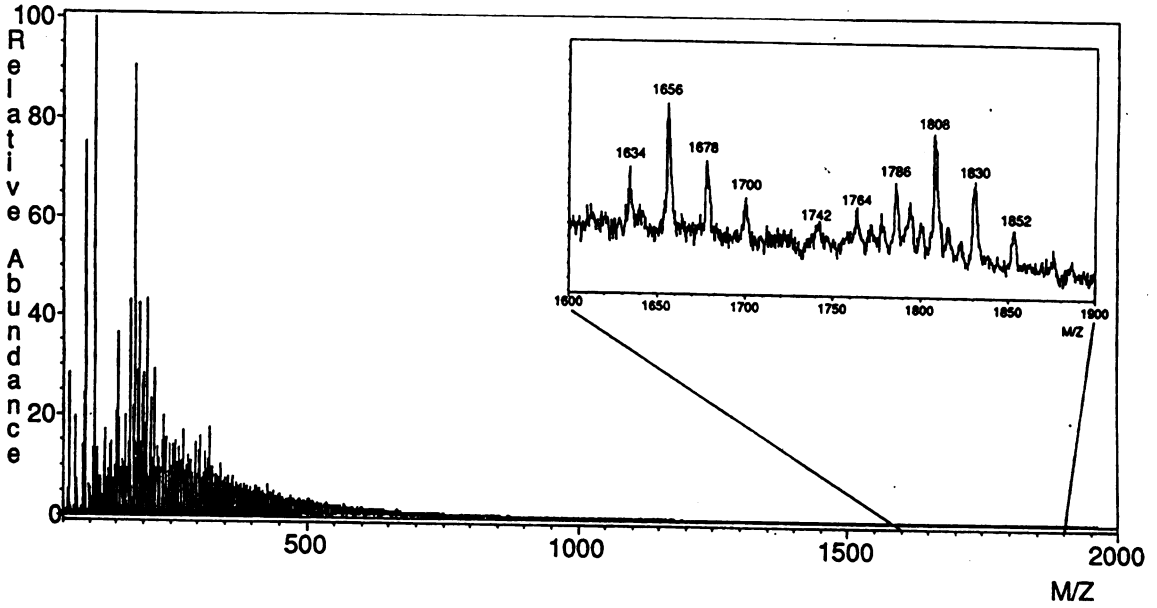


Figure 2.7: Negative mass spectrum of previously characterized, and derivatized oligosaccharide from *Rhizobium trifolii* 843 capsular polysaccharide.

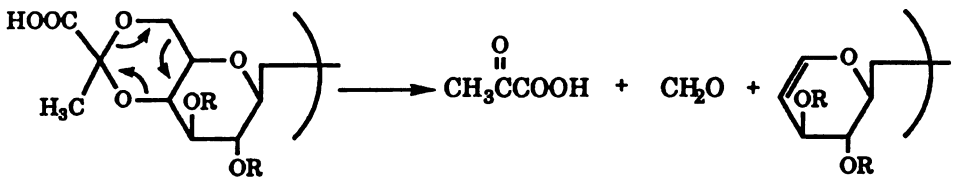


Figure 2.8: Mechanism for rearrangement of acetal bond to produce pyruvate, formaldehyde and an unsaturated glycosyl residue.

2.2.4. Sequence Specific Cleavage of N,N-(2,4-dinitrophenyl)octylamine Derivatives of Hetero and Homooligosaccharides by Negative Ion Fast Atom Bombardment Mass Spectrometry

N,N-(2,4-dinitrophenyl)octylamine (DNPO) derivatives of N,N',N"-triacetylchitotriose, lacto-N-fucopentaose, lacto-N-difucohexaose, and isomaltohexaose were analyzed by negative ion FAB-MS. The lacto-N-fucopentaose and lacto-N-difucohexaose were isolated from human milk. Those compounds were used to test the efficiency of the DNPO derivative in providing sequence information on heterooligosaccharides by (-) FAB-MS analysis. Derivatization increased both the ion intensity and efficiency of useful fragmentation. Sensitivities observed from the negative FAB-MS of these naturally occurring oligosaccharides were often better than 1 pmole. Sequence specific cleavage, which is useful both for distinguishing isomers and the analysis of glycosidic linkages, was observed for all of the oligomers.

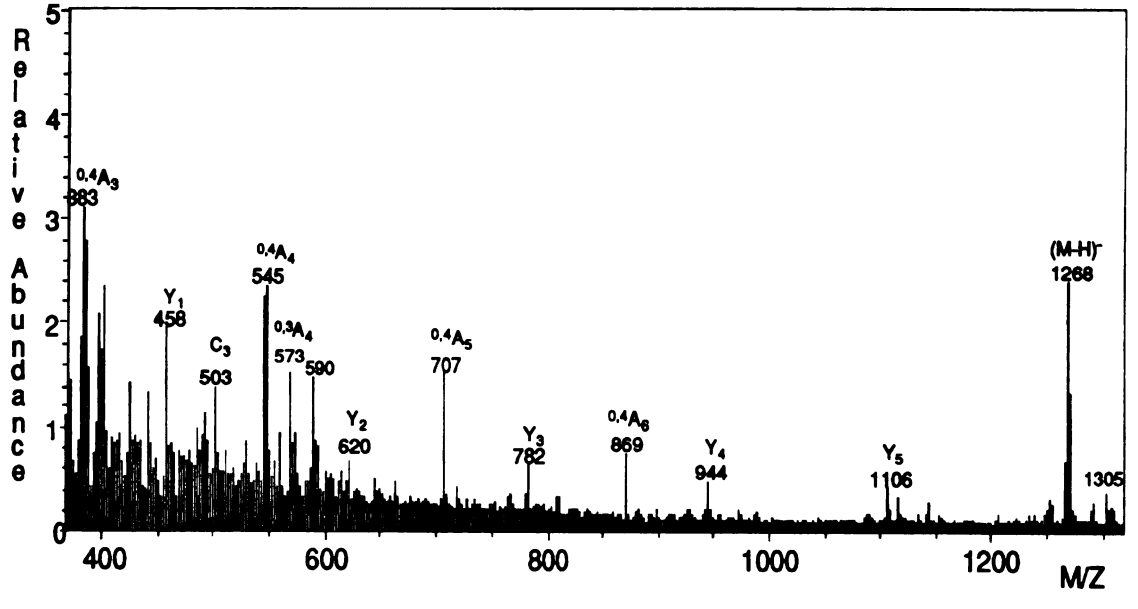


Figure 2.9: Negative FAB-MS of DNPO-derivatized isomaltohexaose

2.2.4.1 Negative FAB-MS of Isomaltohexaose

The derivatization was performed on 41 μg isomaltohexaose (~ 4.1 x 10^{-8} mole), and 0.1 nmol of material was put on the probe for negative FAB-MS analysis. The negative FAB-MS of isomaltohexaose is shown in Figure 2.9. A sharp molecular ion appears at m/z 1268. The lower intensity peak at m/z 1305 is from the cyanoborohydride adduct to the molecule. Comparing the spectrum of isomaltohexaose to Figure 2.2 of maltohexoase, in Y_n-type fragmentation, both 1-4 and 1-6 linked sugars can lead to the same mass losses (-162), regardless of the anomeric configuration (α or β) and, therefore, Y_n-type fragmentation does not appear to be a distinguishing sequence between 1-4 and 1-6 linked isomers. The mechanism for fragmentation of isomaltohexaose is shown in Figure 2.10, which can be compared to that of maltohexsaose in Figure 2.3 a. The two isomers undergo Y_n-type cleavage to yield the same sequence. The $^{0,4}A_n$ sequences in isomaltohexaose (compare Figure 2.3 b) appear at same mass shift with the $^{2,4}A_n$ sequence in maltohexaose. Because of the 1-4 ring linkages in the maltohexaose and the 1-6 linked sugar in isomaltohexaose, the Y_n-type sequences appearing at the same mass shift can be cleaved from different bond order in different ring linkages (Figure 2.11).

Generally, fragmentation and ionization are involved in the two bonds ring cleavage within the cyclic sugar (34-37). The sequences $^{1,4}A_n$ and $^{1,5}X_n$ appearing in the FAB-MS of maltohexaose did not appear in the FAB-MS of isomaltohexaose. The differences in two bonds fragmentation observed for these two isomeric structure, maltohexaose and isomaltohexaose, indicate that this method is potential by quite useful detecting subtle structural differences between oligosaccharides.

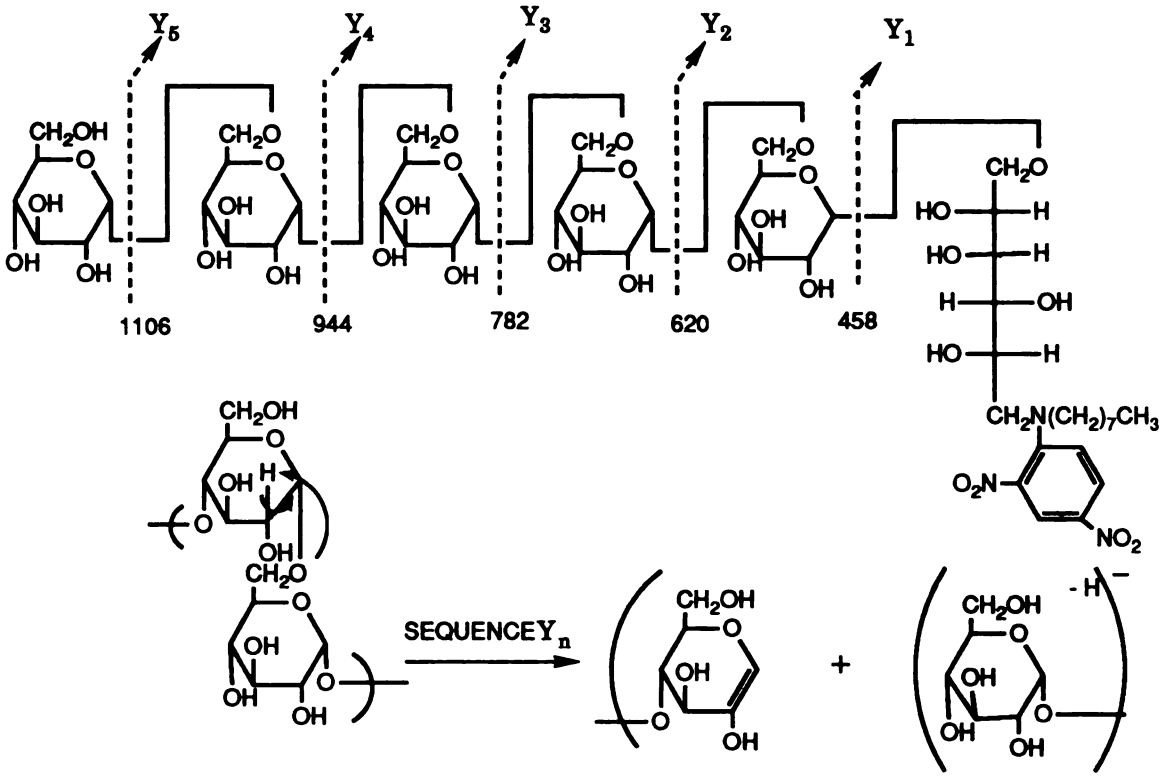


Figure 2.10 Y_n-type fragmentation of derivatized isomaltohexaose.

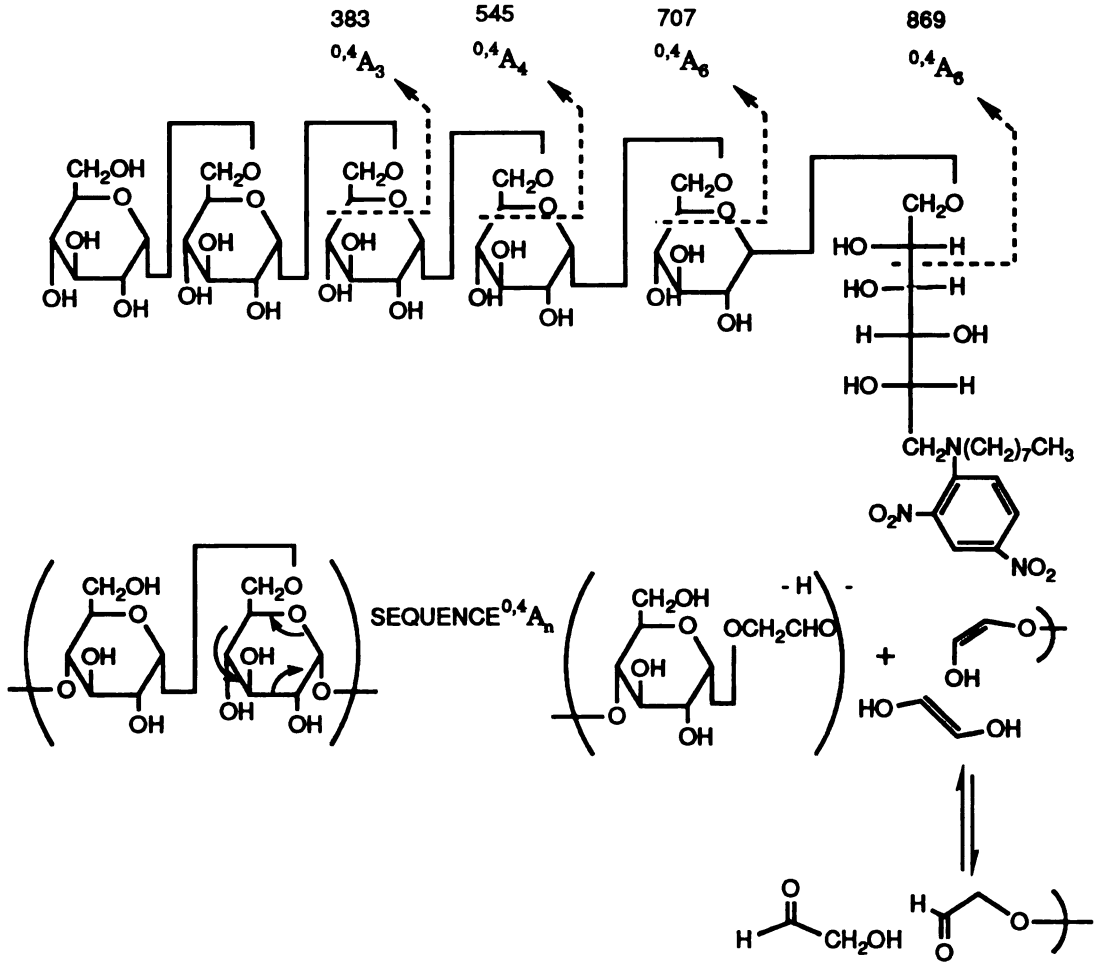


Figure 2.11: Cleavage pattern for generation of $^{0,4}A_n$ sequence in isomaltohexaose.

2.2.4.2. Negative FAB-MS of DNPO Derivative of N,N',N"-triacetylchitotriose

FAB-MS sequence anions were also observed for the DNPO derivative of N,N',N"-triacetylchitotriose, a glucosamine-containing trisaccharide. The structure of the derivative is shown in Figure 2.12.

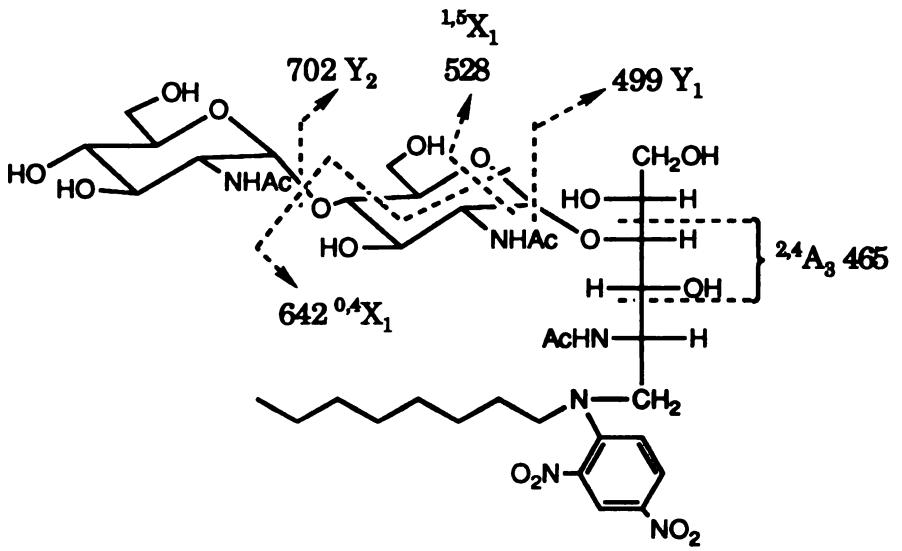


Figure 2.12: The structure of the DNPO-derivative of N,N',N"-triacetylchitotriose.

The derivatization was performed on 69 μg (~ 1 x 10^{-7} mole), and 1 nmol of material was put on the probe. The negative FAB mass spectrum for the DNPO derivative of N,N',N"-triacetylchitotriose is shown in Figure 2.13. The molecular weight of derivatized N,N',N"-triacetylchitotriose is 906, and the highest molecular weight species in the spectrum has the same mass. This molecular ion was resulted from electron capture by the dinitrophenyl group. This same phenomenon has been observed for other extremely electron-deficient centers. This appears more favorable than proton loss for this molecule. In Figure 2.13, the Y_n sequence is observed, yielding the linkage

information for this homotriose. Each Y_n sequence peak is separated by 203 amu, an anhydro-N-acetyl glucosamine residue. The Y_n sequence ions appeared at m/z 499 and 724, and 906, while the peak at m/z 724 possibly is a sodium adduct to Y_2 of m/z 702.

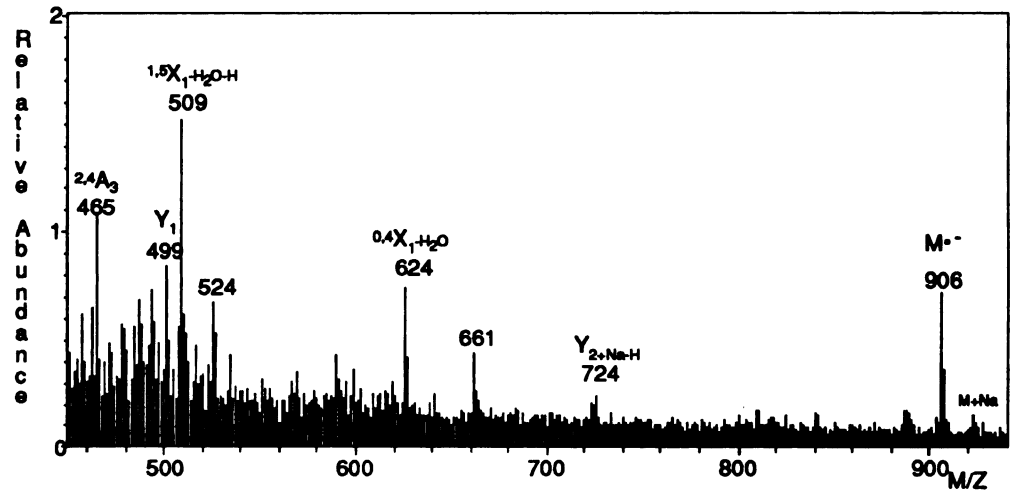


Figure 2.13: Negative FAB-MS of the DNPO-derivative of N,N',N"-triacetylchitotriose.

Beside the Y_n sequence ions, there are other peaks in the mass spectrum. The $^{2,4}A_3$ ion is at m/z 465, the ion at m/z 524 arose from m/z 465 plus a sodium and a cyanoborohydride. All of ions of X sequence appeared with loss of a water. The $^{1,5}X_1$ ion with loss of a water is at m/z 509. The $^{0,4}X_1$ ion with loss of a water is at m/z 624, and $^{0,4}X_1$ plus a cyanoborohydride appears at m/z 661. The cleavage pattern of $^{0,4}X_n$ sequence is shown in Figure 2.14. The $^{0,4}X_n$ cleavage proceeds by a 1, 3 hydrogen rearrangement. The cleavage pattern for anion adduct formation by cyanoborohydride is shown in Figure 2.15.

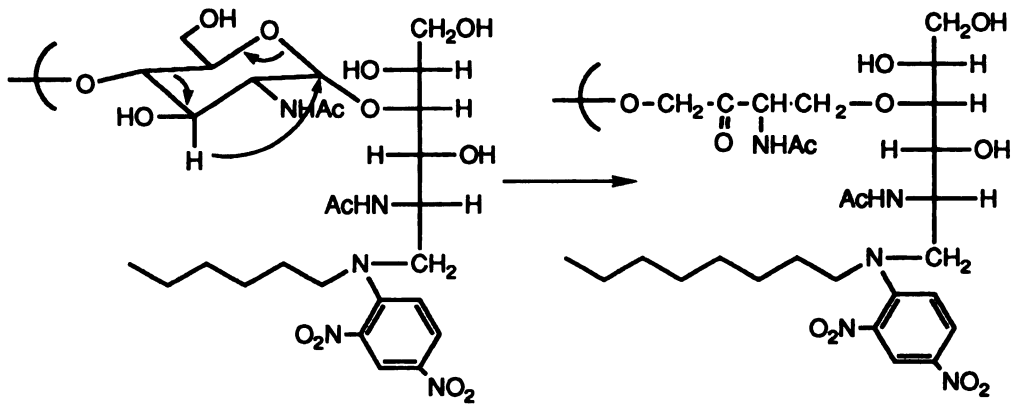


Figure 2.14: The cleavage pattern of $^{0,4}X_n$ sequence.

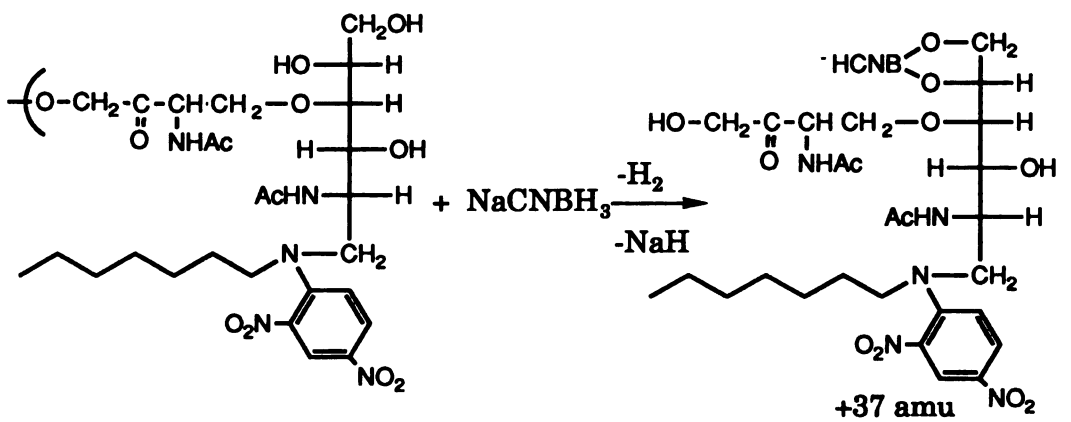


Figure 2.15: The cyanoborohydride adduct formation.

The appearance of peaks corresponding to protonated molecular species is very common in positive FAB-MS. Some investigators have also reported the occurrence of negative salt adducts. Ganguly et al. (10) first reported the anion-adduct formation, the adduct was formed between chloride ions and a glycoside. Prome et. al. (11) also observed abundant chloride-adduct ions mass spectra of glycosides. They observed an ion corresponding

to $(M - H_2O + Cl)^-$ in the negative mass spectrum of a trisaccharide. The ion at m/z 661 in Figure 2.13 is of similar origin $(^{0,4}X_1 - H_2O + CNBH)^-$.

2.2.4.3. The Negative FAB-MS of Lacto-N-Fucopentaose

The structure of its DNPO derivative of Lacto-N-fucopentaose, a heterooligosaccharide from human milk, is shown in Figure 2.16, and its negative FAB-MS is shown in Figure 2.16. Although (as is usually the use for analytes such as these) this spectrum was not obtained from a four-sector tandem mass spectrometer (12), nor by MS/MS link scans, but very good sequence information was obtained from the normal single scan masspectrum. The derivatization was performed

Fucα1-2Galβ1-3GlcNAc β1-3Galβ1-4Glc

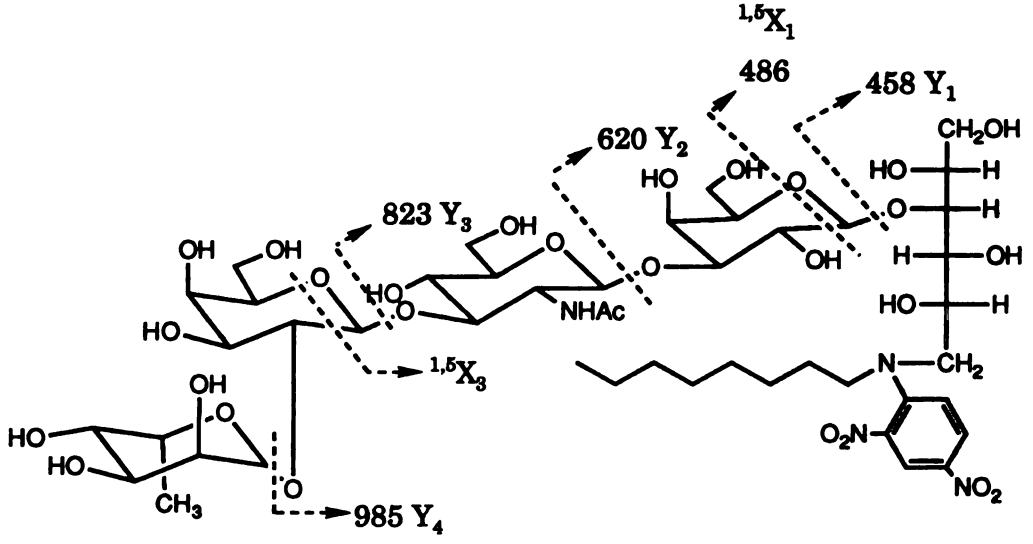


Figure 2.16: The cleavage pattern of the DNPO derivative of lacto-N-fucopentaose.

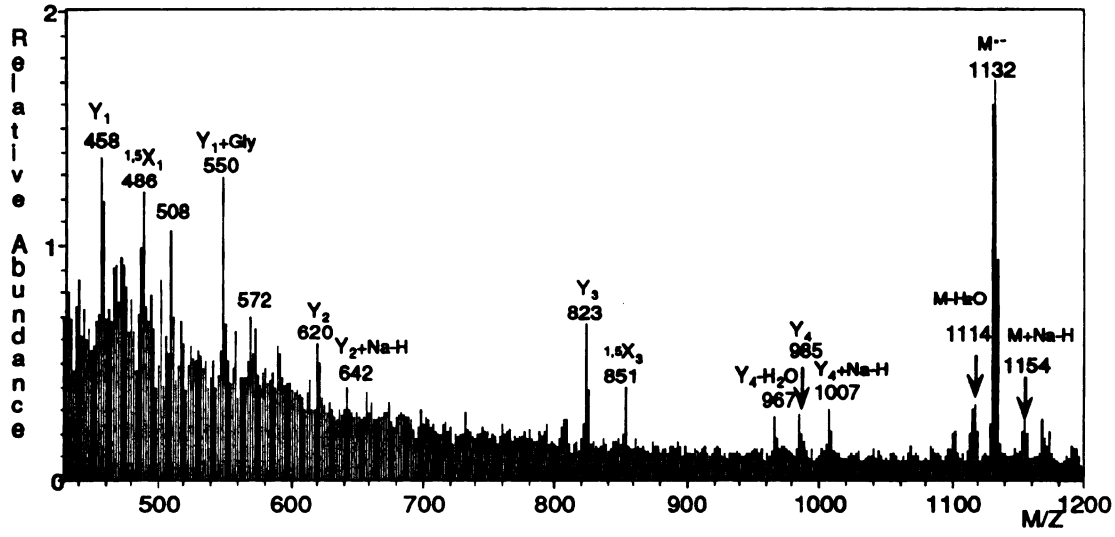


Figure 2.17: Negative FAM-MS of derivatized lacto-N-fucopentaose.

on 33 μg (~ 3.9 x 10^{-8} mole) of material, and ~ 0.3 nmol of derivative was put on the probe.

As in Figure 2.17, the predominant ion in the high mass end of the (-) FAB-MASS spectrum of lacto-N-fucopentaose occurred at m/z 1132. This corresponds to the molecular weight of the compound and indicates that one electron is captured by the molecule to give an $(M + e)^-$ species. The Y_n sequence gave detailed glycosyl sequence information: ion Y_4 at m/z 985 corresponds to a loss of a fucosyl residue (147 amu) from the non-reducing terminus, Y_4 (loss of a water molecule) appeared at m/z 967, and the peak at m/z 1007 is sodium adduct to Y_4, which was rapidly fragmentation from the sodium adduct at the parent ion of m/z 1154. Y_3 at m/z 823 is due to the sequential loss of a fucosyl residue (147 amu) and a galactosyl residue (162 amu). Y_2 at m/z 620 is due to the further loss of an N-acetylglucosaminyl group (203 amu), and the at m/z 642 is sodium adduct of Y_2. Y_1 at m/z 458 is due to the further loss of a galactosyl residue. The peak at m/z 851 is due

to $^{1,5}X_3$ specie. The peak at m/z 572 is due to a sodium adduct to the specie at m/z 550. The peak at m/z 508 is due to a sodium adduct to $^{1,5}X_3$ ion at m/z 486.

The results from FAB-MS analysis of DNPO-derivatized oligosaccharides are quite different from those reported in recent investigations. Cotter *et. al.* (13), based on their study of permethylated and peracetylated derivatives, concluded that glycosyl components which did not have a reducing end did not undergo two-bond ring cleavages and that chemical derivatization generally suppressed fragmentation. Their conclusion precisely pointed out the disadvantages of permethylation and peracetylation methods. Also, they indicated that only β 1-4 and 1-6 linked saccharides showed additional peaks due to the loss of water. None of the α-1-4 and 1-6-linked saccharides showed peaks due to the loss of water. This is not universally true. In our experiments, large molecules showed a greater tendency to lose water. For example, in Figure 2. 17, the additional peaks due to loss of water appeared in the mass range higher than m/z 900.

2.2.4.4 The Negative FAB-MS of Lacto-N-Difucohexaose

Lacto-N-difucohexaose is a branched heterooligosaccharide from human milk. The derivatization of lacto-N-difucohexaose was carried out on 41 μg (~ 4.1 x 10^{-8} mole) of starting material, and 0.1 nmol of final product was put on the probe. The structure of lacto-N-difucohexaose is shown in Figure 2.18 and its (-) FAB-MS in Figure 2.19. In the mass spectrum, the highest ion at m/z 1278 corresponds to the molecular weight of derivatized lacto-N-difucohexaose. This indicates that the 2,4-dinitro phenyl group is again participating in an electron capture reaction. In Figure 2.19, the main sequence ions correspond to the Y_n series and gives complete sequence

Fucα1-2Galβ1-3GlcNAc β1-3Galβ1-4Glc
4
|
Fucα1

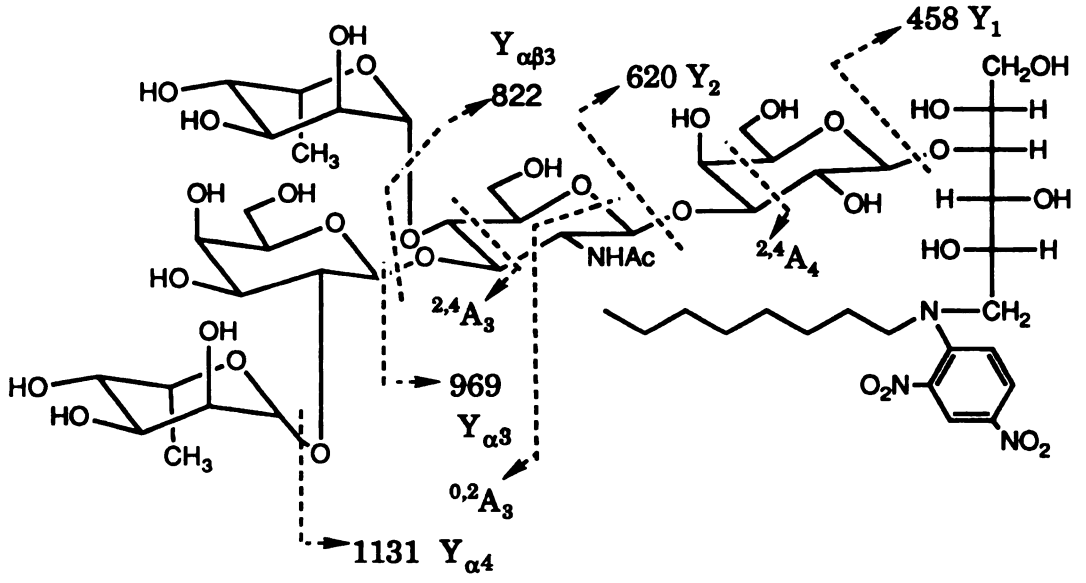


Figure 2.18: The fragmentation pattern of derivatized lacto-N-difucohexaose

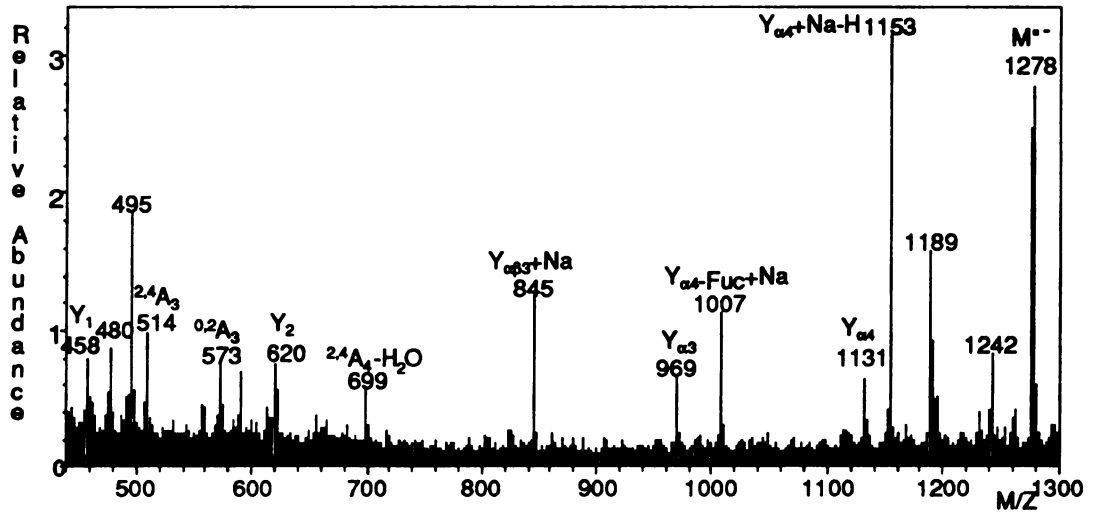


Figure 2.19: Negative FAB-MS of derivatized lacto-N-difucohexaose.

information. The $Y_{\alpha 4}$ ion at m/z 1131 is due to the loss of an anhydrofucosyl group from the non-reducing terminus. The ion at m/z 1153 is sodium adduct to $Y_{\alpha 4}$. The peak at m/z 1007 arises from the loss of the second fucosyl group. The peak at m/z 969 is thought to arise from the $Y_{\alpha 4}$ ion after loss of an anhydrogalactosyl residue. The ion at m/z 845 is due to the cleavage including both α and β branches so-called $Y_{\alpha\beta 3}$ ion. This is due to sequential loss of two fucosyl residues and one anhydrogalactosyl residue from the parent molecule in a metallated (sodium) form. The ion at m/z 620, Y_2 ion, arises from the $Y_{\alpha\beta 3}$ fragment by loss of an anhydroglucosaminyl residue (203 amu). The Y_1 ion is due to the reducing terminus glucose linked to the DNPO group (m/z 458). Other peaks such as m/z 480 and 495, are the sodium and cyanoborohydride adducts to Y_1, respectively. The peak at m/z 1242 is due to the loss of two water moleccules from the parent ion, and the peak at m/z 1189 is the cyanoborohydride adduct to the species at m/z 1153. Some peaks in the lower mass range belong to A_n sequences.

Several sequences are characterized by sodium adduct ions which probably arise from the molecular species after successive electron and Na^+ capture. This species then gives to negative by proton loss ($M^- + Na^+ - H^+$).

It has been noted that the carbohydrate moiety in glycoproteins is essential for their function (14). The naturally occurring oligosaccharides in these molecules are often branched. In this study, the sequence determination of branched carbohydrate derivatives analyzed by (-) FAB-MS demonstrates that this method is an excellent technique for providing useful structural information.

2.2.5. Sequencing of DNPO Derivatized Starch by Negative FAB-MS

2.2.5.1 Introduction

Starch, the storage carbohydrate of the majority of higher plants, occurs as water-insoluble granules that vary in size and shape depending on the species and maturity of the plant. Within these granules, there is normally a mixture of two polysaccharides: amylose and amylopectin. The amylose, amounting to 20-30% of total starch, is largely composed of long linear chains of (1-4)-linked α-D-glucopyranose residues. The major starch component is amylopectin, also a macromolecule containing thousands of shorter branched chains of (1-4) and (1-6) linked α-D-glucose residues.

The structural analysis of the isolated amylose and amylopectin components has been carried out by standard methods based on methylation, periodate oxidation, and partial acid hydrolysis studies (15). The chain length can be studied by the enzymic hydrolysis of the (1-6)-α-D-glucosidic interchain linkages, followed by fractionation of the resulting mixture of linear chains by gel filtration (16). The degree of polymerization is defined as the number of glucose residues per reducing end group, and is normally calculated from physical measurements of the molecular weight.

2.2.5.2 Structure of Amylose

The methods of enzymic degradation are widely used in the structural analysis of amylose. The α-amylase hydrolyses nonterminal (1-4)-α-D-glucosidic linkages in both amylose and amylopectin, giving a mixture of products, while the β-amylase catalyzes the hydrolysis of alternate linkages, giving maltose as the only low-molecular-weight product. None of the above enzymes have any action on (1-6)-α-D-glucosidic interchain linkages; only isoamylase can selectively hydrolyze the 1-6

linkages with no effect on the 1-4 linkages. The amylose has the linear (1-4)-α-D-glucan; most of amylose fraction has secondary chains attached through occasional (1-6)-α branch points and in some species has a few phosphate groups at the C-6 of glucose residues (17).

2.2.5.3. Structure of Amylopectin

The general structure of amylopectin has been known for many years, but the details are still uncertain. Three different diagrammatic representations of the molecule are shown in Figure 2.20 (18). The A chains are linked to the molecule only by the potential reducing group, the B chains are similarly linked and carry one or more A chains, and the C chain ends in the only reducing group R in the molecule. In structure (e), only half of the B chains carry A chains, while the other half of the B chains have their nonreducing end group inside the molecule. A cluster model was further refined by Manners and Matheson (19) in Figure 2.21. In this figure, the majority of the B chains carry more than one A chain, to account for the change in the A:B chain ratio of the partly debranched material. The model in Figure 2.21 is compatible with the physical properties and with various enzymatic degradation studies which show that the branch points are arranged in "tiers" or clusters and are not distributed randomly throughout the macromolecule. The most recent report by Callaghan and Lelievre (20) concludes that the the amylopectin molecules are a large flattened disk shape, consisting of (1-4)-α-D-glucan chains joined by many (1-6)-α-branch points. Takeda *et al.* (21) reported that the phosphate in D-glucose 6-phosphate residues was analyzed by D-glucose 6-phosphate dehydrogenase. The detailed molecular structure of amylopectin seems to have emerged, but the arrangement of the

amylopectin molecules within the starch granule still remains to be determined.

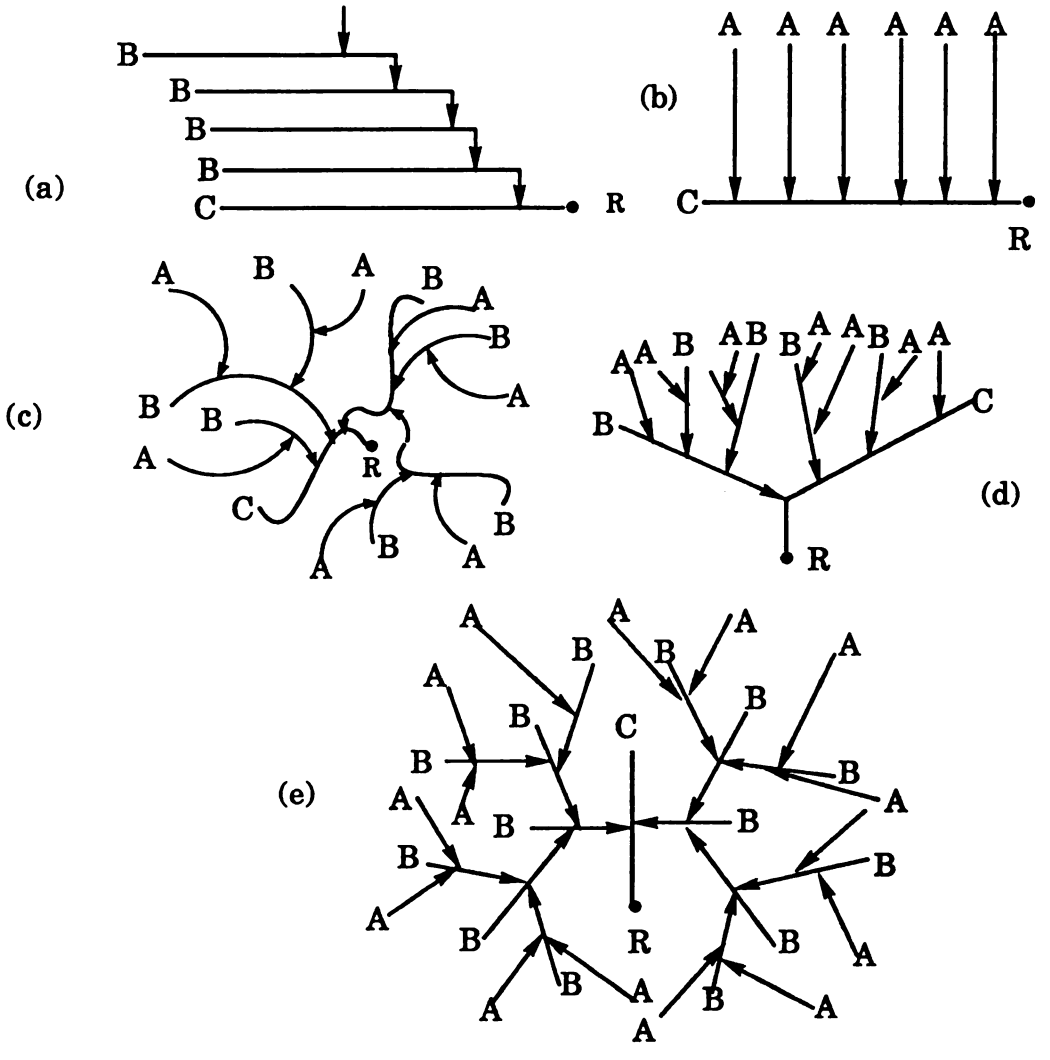


Figure 2.20: The molecular structure of amylopectin proposed (18) by (a) Haworth, (b) Staudinger, (c) Meyer, (d) Meyer, (e) revised Meyer structure. R = reducing end.

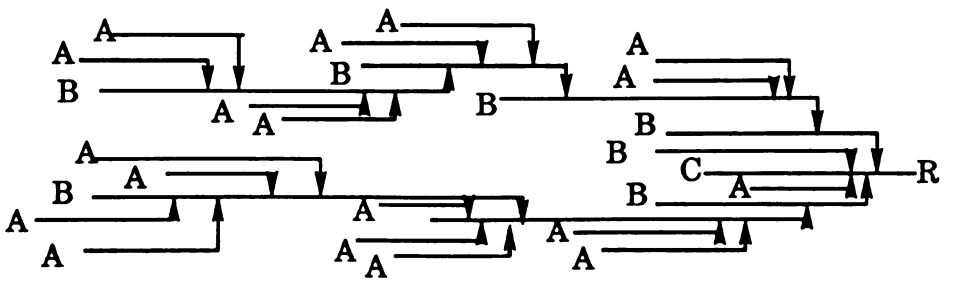


Figure 2.21: The cluster model of amylopectin.

2.2.5.4. Phosphorus in Starch Fractions

Commercially available starches are derived from many botanical sources and contain minor amounts of non-carbohydrate constituents. The content of non-carbohydrate elements associated with starch, in particular phosphorus and nitrogen, show considerable variation. Different phosphorus content in starches is associated with different properties (22). Perhaps this is why considerable research has been conducted to determine the phosphorus content in starches and in their various derivatives.

In their classical studies, Kerr and Severson (23) fractionated non-defatted corn and potato starches. They found 0.013% and 0.088% phosphorus in the more butanol-soluble fraction of corn and potato starches, respectively, and 0.009% and 0.014% phosphorus in the less butanol-soluble fractions, respectively, as compared to 0.020% and 0.081% in the respective parent starches. Posternak (24) detected 3% phosphorus in a hexasaccharide isolated from potato starch after hydrolysis by pancreatic α-amylase. This hexasaccharide yielded D-glucose 6-phosphate upon partial acid hydrolysis, indicating the position of the phosphate. Parrish and Whelan (25) isolated phosphomaltotetraose from potato starch after

digestion by salivary α-amylase. Periodate oxidation showed that the phosphate group is located on the D-glucose unit that is in the third position from the reducing end. Posternak (26) reported that the ratio of phosphate group to D-glucose units was 1:6, which resulted from the calculation that the phosphorus protects a fragment of about 6-7 glucose units against degradation by β-amylase. They also reported that the phosphohexaose is resistant to the action of β-amylase.

Hizukuri *et al.* (27) reported that the phosphorus in starch was estimated to be glucose-6-phosphate, and 60-70% of total phosphorus in potato starch was found to be due to the glucose-6-P residue. The rest of the phosphorus was incorporated as glucose-2-phosphate and/or glucose-3-phosphate. The methods they used were periodate oxidation, borohydride reduction, acid hydrolysis and treatment with alkaline phosphatase. Hizukuri and Tabata (28) isolated three sugar phosphates, 4-O-(α-D-6-phosphoglucosyl)-D-glucose, glucose 3-phosphate, and glucose 6-phosphate, from the α-amylase limit phosphodextrin of potato starch by a partial acid hydrolysis and an ion-exchange resin. From the quantitative yields of glucose 3-phosphate and glucose 6-phosphate, the distribution of esterified phosphate in potato starch was estimated to be 38% on the C-3, and 61% on the C-6 of glucose residues. Their kinetic studies indicated that approximately 1% of the phosphate in starch existed in a form more labile to acid than glucose 3-phosphate, possibly as glucose 2-phosphate. Residues of glucose 6-phosphate are present in some starch fractions, serving as a barrier to β-amylase (29).

2.2.5.5. Negative FAB-MS of DNPO Derivatized Oligosaccharides Digested from Potato Starch

Analysis of the structure of starch is essential for understanding of phosphorus-containing starch. Many processes and treatments have previously been employed on starches, for instance enzymatic digestion, acid hydrolysis, elemental analysis etc. (30). Applications of techniques such as NMR and FAB-MS to starch analysis have not appeared widely.

We have analyzed the N,N-(2,4-dinitrophenyl)octylamine derivative of oligosaccharides resulting from 0.1 *M* TFA hydrolysis of potato starch. By negative FAB-MS analyses, it is possible to see the heterogeneity introduced by the phosphate and the differences between the unsubstituted and the phosphorus-linked carbohydrate chain among the starch-digested oligosaccharides. The (-) FAB-MS spectra are shown in Figure 2.22 a-b.

In Figure 2.22, two main fragmentation patterns, C_n and $^{2,4}A_n$, are observed. Sequence C_n arises from the glucosidic bond cleavage between two sugar units leaving the oxygen atom on the reducing end and the charge on the non-reducing end. A sequence which is derived from the C_n fragmentation sequence by loss of a water molecule (18 amu lower than the C_n sequence) was accompanied with mass number label and without sequence labeling (due to the limited space in the spectrum). The C_{na} sequence is probably due to an adduct of the analyte with the matrix triethanolamine followed by the loss of two water molecules from the C_n sequence. This type of sequence was also observed in the previous study of maltodextrin with the same kind of derivatives. The $^{2,4}A_n$ sequence is another prominent series, which is the most commonly observed series with charge retained on the nonreducing end. The proposed cleavage

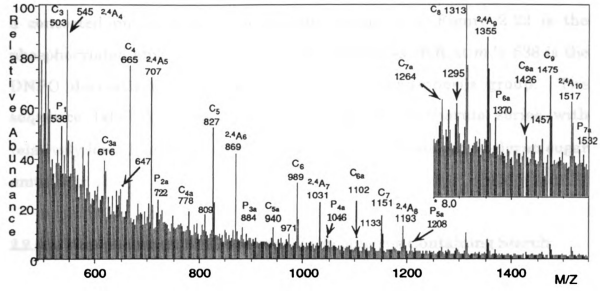


Figure 2.22 a: (-) FAB-MS of N,N-(2,4-dinitrophenyl)octylamine derivative of oligosaccharides digested from potato starch, range 500-1550.

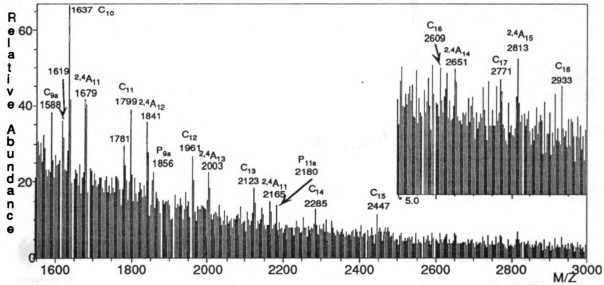


Figure 2.22 b: (-) FAB-MS of N,N-(2,4-dinitrophenyl)octylamine derivative of oligosaccharides digested from potato starch, range 1550-3000.

pattern (shown in Figure 2.3 b) involves a glycosidic bond rearrangement by a concerted mechanism. One specific sequence in Figure 2.22 is the phosphorylated glucan series. The peak with mass shift at m/z 538 is the DNPO derivatized terminus glucose with a phosphorus group. The sequence labeled P_{na} is a sodium phosphate saccharide series with relatively low abundance. The sequence can be followed up to eleven sugar units (m/z 2180).

2.2.5.6. Phosphorus-31 NMR Spectra of Phosphorus Containing Starch

To confirm the existence of the phosphate groups in the oligosaccharides, the acid-hydrolyzed starch was examined by phosphorus-

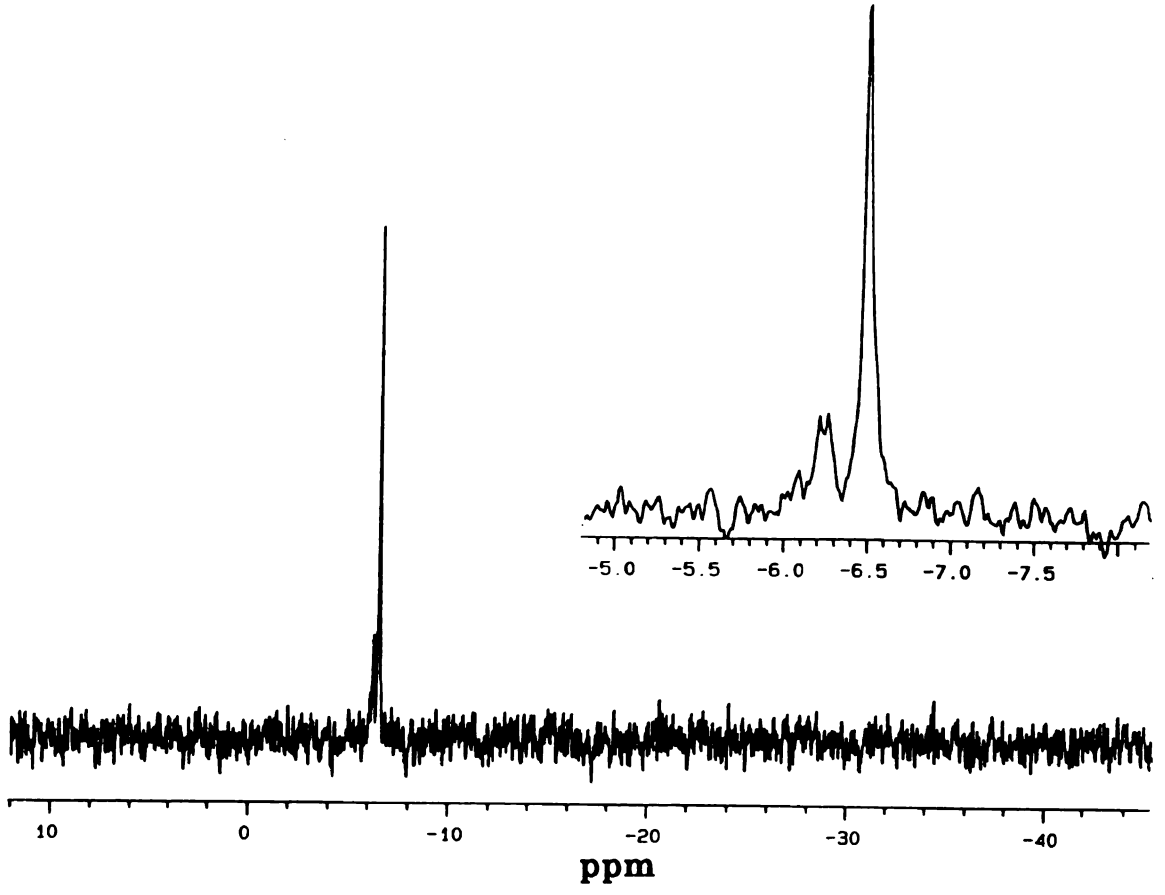


Figure 2.23: ^{31}P NMR spectrum of degraded phosphorus-containing starch (line broadening = 1 Hz)

31 NMR spectroscopy. The degraded starch was dissolved in D_2O, and hexamethylphosphoramide (HMPA) was used as reference. The phosphorus-31 NMR spectrum is shown in Figure 2.23. The line broadening was equal to 1 Hz. In the NMR spectrum, one peak at - 6.5 ppm is possibly due to phosphate group at C-6 position, the lower peaks at - 6.3 with splitting are possibly due to the phosphate groups at C-3 and C-2 positions.

The decoupled ^{31}P NMR spectra in pH 4 and pH 3 solutions showed no large variance from the one in neutral conditions. However in the non-decoupled spectrum (Figure 2.24), there are multiple splittings for the peak at - 6.3 ppm, which is possibly due to the overlap of doublet splitting to the

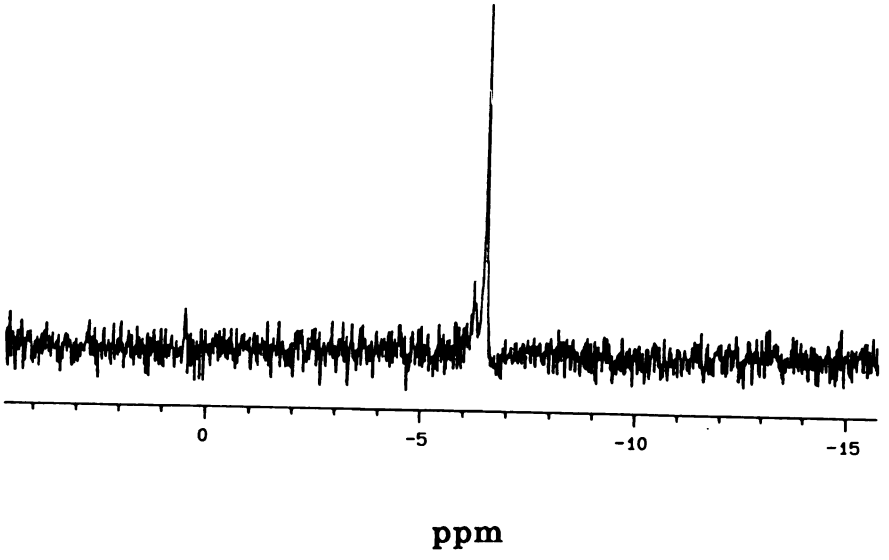


Figure 2.24: Non-decoupled phosphorus-31 NMR spectrum of degraded starch (pH = 4).

phosphate groups at C-2, and C-3 positions. After expanded the spectrum and setting the line brodening to 1 Hz (Figure 2.25). Material from digested starch still is a macromolecule, the rotation of phosphate groups is slow, the spectrum shows that the peak at -6.5 ppm is an unresolved multiplet probably because of differences in environments of the phosphate species to distributions rotamer populations. The result from NMR spectrum confirmed that starch contains phosphate groups, but the location of phosphate groups till needs further studies.

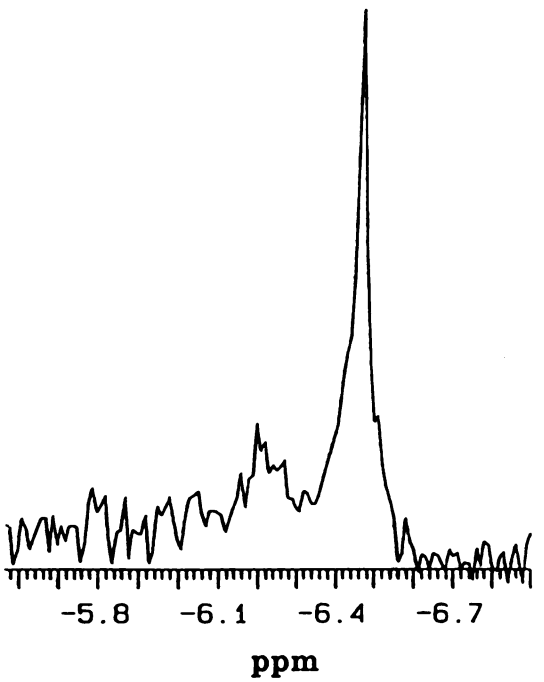


Figure 2.25: Expanded Non-decoupled phosphorus-31 NMR spectrum of degraded starch and setting line broadening to 1 Hz.

2.2.5.7. The Negative FAB-MS of Synthesized Maltohexaose-6-Phosphate

For further investigation of phosphorylated starches, we attempted to prepare the phosphate-containing oligosaccharides at very low levels of phosphorylation. Study of these synthesized phosphorylated

oligosaccharides and comparison of them to the natural product will greatly enhance our understanding of the structure and properties of the phosphorus-containing starch.

The N,N-(2,4-dintrophenyl)octylamine derivatized maltohexaose was treated with phosphorus oxychloride in dimethoxyethane (31). The negative FAB-MS obtained from 0.1 pmole sample is shown in Figure 2.26.

The C-6 primary hydroxyl group is the most active position in the pyranose. The synthesized phosphorylated maltohexaose could be the maltohexoase-6-phosphate. The negative FAB-MS indicates the presence of both monophosphorylated and nonderivatized maltohexaose. The low level derivatization could possibly be controlled by the shortness of the reaction time or the hydrolysis of phosphorus oxychloride. The ion at m/z 1269 corresponded to the DNPO derivatized maltohexaose. The ion at m/z 1330 is

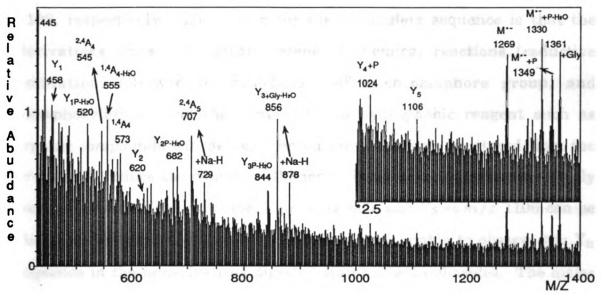


Figure 2.26: The negative FAB-MS of synthesized phosphorylated maltohexaose.

due to the loss of a water molecule form monophosphorylated maltohexaose. The peak at m/z 1349 arises from the monophosphorylated maltohexaose. The peak at m/z 1361 is due to the unphosphorylated material plus glycerol, the matrix compound. In the case of DNPO derivatized maltohexaose, the dinitrophenyl group stabilizes the negative charge, so the negative FAB-MS shows that the ion of the unphosphorylated molecule had the same molecular weight, 1269. Generally, in negative FAB-MS, the mass unit of the adduct from glycerol, which arises by loss of a proton from the glycerol molecule, should be 91 amu higher than $M^{-\cdot}$. In Figure 2.27, the glycerol adduct peak at m/z 1361 is 92 amu higher than m/z 1269 due to the electron capture of DNPO group. The phosphorylated peaks, Y_{np} ions with loss of a water molecule, appeared at m/z 520, 682, 844, and 1330, corresponding to the one, two, three and six sugar units. The phosphorylated Y_4 and unphosphorylated Y_5 ions appeared at m/z 1024 and 1106, respectively. The reason for the incomplete sequence is that the derivatives underwent several steps of chemical reactions (reductive amination, followed by attachment of a chromophore group, and phosphorylation), and that any salt and hydrophobic reagent such as organic compounds will depress the amount of sample sputtered from the probe and reduce the density of the ions. The Y_n sequence was relatively weak; only ions Y_1 at m/z 458, Y_2 at m/z 620, and Y_5 at m/z 1106 can be observed, which is not surprising since we were unable to observe the Y_n sequence in the underivatized digested starch oligosaccharides. The major sequence from the non-reducing end was the $^{2,4}A_n$ sequence of $^{2,4}A_4$ at m/z 545, and $^{2,4}A_5$ at m/z 707, similar to the $^{2,4}A_n$ sequence observed in the negative FAB-MS of derivatized oligosaccharides digested from potato starch (Figure 2.21). However, in Figure 2.21 there is better sensitivity and

completion of sequence. The spectrum from synthetic phosphorylated maltohexaose was just a preliminary result.

We also attempted to derivatize and phosphorylate the oligosacchrides digested from potato starch by the enzyme α-amylase without much success, even after prolonging the reaction to three days. One possible reason is that the starch was not very soluble in water, and so the reaction mixture was heterogeneous. Another explanation is that the enzyme we used was not active enough.

In conclusion, the starch appears to be heterogeneous in nature and contains phosphorus associated compounds. The origin and structure of this phosphate is still unknown. In negative FAB-MS, sequences can be observed both with and without phosphophate linkages, indicating that the starches of some strains contain phosphophate groups, while others do not. This also explained the previous reports in which 3% of phosphorus was found in hexasaccharide, isolated from potato starch (24). The phosphorus-31 NMR spectra are evidence indicating the satrch contained phosphorus. From the multiplicity of non-decoupling ^{31}P NMR, the phosphorus possibly located at C-6, C-3, and C-2 positions.

Because of the complex and delicate structure of starch granules and their derivatives, starch carbohydrate cannot be dissociated from the minor non-carbohydrate components that affect the properties of starches and products containing the starches. Realization of the important effect of non-starch components such as phosphorus on the properties of starch products will lead to a better understanding of starchy products and will aid in the development of better commercial methods for producing specific products.

2.2.6. The Negative FAB-MS of N,N-(2,4-DinitroPhenyl)octylamine Derivatized oligosaccharide isolated from Glycoprotein Fetuin

Glycoproteins are an important cell surface component. Determination of the structure of glycoproteins will have a considerable impact on our understanding of the biological function of the cell surface. In 1985, Naik et. al. (32) reported the determination of carbohydrate chains present in glycoproteins. Their method was based on the acetolysis of the intact glycocojugate, extraction of the peracetylated carbohydrate fragments, and analysis by FAB-MS. They reported the molecular ion of fully peracetylated species, and the Y_n sequence (which was called A-type oxonium ions). The highest molecular ion they reported, at m/z 1958, was a peracetylated species, possibly from hexopyranose.

For further study, the DNPO method was applied to structure elucidation of the carbohydrate moiety of the glycoprotein fetuin. Enzyme degradation was applied to the intact fetuin by Robert A. Cedergren of Dr. R. I. Hollingsworth's lab at MSU. After enzyme degradation, the sample was derivatized by reductive amination and derivatization, followed by gel filtration to purify the derivatized glycan. The fractions collected from the gel filtration column were tested by the phenol sulfuric calorimetric method. The derivatized oligosaccharides were analyzed by negative FAB-MS. The resultant (-) FAB-MS is shown in Figure 2.27. In Figure 2.27, there are many peaks between m/z 500-1000, some of them possibly due to the unpurified peptide residue. Some higher mass peaks were observed from m/z 1200-2800, which possibly arise from the derivatized oligosaccharides isolated from fetuin. Because of the limited amount of sample, further analyses were not performed. This preliminary result indicates that the DNPO method can be applied to detect higher molecular

weight oligosaccharides from a glycoprotein than the literature reported (32) peracetylation method in glycocojugate analysis.

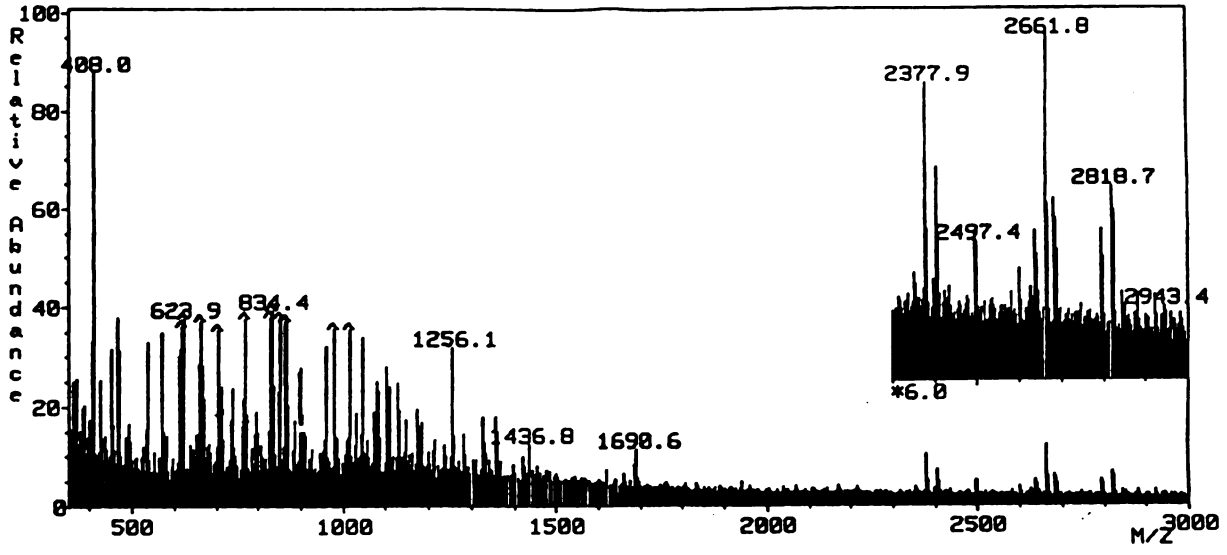


Figure 2.27: The negative FAB-MS of N,N-(2,4-dinitrophenyl)octylamine derivatized oligosaccharide isolated from fetuin.

2.2.7. N,N-Dansyloctylamine Derivatives for the Isolation, Purification, and Mass Spectrometric Characterization of Oligosaccharides

2.2.7.1. Preparation of N,N-Dansyloctylamine Derivatives of Oligosaccharies

The procedure for preparing the N,N-dansyloctylamine derivatives of oligosaccharides is similar to that of the N,N-(2,4-dinitrophenyl)octylamine derivatives and shown in Figure 2.28.

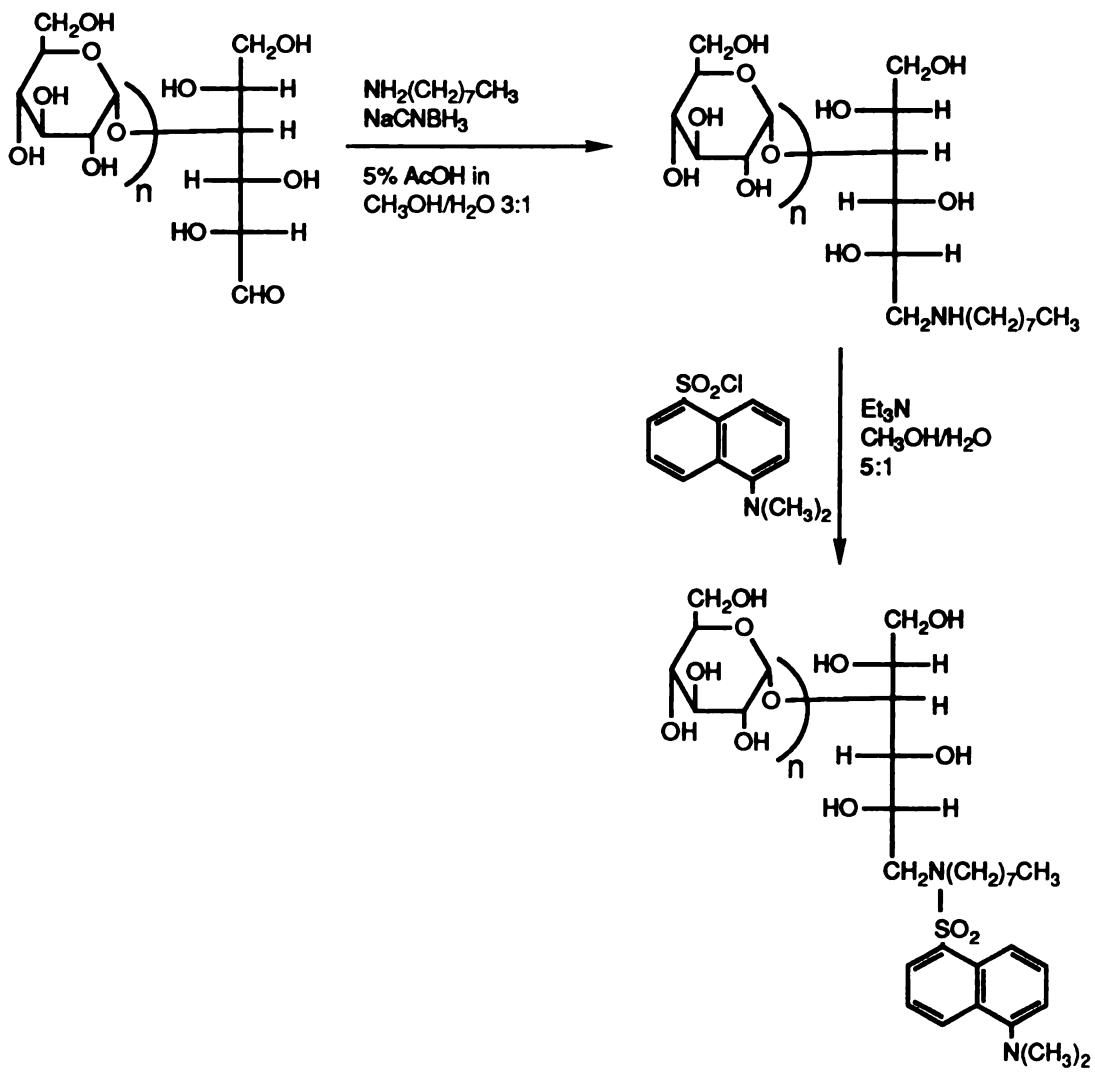


Figure 2.28: Preparation of N,N-dansyloctylamine derivatives.

The chromophore is a dansyl [5-(N,N-dimethyl)aminonaphthalene sulfonyl] group which aids UV detection during HPLC and also stabilizes positive ion species formed during analysis by fast atom bombardment mass spectrometry. The hydrophobic tail provided by an octyl group allowss the derivatized oligosaccharides amenable to be separated by reverse-phase chromatography using a C-18 stationary phase. This method was applied to analyze a mixture of starch maltodextrins of degree of polymerization 3-14, maltohexaose, and maltose.

2.2.7.2. Positive FAB-MS of N,N-Dansyloctylamine Derivatized Maltose

The (+) FAB-MS of N,N-dansyloctylamine derivatized maltose is shown in Figure 2.29. The ion at m/z 711 is the sodium salt adduct of N,N-dansyloctylamine derivatized maltose. For this spectrum, sodium hydroxide was used as the base to attach the dansyl group. Because the spectrum contained many sodium adduct peaks, an improved method was used with the triethylamine as base. The peak at m/z 362 is due to the

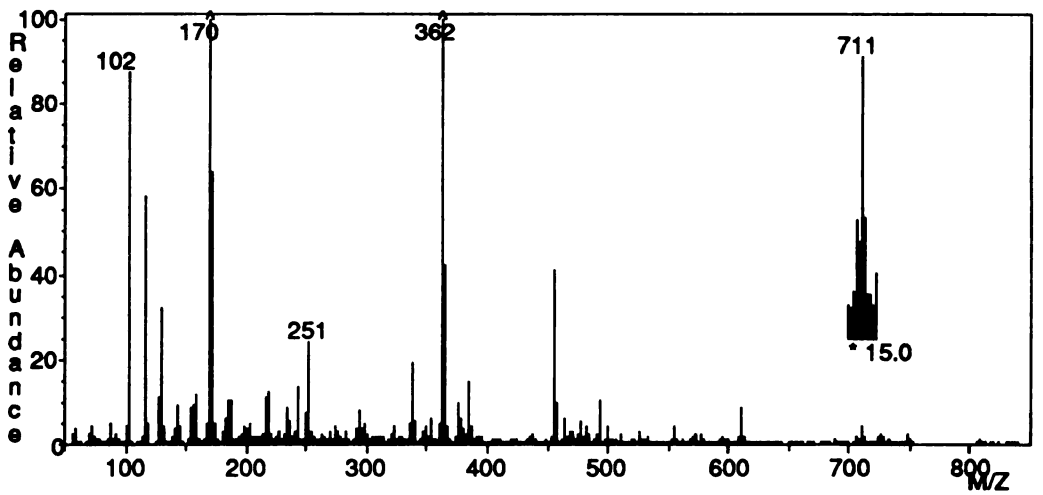


Figure 2.29: (+) FAB-MS of N,N-dansyloctylamine derivatized maltose.

derivatizing reagent. The peak at m/z 102 is due to the protonated thioglycerol matrix.

2.2.7.3. Positive FAB-MS of N,N-Dansyloctylamine Derivatized Maltohexaose

The (+) FAB-MS of N,N-dansyloctylamine derivatized maltohexaose is shown in Figure 2.30. For making the derivatives, 47 mg (~ 4.5 x 10^{-7} mole) of maltose was used as starting material and 0.1 pmol derivatized material was put on the probe. The peak at m/z 1337 is due to the protonated N,N-dansyloctylamine derivatized maltohexaose molecule. The next peak is the strong sodium adduct peak at m/z 1359. In the lower mass range, most peaks arise from the matrix and organic reagents. The peak at m/z 102 is protonated thioglycerol, m/z 363 and 454 are from protonated N,N-dansyloctylamine and its glycerol adduct, respectively. In contrast to Figure 2.2 of the negative FAB-MS from DNPO derivatized maltohexaose, the positive FAB-MS showed no sequence peaks at all. In preparing dansyl

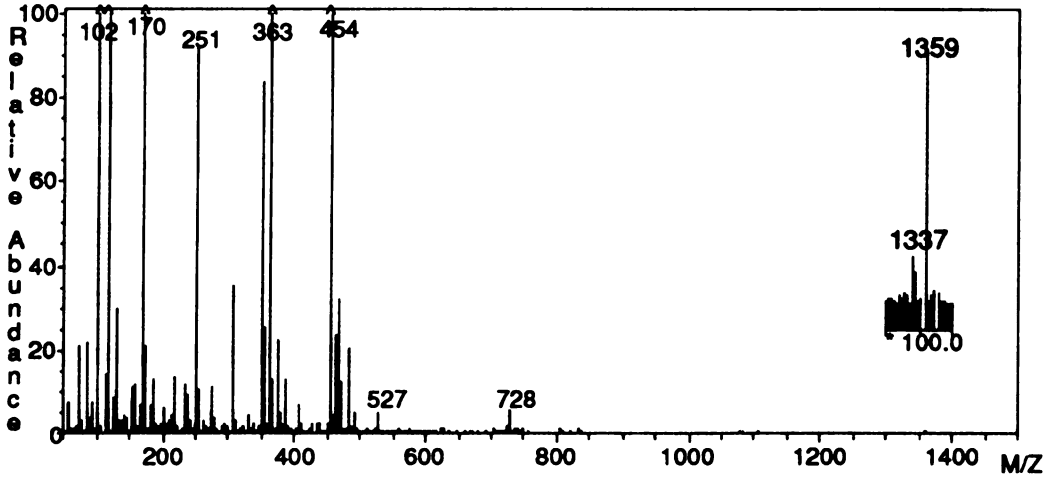


Figure 2.31: Positive FAB-MS of derivatized maltohexaose.

derivatives, sodium hydroxide was used as the base. Even after the sample was purified by a C-18 reversed phase cartridge, the trace amounts of alkali metal ions still decreased the fragmentation abundance.

2.2.7.4. Positive FAB-MS of N,N-dansyloctylamine Derivatized Maltodextrin

The (+) FAB-MS of N,N-dansyloctylamine derivatized maltodextrin is show in Figures 2.321a and b. The maltodextrin is a starch-derived mixture of oligomers. In the positive mode, a mixture of dextrin with degree of polymerization 1-15 can be detected. Components present at the level of 0.1 pmol could be detected. This method gave a good yield of cations for the derivatized maltodextrin. In Figure 2.31, the sequence ions observed are protonated Y_n ions with an adduct of matrix compound glycerol. Molecular ions from components of the mixture of oligomers which have the same ions as same the Y_n ions. In the lower range of 520-1800, the abundance of the Y_1 ion is as high as 100% and other Y_n ions are still dominant. The other sequence ions are $^{1,5}X_n$ ions from the protonated molecular species. They are low in abundance and disappear above approximately m/z 1600. There are sequence ions which arise from the main Y_n ions by loss of a water. These appear at m/z 600, 762, 924, 1086, 1248, 110, 1572, 1734 without specific labeling. In another sequence, the ions are a little higher than the main ions. This sequence is due to the sodium salt adduct to the main Y_n ions and the fragments appear at m/z 640, 802, 964, 1126, 1288, 1450, 1612, 1774, 1936, 2098, 2260, 2422, 2584, 2746 and 2908. A few ions appeared at m/z 647, 809, 971, corresponding to the $^{1,5}X_n$ ions plus the matrix compound glycerol. The mixture of derivatized oligomers could easily be separated by HPLC. Based on the relative proportions of the

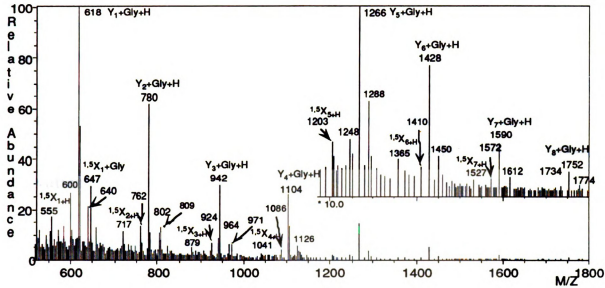


Figure 2.31 a: (+) FAB-MS of N,N-dansyloctylamine derivatized maltodextrin.

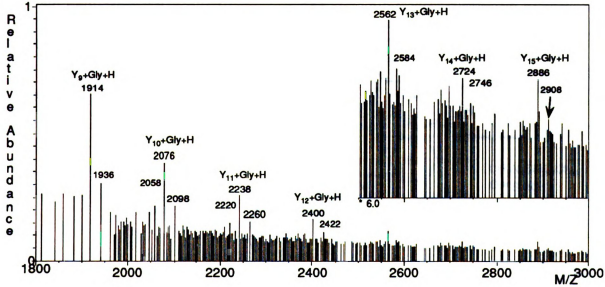


Figure 2.31 b: (+) FAB-MS of N,N-dansyloctylamine derivatized maltodextrin.

individual oligomers calculated from this analysis, the highest oligomers were present in amounts of about 0.1 pmol; however, they could easily be detected in mass spectra of the entire mixture. This represents an improvement in sensitivity of between 100- and 1000-fold compared to the aminobenzoic acid alkyl ester method (16). In our analyses, single scans were taken and no attempt was made to improve signal-to-noise by signal averaging. In some cases, there was no sample work-up after derivatization.

2.2.7.5. FAB-MS of N,N-dansyloctylamine Derivatized Oligosaccharide from the Bacteria Surface

As a further test of the efficiency of this derivatizing procedure for enhancing signal intensity for positive FAB-MS, a previously characterized oligosaccharide from *Rhizobium trifolii* 843 capsular polysaccharide (1),

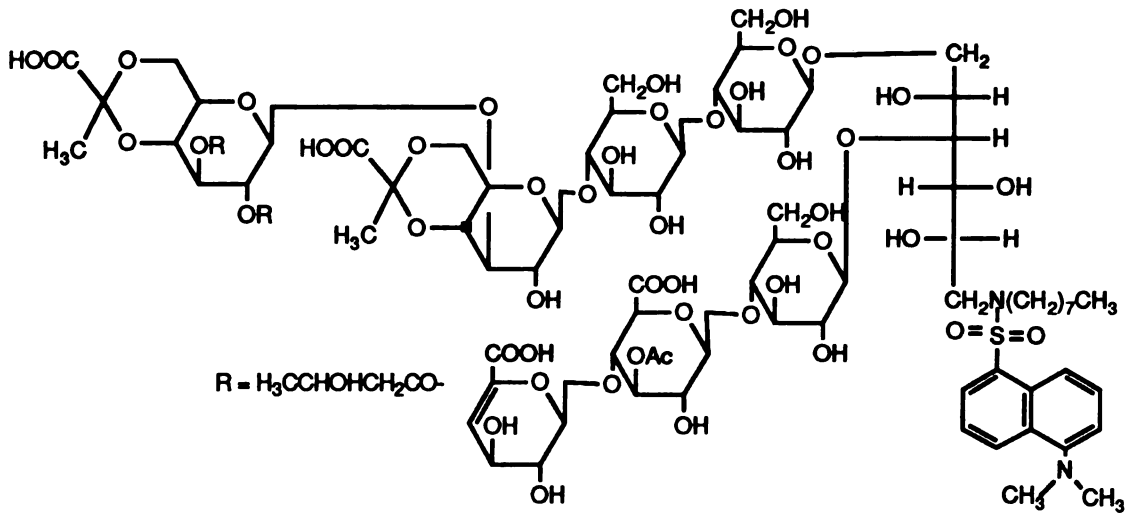


Figure 2.32: N,N-dansyloctylamine derivatized oligosaccharide from a bacterial cell surface.

was derivatized and analyzed. The structure of the N,N-dansyloctylamine derivatized oligosaccharide is shown in Figure 2.32.

The reaction was performed on 64 μg (~ 3.8 x 10^{-7} mole) of isolated pure oligosaccharide, and about 3 pmol of derivatized material was put on the probe. The positive FAB-MS of N,N-dansyloctylamine derivatized oligosaccharide from the bacterial surface is shown in Figure 2.33.

Since base treatment was involved in the derivatization procedure, loss of the acetyl and 3-hydroxybutyryl groups in the derivatized molecule was expected. There are four carboxyl groups in the derivatized

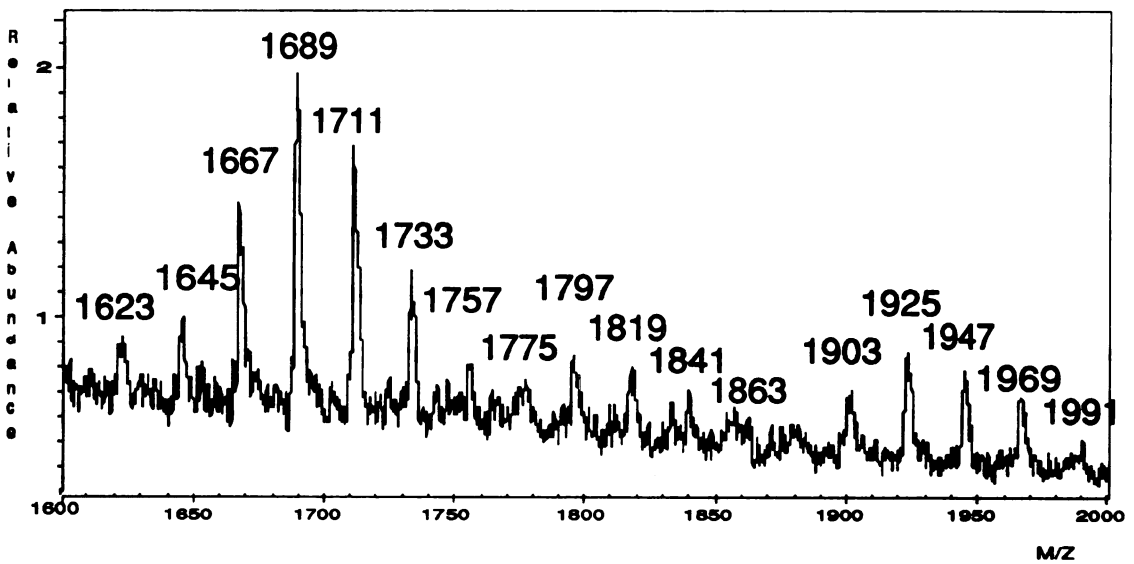


Figure 2.33: (+) FAB-MS of N,N-dansyloctylamine derivatized oligosaccharide from the bacterial cell surface.

oligosaccharide; therefore, clusters of sodium salt adducts were observed in the spectrum. The peak at m/z 1903 arises from the derivatized molecule (molecular weight 1810) plus a matrix compound glycerol (93 amu). The clusters at m/z 1925, 1947 1969 and 1991 arise from sodium adducts of the

four carboxyl groups on the compound. The peak at m/z 1775 arises from the protonayed derivatized molecule with loss of two water molecules, and the clusters at m/z 1797, 1819, 1841 and 1863 are sodium adducts to the four carboxyl groups. The ion at 1757 is loss of a water molecule from m/z 1775. The ion at m/z 1623 arises from the protonated derivatized molecule loss of a terminal unsaturated hexaose (159 amu) and an aldehyde group. The clusters at m/z 1645, 1667, and 1689 are the sodium adduct to the three carboxyl groups, the ions at m/z 1711 and 1733 are sodium adducts on the molecule.

2.2.8. Conclusion

The method described here provides an excellent yield of derivatized material, without underivatization or any side reaction being observed in the spectrum, in a very efficient fashion with a minimum of sample preparation and purification before analysis. The molecular weights of the derivatized compounds are much lower than those obtained from the peralkylation method. This method also has a minimal effect on the sensitivity. Useful structural information was obtained from the ion series observed, without using MS/MS conditions. The sensitivity obtained using this derivative is 100 fold superior to that reported from other methods, and the derivatives are "well-behaved" during HPLC analysis; the peak shapes on HPLC are good with no tailing, and separation can be effected using very simple solvent systems. The sensitivity in FAB-MS analysis is also very good, the limit of detection is as low as 0.01 pmol. The spectra obtained from large oligosaccharide have relatively high ratio of signal to noise. One reason for this is due to the direct analysis of the sample without purification, another reason is that the spectra was done in a single scan

without peak average, also large molecules have lower effect from the derivative, the surface activity becomes lower, and the appearance of the sputtered ionization occurs with more difficulty.

To distinguish the spectra of isomers, comparision of the spectra of maltohexaose and isomaltohexsaose, Y_n sequence was performed. Four sequences were observed from maltohexaose and two sequences were present in the spectrum of isomaltohexaose. The $^{1,4}A_n$ from maltohexaose is impossible to obtain in the isomaltohexaose, because after loss of the $^{1,4}A_n$ species, the isomaltohexaose is still in the aggregated form, like loss of functional groups, which does not release enough energy from the molecule. So the type of fragmentation sequences obtained from the molecules provide information about sequence and some hint of linkages.

In the case where there are monosaccharides with different ring size in the polysaccharide, such as furanose and pyranose rings, we can predict the rings with different size are going to give different kind of fragments, even though we have no experince with this case. Also both pyranose and furanose forms will give the same Y_n sequence.

In the case of large oligosaccharides, such as maltodextrin and starch hydrolyzed oligosaccharides, the most popular sequences observed are C_n sequences. The C_n sequence is formed from the bond cleavage between the oxygen and C_4. The reason we obtain C_n sequence only in large oligosaccharides is possibly due to the fact that this fragmentation sequence can release more energy from the molecules.

In most cases, the desorption techniques such as FAB-MS, plasma desorption, and field desorption, generally provide only molecular mass information for many large nonvolatile compounds, revealing only minor structural information from specific fragmentation. The sequence ions

obtained in this method are very useful for studying molecular structure, ion structure, and ion fragmentation patterns, which are used to predict and interpret fragmentation of unknown oligosaccharides, glycosides, and also can be extended to more complex hetero-oligosaccharides, especially those from glycoproteins isolated from low-abundance biological samples.

This method is that it can give very good sequence information, and some clue for the linkages, but the weakness is to obtain the real linkage information, methylation analysis is still required. Also the configuration of anomeric proton can not be obtained from this method, but it is easily obtained from NMR spectrometry.

2.3 References

1. Hollingsworth, R. I., and Dazzo, F. B., *Carbohydr. Res.*, (1988) 172 97-112.
2. McIntire, F. C., Clements, L. M. and Sproull, M., *Anal. Chem.* (1953) 25 1757-1758.
3. Coates, M. L., Wilkins, C. L., *Biomed. Mass Spectrom.* (1985) 12 424-428.
4. Coates, M. L., Wilkins, C. L., *Anal. Chem.* (1987) 59 197-200.
5. Takayama, K., Qureshi, N., Hyver, K., Honovich, J., Cotter, R. J., Mascagni, P., Schneider, H. J., *J. Biol. Chem.* (1986) 261 10624-10631.
6. Cotter, R. J., Honovich, J., Qureshi, N., Takayama, K., *Biomed. Environ. Mass Spectrom.* (1987) 14 591-598.
7. Domon, B., and Costello, C., *J. Glycocojugate,* (1988) 5 397-409.
8. Hendrickson, J. B., *Angew Chem. Int. Ed. Engl.* (1974) 13 47-76.
9. Poulter, L. J., Earnest, J. P., Strond, R. M., and Burlingame, A. L. *Biomed. Environ. Mass Spectrom,* (1988) 16 25-30.

10. Ganguly, A. K., Capucino, N. F., Fujiwara, H., and Rose, A. K., *J. Chem. Soc. Chem. Comm.*, (**1979**) 148-149.

11. Prome, D., Prome, J.-C., Puzo, G., and Aurelle, H., *Carbohydr. Res.* (**1985**) 140 121-129.

12. Biemann, K., *"Biological Mass Spectrom."* Burlingame, A. L., and McCloskey, J. A., (Eds), Elsevier Science Publishers B.V., Amsterdam.

13. Spengler, B., Dolce, J. W., and Cotter, J. R., *Anal. Chem.* (**1990**) 62, 1731-1737.

14. Sasaki, H., Norinichi, O., Dell, A., and Fukuda, M., *Biochemistry* (**1988**) 27 8618-8626.

15. Williams, J. M., *"Starch and its Derivatives"* (**1968**) 91-138, (Radley, J. A., ed.), 4th ed. Chapman & Hall, London.

16. Lee, E. Y. C., Mercier, C., and Whelan, W. J., *Arch. Biochem. Biophys.* (**1968**) 125 1028-1030.

17. Morrison, W. R., and Karkalas, J., *"Methods In Plant Biochemistry"* (**1990**) 2 323-352.

18. Manners, D. J., *Essays Biochem.* (**1974**) 10 37-71.

19. Manners, D. J., and Matheson, N. K., *Carbohydr. Res.* (**1981**) 90 99-110.

20. Callaghan, P. T., and Lelievre, J., *Carbohydr. Res.* (**1987**) 62 33-40.

21. Takeda, Y., and Hizukuri, S., *Carbohydr. Res.* (**1987**) 168 79-88.

22. Manners D. J. *"Biochemistry of storage carbohydrates in green plants"* (**1985**) 154, Academic Press, London.

23. Kerr, R. W., and Severson, G. M., *J. Am. Chem. Soc.*, (**1943**) 65, 193.

24. Posternak, T., *Helv. Chim. Acta.*, (**1935**) 18, 1351.

25. Parrish, F. W., and Whelan, W. J., *Staerke*, (**1961**) 13 231.

26. Postnak, T., *J. Biol. Chem.*, (**1951**) 188 317.

27. Hizukuri, S., Tabata, S., Kagoshima, and Nikuni Z., *Die Starke* (**1970**) 10 338-343.

28. Tabata, S., and Hizukuri, S. *Die Starke* (**1971**) 8 267-272.

29. Hizukuri, S., Takeda, Y., and Yasuda, M. *Carbohydr. Res.* (**1981**) 94 205-213.

30. Westler, R. L., *"Methods in Carbohydrate Chemistry"* (**1964**) 4 Westler, R. L., ed., Academic Press, New York.

31. Khorana, H. G., *"Some Recent Developments in the Chemistry of Phosphate Esters of Biological Interest"* John Wiley & Sons, Inc., New York · London, (1961) 13-43.

32. Naik, S., Oates, J. E., Dell, A., Talor, G. W., Dey, P. M., and Pridham, J. B., *Biochem. Biophys. Res. Comm.* (**1985**) 132 1-7.

CHAPTER 3

EXPERIMENTAL FOR PART I

3.1 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Maltose

Nine mg (2.5×10^{-5} mole) of maltose, 0.1 mL water, 3 equivalents octylamine, and 4 equivalents sodium cyanoborohydride were placed in a Teflon-lined, screw cap, one dram vial. Two-tenths mL 5% acetic acid in a solution of water/methanol (3:1) was then added; the vial was capped and heated at 70° C for 1 hour, then blown to dryness. Afterwards, 0.1 mL of 0.2*M* sodium bicarbonate and 2 equivalents of 2,4-dinitrofluorobenzene in 2 mL methanol were added (1). The mixture was heated for 25 minutes at 60°C, and subsequently 0.4 mL of 0.2 *M* sodium hydroxide in dioxane containing trace amounts of water were added to dissolve the base. The mixture was then heated at 60° C for 1 hour. After cooling to room temperature, 5 mL water and 5 mL chloroform were added. The contents of the aqueous layer were adsorbed on a C18 reversed phase cartridge column. The column was eluted with 1 mL water, 1 mL water/methanol (1:1), and 1 mL methanol. All three fractions of elutant were analyzed by NMR and FAB-MS. Most of the carbohydrate was contained in the water fraction.

3.2 UV Spectrum of N,N-(2,4-dinitrophenyl)octylamine Derivatized Maltose

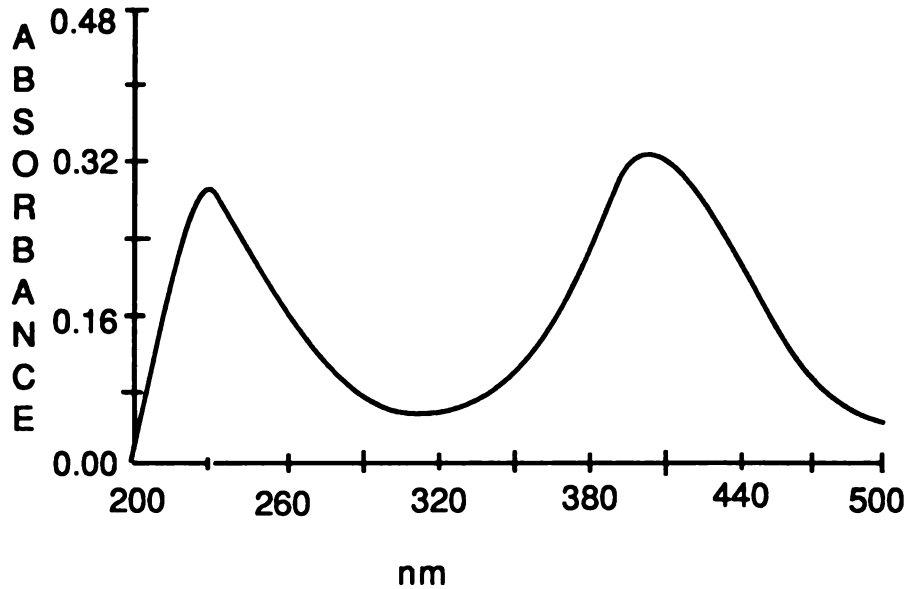


Figure 3.1: UV spectrum of N,N-(2,4-dinitrophenyl)octylamine derivatized maltose.

3.3 Preparation of N,N-(2,4-dinitrophenyl))octylamine Derivatized Maltohexaose

Maltohexaose (650 mg, ~ 6.3 x 10^{-7} mole) was derivatized as in section 3.1.

3.4 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Maltodextrin

Maltodextrin of 17 mg (~ 9.6 x 10^{-6} mole) was derivatized as in section 3.1.

3.5 UV Spectrum of N,N-(2,4-dinitrophenyl)octylamine Derivatized Maltodextrin

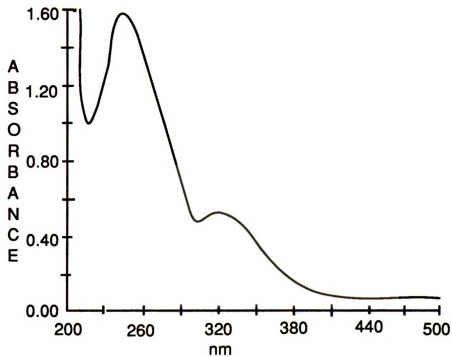


Figure 3.2: UV spectrum of N,N-(2,4-dinitrophenyl)octylamine derivatized maltodextrin

3.6 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Oligosaccharide Separated From the Bacteria Surface

The 850 mg (~ 5.1 x 10^{-7} mole) of oligosaccharide from the bacteria surface was derivatized as in the section 3.1.

3.7 Preparation of N,N-Densyloctylamine Derivatized Maltose

A sample of maltose, 16 mg (~ 4.8 x 10^{-5} mole), ~3 equivalents octylamine and 4 equivalents sodium cyanoborohydride were added to a one dram vial, after which 0.2 mL 5% acetic acid in a solution of water/methanol (3:1) was added. The Teflon-lined vial was capped and

heated at 70° C for 1 hour. Subsequently, the solution was reduced to dryness. Saturated sodium bicarbonate 0.2 mL, 3 equivalents dansyl chloride (5-N,N-dimethyl naphthalene sulfonic chloride), and 0.2 mL methanol/water (4:1) were added consecutively (2), and the mixture heated at 45° C overnight. Samples were analyzed directly without purification.

3.8 Preparation of N,N-Densyloctylamine Derivatized Maltodextrin

The 93 mg (~ 6.3 x 10^{-5} mole) of maltose was derivatized as in the section 3.7.

3.9 Preparation of N,N-Densyloctylamine Derivatized Oligosaccharide Separated from Bacteria Surface

The 640 mg (~ 3.8 x 10^{-7} mole) of maltose was derivatized as in the section 3.7.

3.10 Preparation of N,N-Densyloctylamine Derivatized Maltohexaose (Method II)

The 770 μg (~ 7.4 x 10^{-7} mole) maltose, ~ 5 equivalents octylamine and 4 equivalents sodium cyanoborohydride were added to a 1 dram vial; then 0.1 mL 5% acetic acid in a solution of water/methanol (3:1) was added, the Teflon-lined vial was capped and heated at 70° C for 1 hour. Subsequently, the solution was blown to dryness. Triethanolamine (0.1 μL), 2 equivalents dansyl chloride (5-N,N-dimethyl naphthalene sulfonic chloride), and 0.2 mL methanol/water (4:1) were added consecutively, and the mixture heated at 45° C overnight. Samples were analyzed directly without purification.

3.11 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Maltohexaose (Method I)

The 730 μg (~ 7 x 10^{-7} mole) of maltohexaose, 5 equivalents octylamine, and 6 equivalents sodium cyanoborohydride were placed in a one dram vial. Two-tenths mL 5% acetic acid in a solution of water/methanol (3:1) was then added. The vial was capped and heated at 70° C for one hour, then blown to dryness. Afterwards, 1 mL of triethylamine and 2 equivalents of 2,4-dinitrofluorobenzene in 2 mL methanol, 0.2 mL water were added. The mixture was then heated at 45° C overnight. After cooling to room temperature, 5 mL water and 5 mL ethyl acetate were added. The contents of the aqueous layer were analyzed directly by FAB-MS.

3.12 Preparation of N,N-Densyloctylamine Derivatized Maltohexaose (Method I)

The 470 μg (~ 4.5 x 10^{-7} mole) of maltose, ~5 equivalents octylamine and 4 equivalents sodium cyanoborohydride were placed in a one dram vial; and 0.1 mL 5% acetic acid in a solution of water/methanol (3:1) was added, the Teflon-lined vial was capped and heated at 70° C for one hour. Subsequently, the solution was blown to dryness. Triethylamine (0.5 μL), 0.5 μL 2,4-dinitrofluorobenzene, and 0.2 mL methanol/water (4:1) were added consecutively, and the mixture heated at 45° C overnight. Samples were analyzed directly without purification.

3.13 Preparation of 6-Phosphate Maltohexoase

One-tenth mL of 0.014 *M* N,N-(2,4-dinitrophenyl)octylamine derivatized maltohexaose was placed in a one dram vial and blown to

dryness to remove excess solvent; the vial was kept in an ice bath, and 0.1 mL dimethoxyethane (DME), and 1 μL $POCl_3$ were added. The reaction was kept in an ice bath for one hour, then 0.02 mL water was added. The sample was analyzed directly by FAB-MS, without purification.

3.14 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized N,N',N"-Triacetylchitotriose

The 0.0027 μ*M* aqueous solution of N,N'N"-triacetylchitotriose (40 μL, ~ 1.1 x 10^{-7} mole) was placed in a one dram vial, followed by addition of 25 μL of 24 μ*M* octylamine in 5% acetic acid in water/methanol (3:1) solution and addition of 25 μL of 4 μ*M* sodium cyanoborohydride in 5% acetic acid water/methanol (3:1) solution. The mixture was heated at 40° C overnight. After cooling to room temperature, 25 μL of 56 μ*M* triethylamine in methanol and 25 μL of 64 μ*M* 2,4-dintrofluorobenzene were added. The mixture was heated at 45° C for two days and analyzed by FAB-MS without further purification.

3.15 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Lacto-N-Fucopentoase

The 7.8 x 10^{-4} μ*M* aqueous solution of lacto-N-fucopentoase (40 μL, ~ 1.1 x 10^{-7} mole) was derivatized as in section 3.14.

3.16 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Lacto-N-Difucohexaose

The 8.2 x 10^{-4} μ*M* aqueous solution of lacto-N-difucohexaose (40 μL, ~ 1.1 x 10^{-7} mole) was derivatized as in section 3.14.

3.17 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized Isomaltohexaose

The 8.2 x 10^{-4} μ*M* aqueous solution of isomaltohexaose (40 μL, ~ 1.1 x 10^{-7} mole) was derivatized as in section 3.14.

3.18 Enzyme Digestion of Potato Starch I

Potato starch 54 mg, and 0.7 mL (0.15 unit/mL) of α-amylase were placed in a one dram vial. The α-amylase was suspended in a 2.9 *M* NaCl solution containing 3 m*M* $CaCl_2$. Acetic acid (0.7 mL of 0.02 μ*M*) buffered with NaOAc (pH 5) was added. The sample was heated at 100° C for 60 minutes and quenched with 5% acetic acid solution. The digested starch was derivatized by N,N-(2,4-dinitrophenyloctyl)amine and analyzed by FAB-MS. The spectrum was not as good as that of maltodextrin, possibly because the starch is not sufficiently water-soluble, and the reaction was not complete.

3.19 Acid Hydrolysis of Starch

The 72 mg of starch was placed in a one dram vial, followed by 100 μL of 0.1 *M* trifluoroacetic acid. The sample was heated at 75° C for 45 minutes, then N,N-(2,4-dinitrophenyl)octylamine was derivatized as previously described. The sample was purified by mixed bed resin, and sequence FAB-MS spectrum was obtained.

3.20 Enzyme Digestion of Potato Starch II

Potato starch 100 mg and 2 mL DMSO were placed in a one dram vial. The sample was placed at room temperature for 2 days, then 20 μL of α-amylase (5 units/μL in 50 *mM* citrate, pH 4.6) was added. The mixture

was heated at 45° C and completion of the reaction was indicated by iodine solution (I_2 0.002%, KI 0.2 % in 0.2 *M* acetate buffer, pH 4.8). After approximately 20 hours, the color of the iodine solution became pale, the sample was blown to dryness, and derivatized by N,N-(2,4-dinitrophenyl)octylamine using the same procedure described earlier.

3.21 Preparation of N,N-(2,4-dinitrophenyl)octylamine Derivatized oligosaccharide isolated from glycoprotein Fetuin

Two-tenth mL enzyme digested fetuin was placed in a one dra vial. The sample was lyophilized to dryness and derivatized as in section 3.5.

3.22 Mass Spectrometry

FAB-MS was performed on a JEOL HX110 (Peabody, MA) double-focusing mass spectrometer (EB configuration) equipped with a high field magnet operated in the negative ion mode. Ions were produced by bombardment with a beam of Xe atoms (6 kV). The accelerating voltage was 10 kV, and the resolution was set at 1000 or 3000 according to the mass range of interest. The samples were dissolved in a 1:1 (v/v) H_2O-MeOH solution; generally 1-1.5 μL sample was mixed with 1 μL of glycerol/thioglycerol (1:1) or triethanolamine on the FAB-MS stainless steel probe tip. The solvent was pumped away by the mass spectrometer before analysis. Calibration was performed using Ultramark (443 and 6121) or $(CsI)_nI^-$ cluster ions, depending on the mass range of interest. A JEOL DA-5000 data system recorded the spectra. The sample was scanned at a scan slope of 2 min from m/z 0-3000. Data presented were acquired in a single scan and found to be reproducible.

3.3 References

1. McIntire, F. C., Clements, L. M. and Sproull, M (**1953**) *Anal. Chem.* 25, 1757-1758

2. Renner, D., and Spiteller, G., *Angew. Chem. Int. Ed. Engl.* (**1985**) 24 408-409.

PART II: STRUCTURE ELUCIDATION AND SYNTHETIC TRANSFORMATIONS RELATED TO THE INNER AND OUTER CORE OF THE LIPOPOLYSACCHARIDES OF *RHIZOBIUM MELILOTI* 41 AND *LEGUMINOSARUM* BIOVAR *VICIAE* VF-39 MUTANTS

ABSTRACT

PART II: STRUCTURE ELUCIDATION AND SYNTHETIC TRANSFORMATION RELATED TO THE INNER AND OUTER CORE IN THE LIPOPOLYSACCHARIDES OF *RHIZOBIUM MELILOTI* 41

by

Yuanda Zhang

The structure of the inner and outer core of the LPS of *Rhizobium meliloti 41* has been analyzed by a combination of chemical and spectroscopic methods. An unusual sugar component has been studied. The structure was investigated by J-correlated, ^{1}H-NMR spectroscopy. The composition and structure of the carbohydrate chain was analyzed by methylation analysis using GC-MS. Conformation of the structure (but not configuration) of the discovered compound was attempted by synthesis.

The LPS of a wild type strain of *Rhizobium leguminosarum* biovar *viciae* (strain VF-39) and two symbiotically defective Tn5 mutants (VF-39-32 and VF-39-86) have been studied. The LPS of the mutants reflected impaired synthesis of the *O*-antigen. This study clearly defines a role for the bacterial LPS in the proper functioning of the *Rhizobium* legume symbiosis.

A trisaccharide containing an α-linked *O*-mannopyranosyl residue on the 3-hydroxymethyl group and an α-linked *O*-galactopyranosyl residue at the 5 position of a new branched pentafuranose sugar has been isolated from strain VF 39-32, a symbiotically defective strain of *Rhizobium leguminosarum* biovar *viciae* impaired in normal lipopolysaccharide

synthesis. The branched glycosyl residue has been identified as a 2-deoxy-3-C-(hydroxymethyl)pentofuranose and appears to have the threo-configuration. The structure of the new glycosyl component was determined by ^{1}H-NMR spectroscopy using spin decoupling to determine proton connectivities and by mass spectrometry. The positions of substitution were deduced by proton n.O.e. studies on the native oligomer and the configuration determined by an analysis of coupling constants, n.O.e., and spin decoupling data. The ring size of the branched chain glycosyl component was deduced from coupling constants. Signals for small quantities of this unusual component can be discerned in ^{1}H-NMR spectra of lipopolysaccharide fragments from the parent wild-type organism, suggesting that its synthesis is repressed by that of the normal lipopolysaccharide and significantly repressed in the mutant strain. Traces of a trisaccharide which is likely to be a truncated version of a tetrasaccharide, normally synthesized by the parent strain, containing galactose, galacturonic acid, mannose and 3-deoxy-2-octulosonic acid, were also detected.

CHAPTER 4

STRUCTURE ELUCIDATION AND SYNTHETIC TRANSFORMATION RELATED TO THE THE INNER AND OUTER CORE IN THE LIPOPOLYSACCHARIDES OF *RHIZOBIUM MELILOTI* 41

4.1 Introduction

Carbohydrate molecules found naturally or obtained by synthesis are playing a very important role in current chemistry, biochemistry, microbiology, and other areas. The project described in part II has two objectives: the structure elucidation and synthesis of key components of the lipopolysaccharide (LPS) of the nitrogen-fixing bacterium *Rhizobium meliloti* 41. This organism effects the reduction of dinitrogen gas into ammonia, an essential step in providing plant protein and nutrition.

The interest in this research is the chemical structure of the inner and outer core lipopolysaccharide from *Rhizobium meliloti 41 LPS* mutant. *Rhizobium* is a species of Gram-negative bacteria which has a nitrogen-fixing symbiotic relationship with legume plants (1). Like most Gram-negative bacteria, *Rhizobium* has usual surface polysaccharides consisting of lipopolysaccharides (LPSs), extracellular (EPSs) and capsular polysaccharides (CPSs). The major surface polysaccharides of these bacteria are lipopolysaccharides. The symbiosis is characterized by a high degree of specificity with respect to host-range. The bacteria from a given species will selectively infect only one or two host species. In order to assess the role of *Rhizobium* lipopolysaccharides in the specific recognition between the symbiont bacterium and the host cell, we propose to characterize these molecules structurally.

Lipopolysaccharides are macromolecules which determine the surface chemistry of Gram-negative bacteria. Lipopolysaccharides are located in the cell envelope (Figure 4.1) (2). They are composed of three main structural

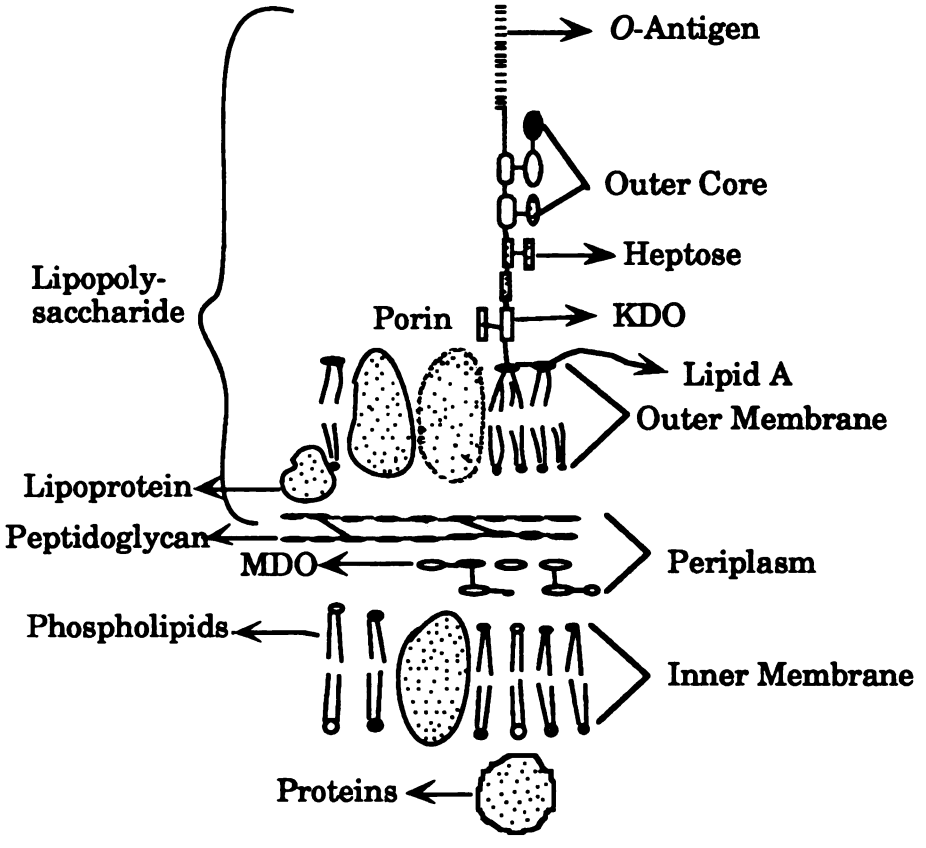
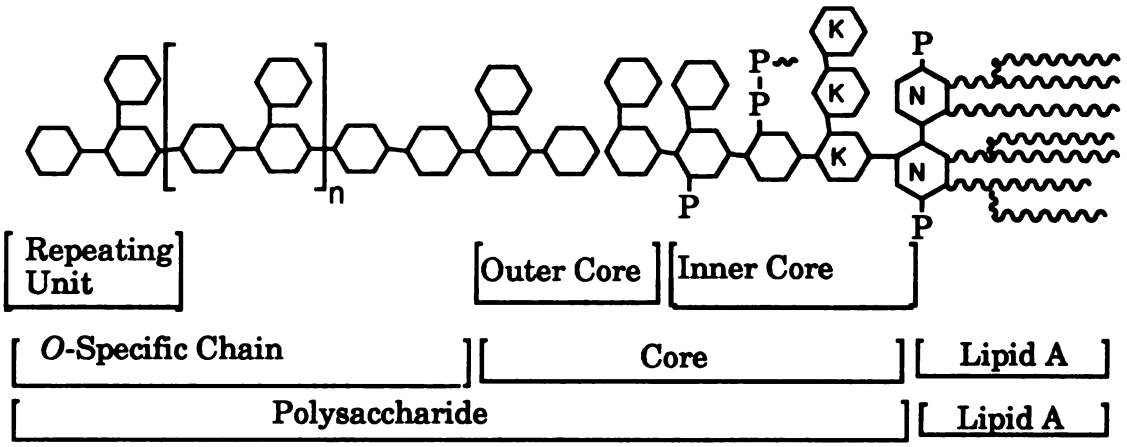


Figure 4.1: Schematic molecular representation of the *E. coli* envelope. Ovals and rectangles depict sugar residues. Circles represent the polar head groups of phospholipids. MDO are membrane-derived oligosaccharide, and KDO is 3-deoxy-D-manno-octulosonic acid. KDO and heptose make up the inner core of LPS. The figure was taken from Christian R. H. Raetz (2) with modifications.

regions. The first region of LPS is the *O*-antigen, which is usually a repeating sequence of carbohydrate molecules which determines the serology of the organism (Figure 4.2) (3). The *O*-antigen is, typically, highly variable.

The *O*-antigen is attached to a second carbohydrate region, which is much smaller and not a repeating sequence, called the R-core. This core region contains the LPS, marker sugar called KDO, and is usually conserved within a given genus. It might be expected that the core structures of *Rhizobium trifolii* and *Rhizobium meliloti* might be the same. The third region of the LPS called lipid A, is made up of a carbohydrate group (usually a disaccharide) to which a variety of 12 to 16 carbon 3-hydroxyalkanic acids are attached. The lipopolysaccharide molecule is anchored to the outer membrane of the bacterium via the fatty acids of the lipid A region.

It has been reported that the lipopolysaccharide of *Rhizobium trifolii*



Sugar residue, l-Glycero-D-manno-heptose,
K 2-Keto-3-deoxy-D-manno-octonate, N D-Glucosamine, P Phosphate,
Ethanolamine, 3-Hydroxyacyl group,
3-Acyloxyacyl group

Figure 4.2: Schematic structure of a *Salmonella* lipopolysaccharide. The figure was taken from the review of Westphal *et al.* (3) with modifications.

ANU843 (a strain of a related species) contains a trisaccharide (Figure 4.3) (4) and tetrasaccharide (Figure 4.4) (5) as the components in the core

oligosaccharide. According to the above reports, both the *O*-antigen and the core oligosaccharides contain 3-deoxy-D-manno-2-octulosonic acid (KDO) at their reducing end, and are linked to the rest of the LPS molecule via KDO.

In this study, we attempt to characterize the structure of the core oligosaccharide of the LPS of *Rhizobium meliloti* and describe efforts to determine the structure of one unusual glycosyl component. This work will also allow us to relate host range to LPS structure by studying organisms from two different species.

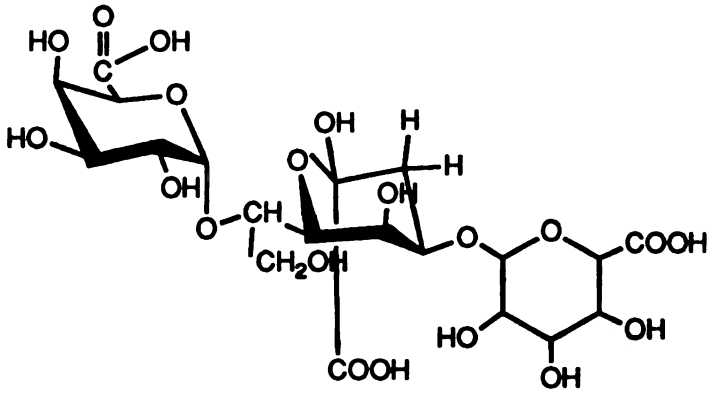


Figure 4.3: Structure of trisaccharide from *Rhizobium trifolii ANU843* (4).

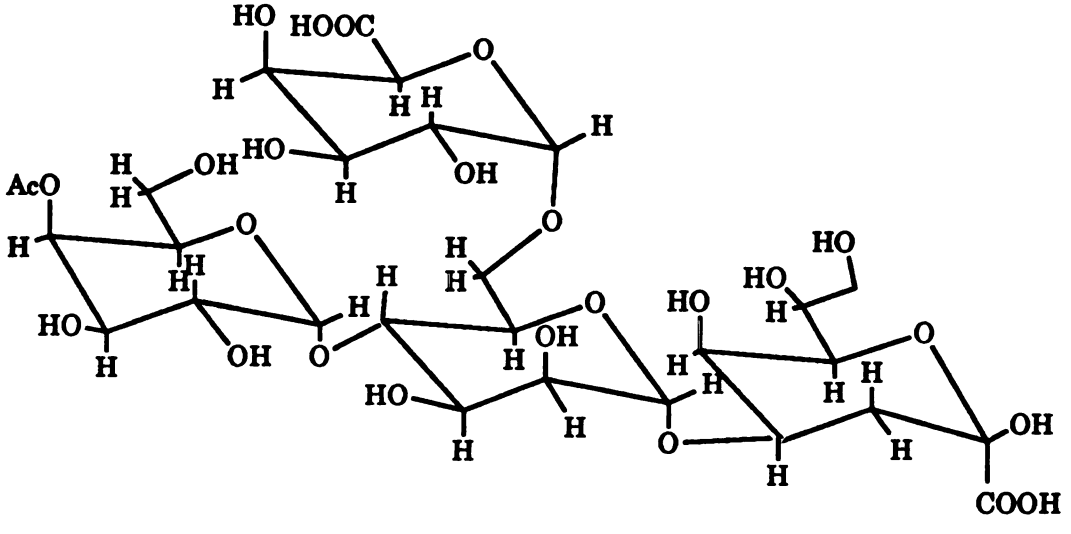


Figure 4.4: Structure of tetrasaccharide from *Rhizobium trifolii* ANU 843 (5).

4.2 Materials and General Methods

4.2.1 Further Purification of the Core Oligosaccharides (4):

The core oligosaccharides were initially isolated by mild acid (1% acetic acid) hydrolysis of lipopolysaccharides, followed by gel filtration chromatography of the water-soluble fraction on a Bio-Gel P-2 column using 1% formic acid as eluant. The fractions containing carbohydrate were identified by colorimetric assay, pooled, then lyophilized for NMR.

4.2.2 Composition Analysis (5):

The reducing terminus and keto-groups on the oligosaccharide were first reduced to alcohols with sodium borodeuteride. Carboxyl groups were converted to methyl esters by 1% HCl in methanol, then transformed to alcohols by reduction with sodium borohydride or sodium borodeuteride. The reduced oligosaccharides were hydrolyzed to monosaccharides with aqueous trifluoroacetic acid (TFA), reduced to alditols with sodium borohydride, and peracetylated with acetic anhydride/pyridine, followed by GC-MS analysis.

4.2.3 Methylation Analysis (6):

The oligosaccharide (1 mg) was pre-reduced with 10% aqueous sodium borohydride, then converted to methyl esters by using 1% methanolic HCl. The methyl esters were reduced with lithium alumina deuteride (LAD). The sample was permethylated using sodium methylsulfenyl anion and methyl iodide. The resulting oligosaccharide was hydrolyzed by 1 *M* TFA, and converted to alditols by reduction with sodium borodeuteride. Peracetylation was performed by dissolving the oligomers in a minimum volume of TFA and an equal volume of a 2:1 mixture of acetic acid and trifluoroacetic acid anhydride. The mixture was left at room temperature for 6 hours, and

concentrated to dryness under a stream of nitrogen. The residue was mixed with chloroform, and partitioned between chloroform and 1 *M* HCl. The chloroform layer was dried with sodium sulfate and evaporated to dryness for GC and GC-MS analysis.

^{1}H NMR spectra were recorded with a Varian VXR300 and VXR500 spectrometer operated at 300 MHz and 500 MHz respectively, and ^{13}C NMR spectra were recorded with the Varian VXR300 spectrometer operated at 75.4 MHz. All chemical shifts are referenced to Me_4Si. GC-MS analyses were performed on a JEOL JMS-AX505H spectrometer interfaced with a Hewlett-Packard 5890A Gas Chromatograph using an ionizing voltage of 70 eV. Electron impact analysis of the per(trideutero)acetylated oligomer was performed on the same instrument at the same voltage by direct insertion probe.

4.3 Results and Discussion

4.3.1 Structure Elucidation of the Inner and Outer Core *Rhizobium Meliloti* 41

From the analysis of inner and outer core components of the lipopolysaccharide of *Rhizobium meliloti* 41, some primary information was obtained which suggested that one of the components was an unusual sugar.

An earlier report (4) indicated that, the core regions of LPS from *R. trifolii* are composed largely of galacturonic acid (7). More detailed studies of *R. trifolii* LPSs has shown that both the *O*-antigen and the core oligosaccharides contain KDO at their reducing end. This result indicates that *Rhizobium* LPSs are quite different from those from *Salmonella* and *E. coli*. The major oligosaccharide from the core region of the lipopolysaccharide from *R. trifolii* ANU 843 was isolated and its structure determined to be a

trisaccharide consisting of two galacturonic acid and a KDO residue (4). The ^{1}H-NMR spectral analysis (Figure 4.5) of the trisaccharide from *Rhizobuim trifolii* ANU 843 showed that the proton of the methylene group at 3-position of the KDO appeared in the proton spectrum as a doublet of doublets (J = 12.5 and 2.9 Hz) at δ1.82 and a triplet (J = 12.5 Hz) at δ2.17. The triplet was assigned to the axial proton which should have a large splitting from the other geminal proton. The equatorial proton should have a comparatively small vicinal coupling, due to being *cis* to the proton on C-4.

Whether or not *Rhizobium* LPSs play a role in the specific recognition between the symbiont bacterium and the host cell is still uncertain (8, 9), but the LPSs may be important in later symbiont events, because the mutants lacking the *O*-antigen portion of the LPS are defective in nodulation (10). In order to understand the roles of *Rhizobium* LPSs in symbiosis, the structural characterization of these molecules is very important. A further study reported the second core oligosaccharide from the LPS of *R. trifolii* ANU 843 is a tetrasaccharide composed of galactose, galacturonic acid, mannose, and a KDO residue. The mannose residue is α-linked to the 4-position of the KDO residue, and the galacturonic acid residue is α-linked to the 6-position of mannose. The galactose residue, acetylated at the 4-position, is attached to the 4-position of mannose by α-linkage. The ^{1}H-NMR spectra (5) (Figure 4.6) of the tetrasaccharide from *R. meliloti* 41 sugestion contained a triplet (J = 11.4 Hz) at δ2.06 and a distorted doublet of doublets with a larger splitting of 11.4 Hz between δ1.74 and 1.86 ppm. These signals were assigned to the axial and equatorial protons, respectively, at the 3-position of a KDO residue.

One of the oligosaccharides from the LPS of *Rhizobium meliloti 41* had an unusual NMR spectrum. In this component there were no usual peaks between δ1.82-δ2.17 for KDO as in Figures 1 and 2, but instead, the 500 MHz

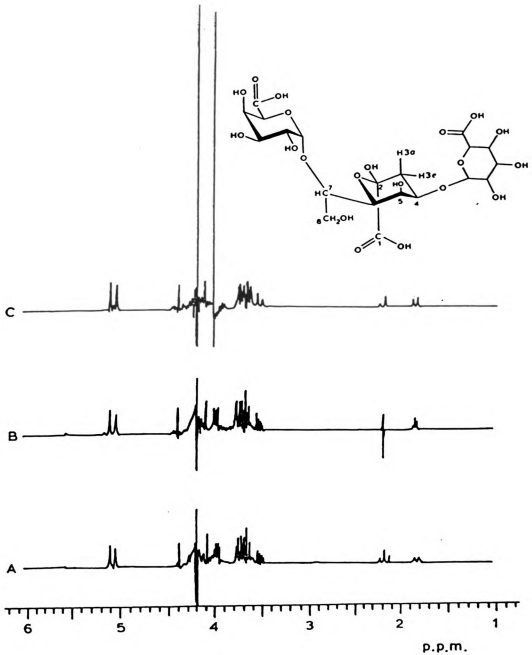


Figure 4.5: Structure of trisaccharide, and (A) Proton NMR spectrum in D_2O from *Rhizobium trifolii* ANU 843, (B) Spectrum after irradiation of the triplet at δ2.17, note the collapse of the doublet of doublets at δ1.82 and the perturbation of the signals at δ4.05, (C) Spectrum of irradiation at δ4.05, note the collapse of the triplet at δ2.17 and the doublet of doublets at δ1.82, reproduced with permission from Dr. Rawle I. Hollingsworth (4).

^{1}H-NMR spectrum showed the protons of a methylene group which appeared only as a doublet of doublets at δ 2.1- 2.25, indicating that they were due to a deoxy group with no neighboring hydrogens and probably, next to a carbonyl group (Figure 4.7). The larger splitting (J = 14.5 Hz) was from geminal proton, and the smaller splitting (J = 1.9 Hz) was due to the long-range coupling from a equatorial proton at the 5-position; therefore, this component should have the galacto-configuration.

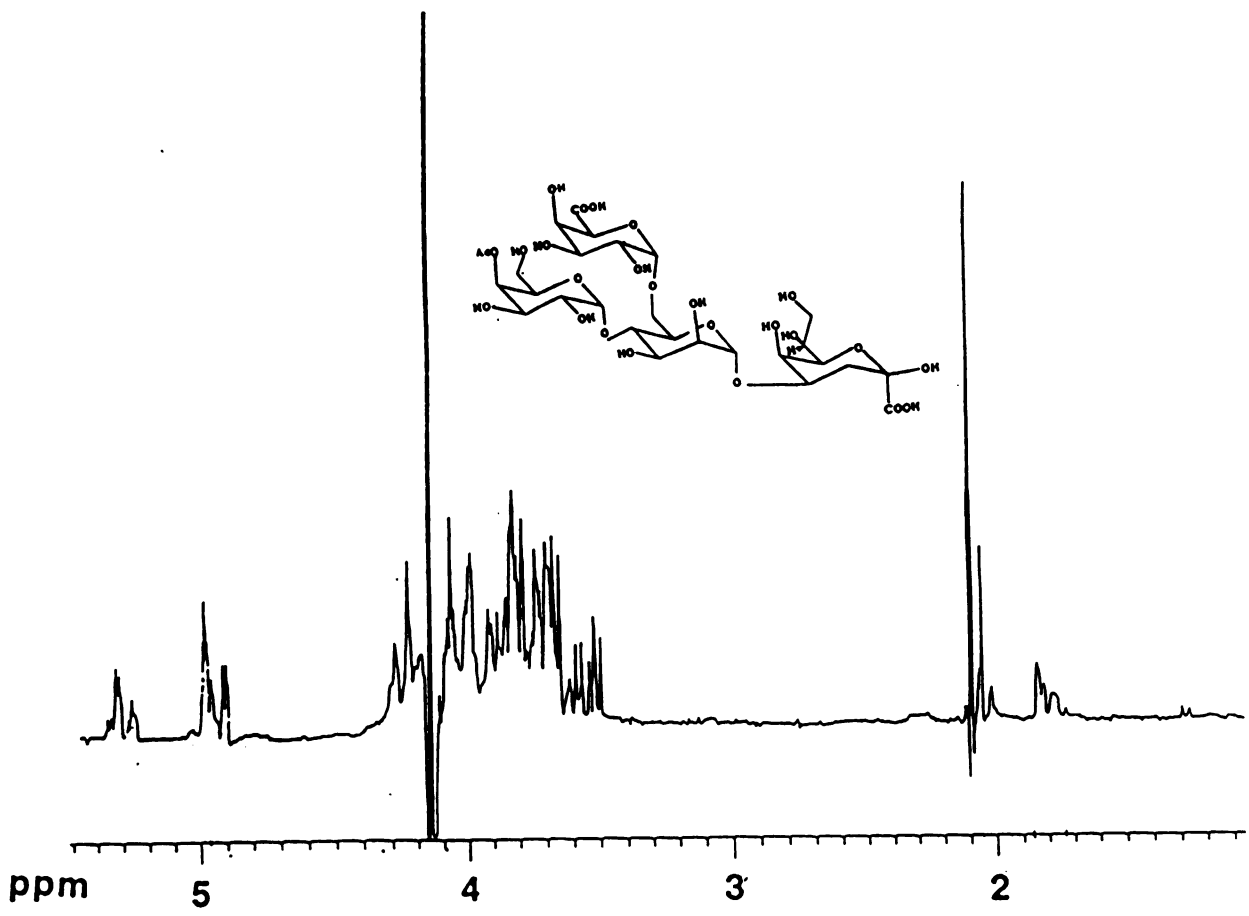


Figure 4.6: Proton NMR spectrum and structure of tetrasaccharide from *Rhizobium trifolii* ANU843, reproduced with permission from Dr. Rawle I. Hollingsworth (5).

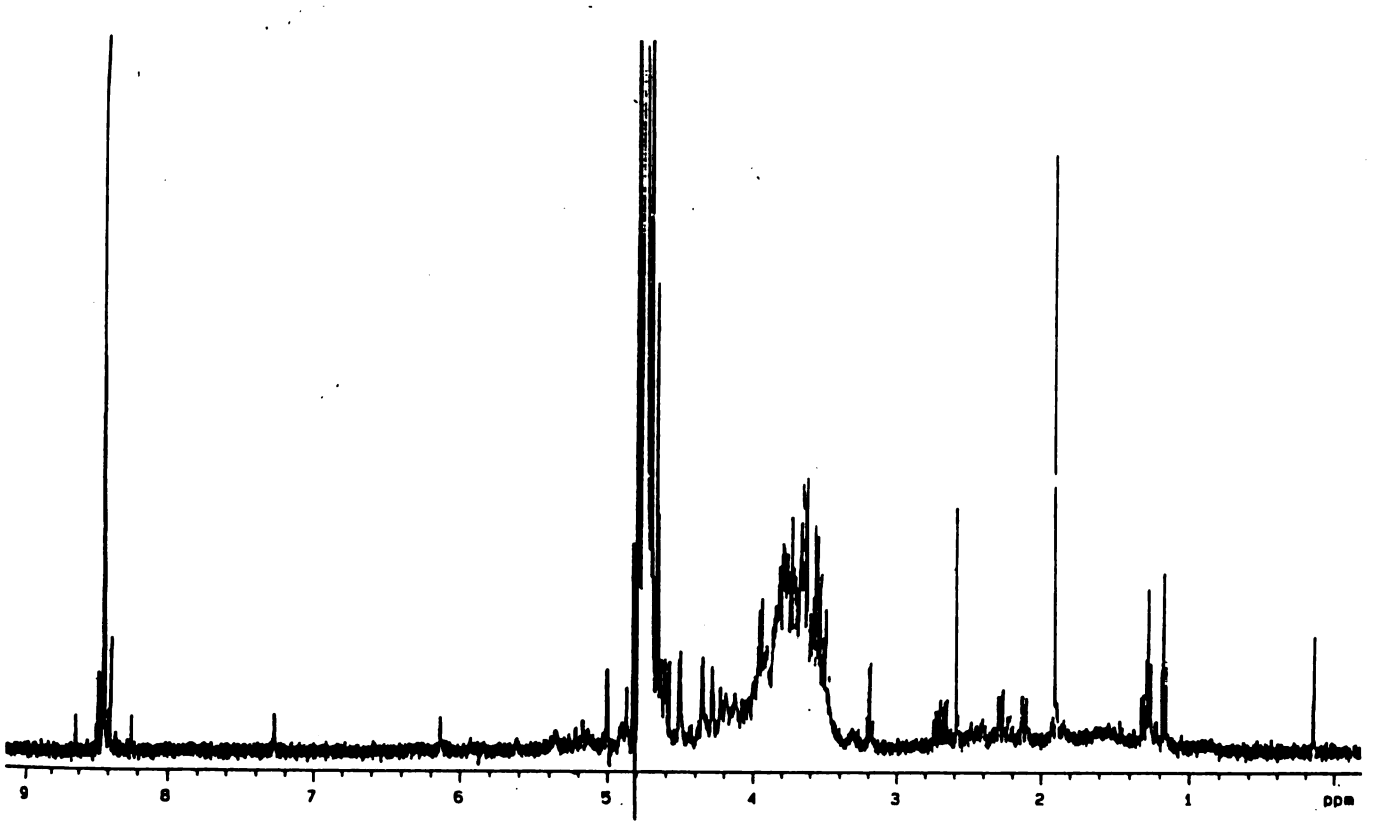


Figure 4.7: Proton NMR spectrum of an unusual component in *Rhizobium meliloti* 41

The components from *Rhizobium meliloti 41* were reduced with 10% aqueous borohydride, which allows the terminal ring to remain in the linear form. The permethylation was effected with 1% methanolic HCl, followed by reduction with borodeuteride. In this stage, the carboxyl group was reduced to CD_2OH. The hydrolysis was achieved by 1 *M* trifluoroacetic acid, which cleaved the oligosaccharides to carbohydrate monomer. The peracetylation was performed by pyridine/acetic anhydride (1:1), or pyridine/acetic anhydride-d_6 (1:1) (Figure 4.8). Finally the undeuterated and deuterated components were analyzed by GC-MS. From GC/MS analysis of alditol acetates of these oligosaccharides, there were four late-eluting components with similar spectra. The retention times of these components were approximately 40 minutes, higher than that of the hexoses such as glucose (Figure 4.9). The presence of four components with similar spectra is consistent with a diketo sugar, since in the diketo sugar there are two chiral centers, and therefore four possible isomers.

In the electron compact (EI) mass spectrum (Figure 4.10), there is a strong cluster at m/z 240. This corresponded to a structure such as the alditol acetate of 3-deoxy-erythro-2,4-heptodiulosonic acid (Figure 4.11). The cluster at m/z 240 corresponds to the starting material with loss of three acetic acid and one formaldehyde molecules. After peracetylation by acetic anhydride-d_6 (Figure 4.12), the cluster at m/z 240 shifted to m/z 247. This indicated the presence of three deuterated acetyl groups (Figure 4.13).

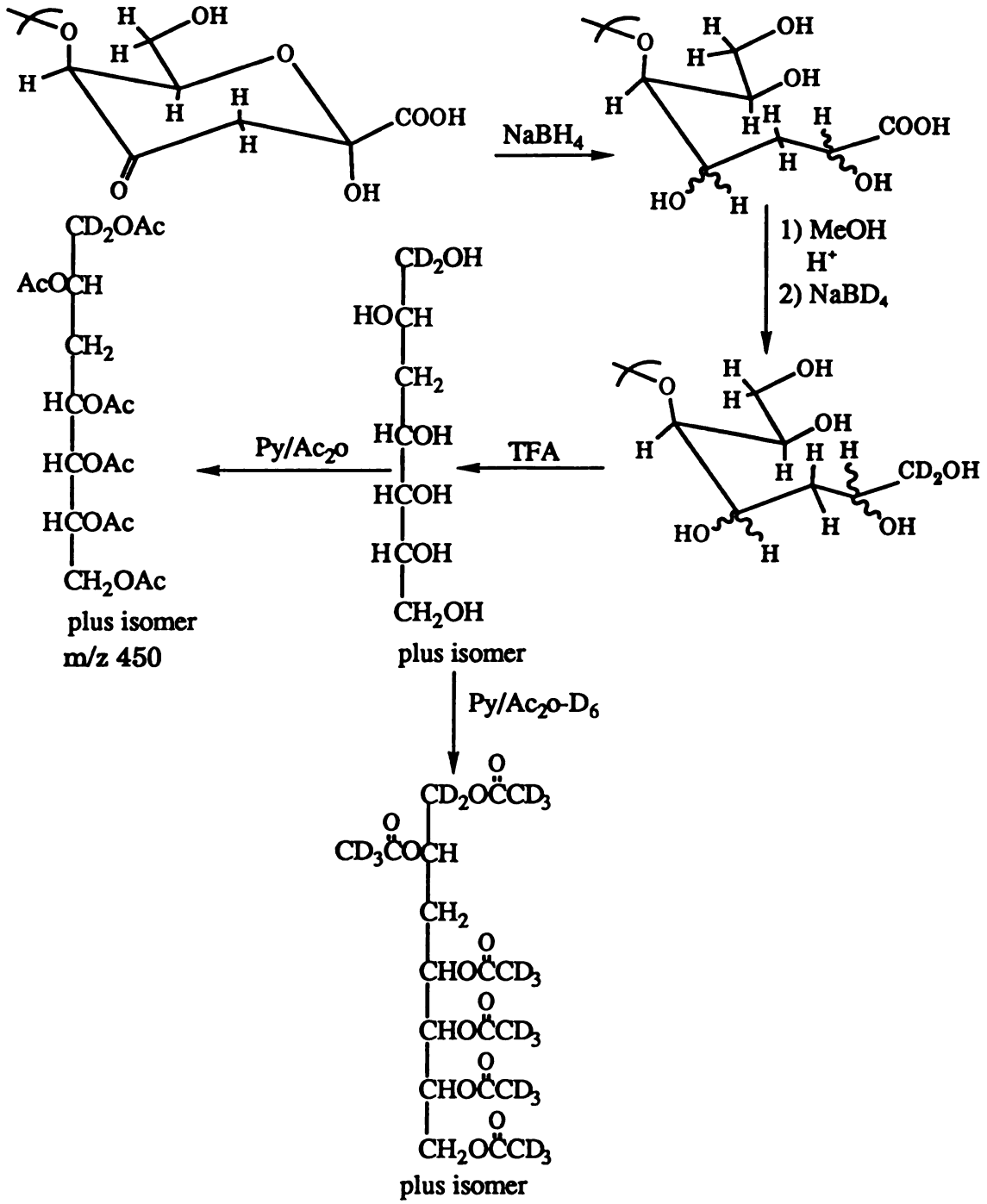


Figure 4.8: Deuterium-labeled, carboxyl-reduced component

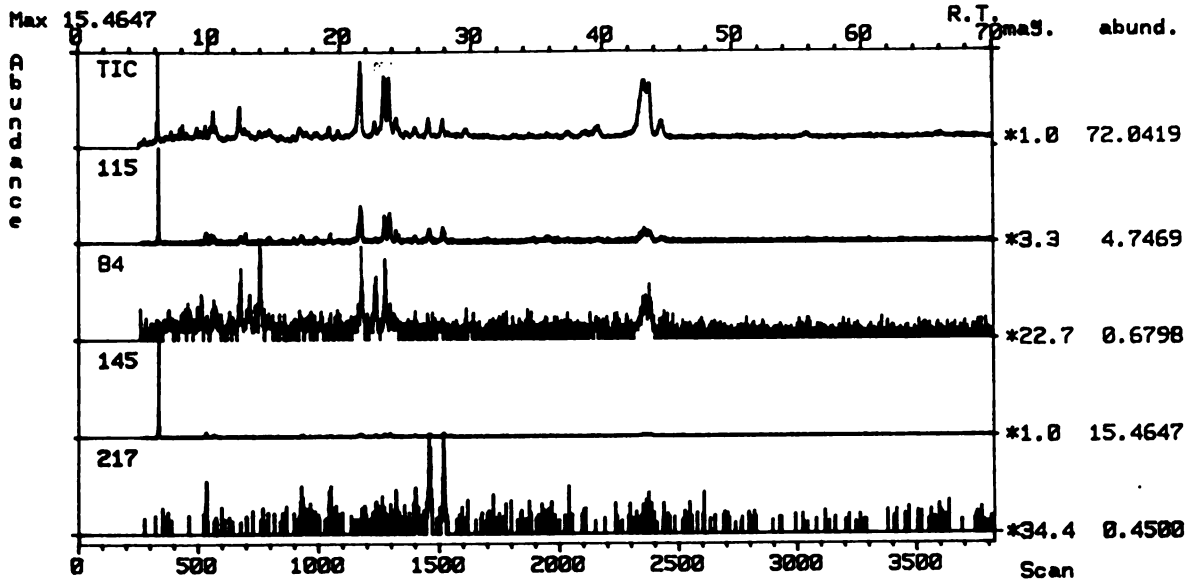


Figure 4.9: GC trace of 3-deoxy-erythro-heptodiulosonic acid derivatives

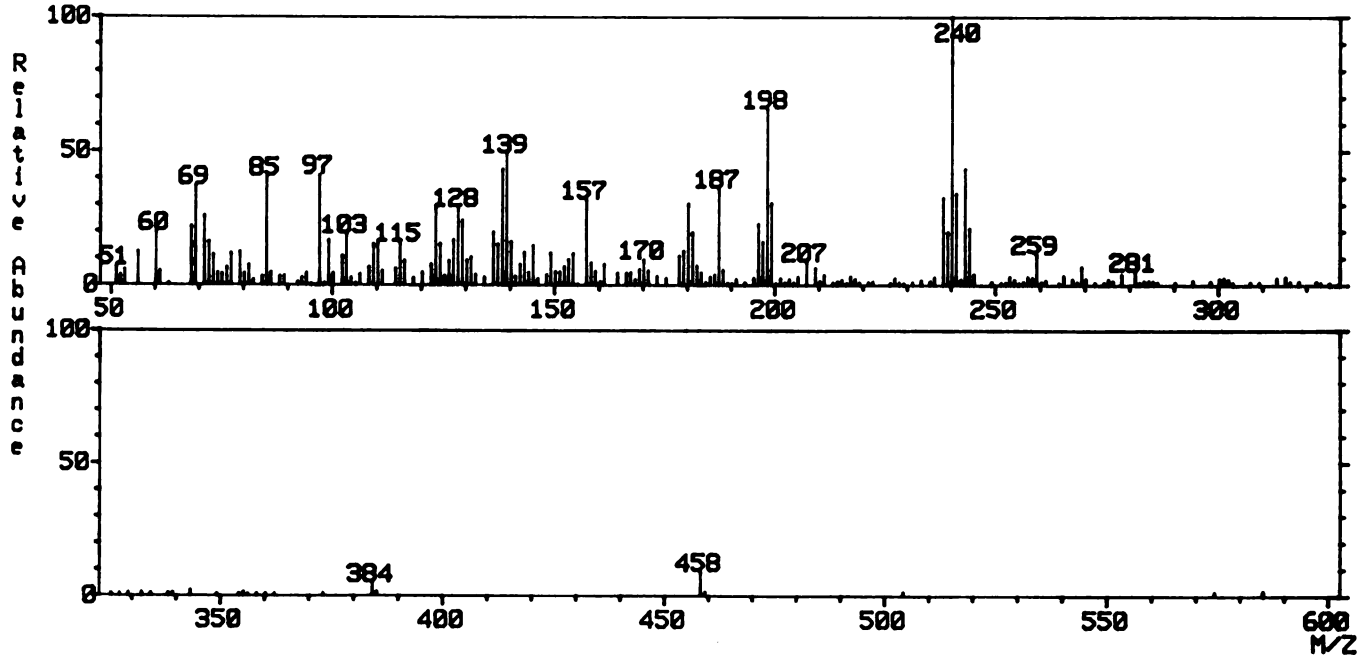


Figure 4.10: EI mass spectrum of 3-deoxy-erythro-2,4-heptodiulosonic acid derivatives

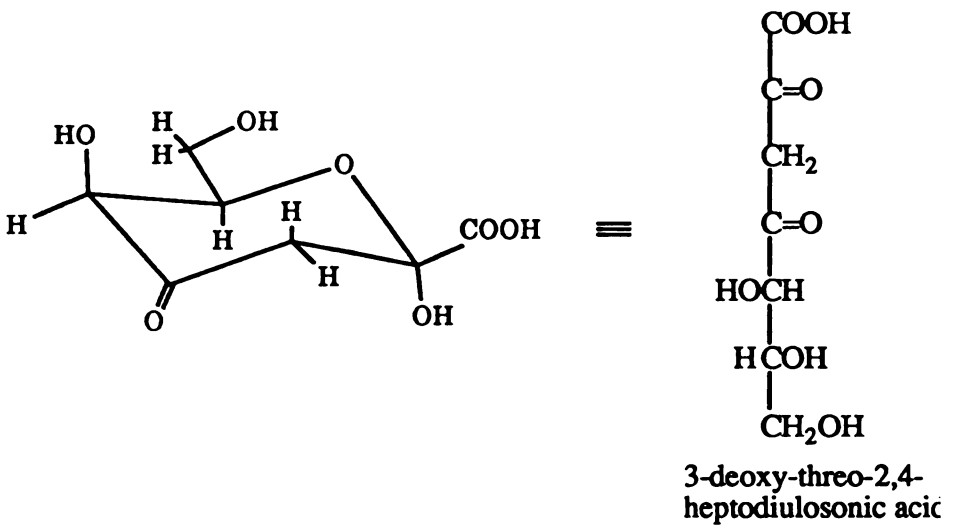


Figure 4.11: The structure of 3-deoxy-threo-2,4-heptodiulosonic acid.

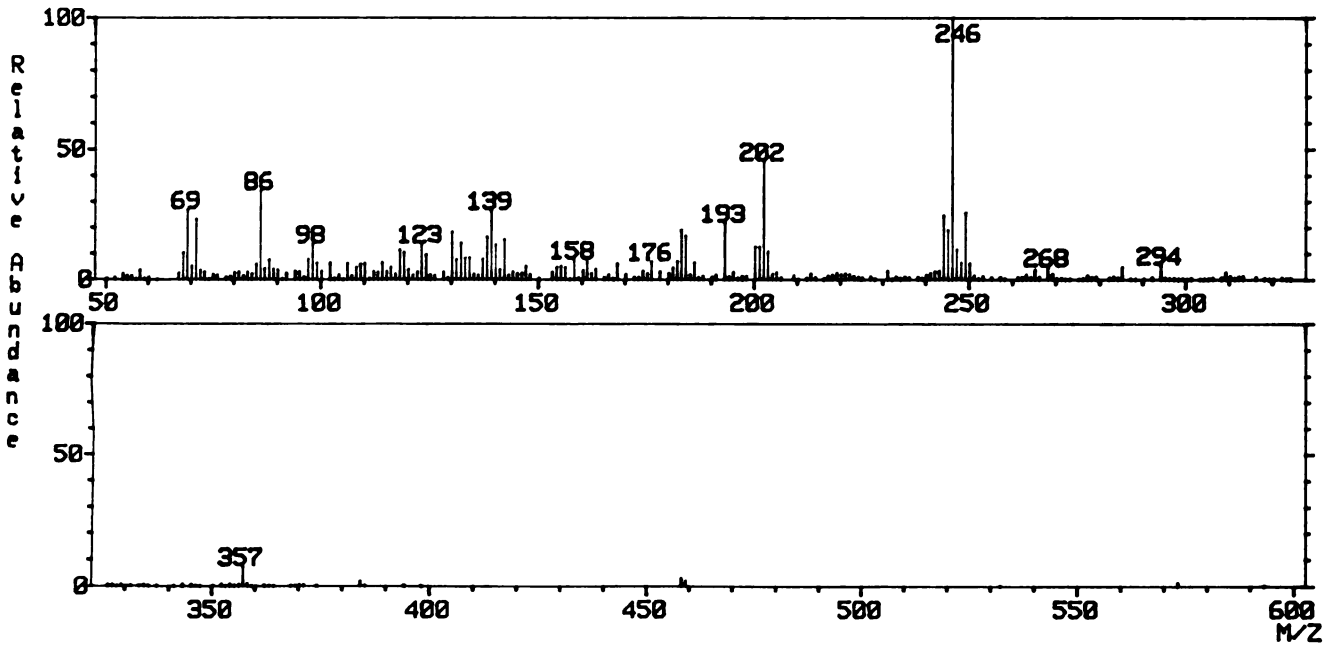


Figure 4.12: EI mass fragmentation of 3-deoxy-erythro-2,4-heptodiulosonic acid deuterated derivatives

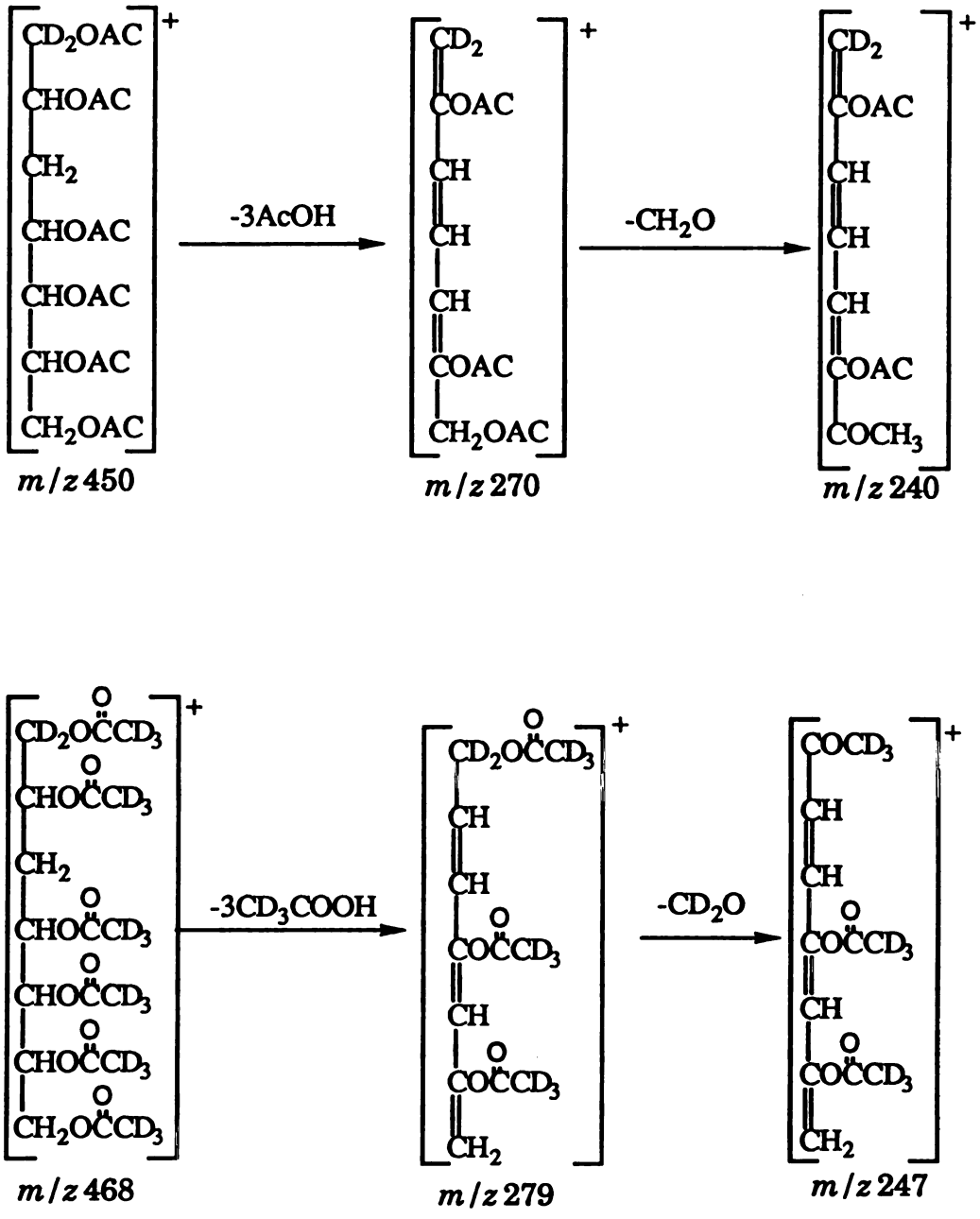


Figure 4.13: EI mass fragmentation of 3-deoxy-heptitol derivatives

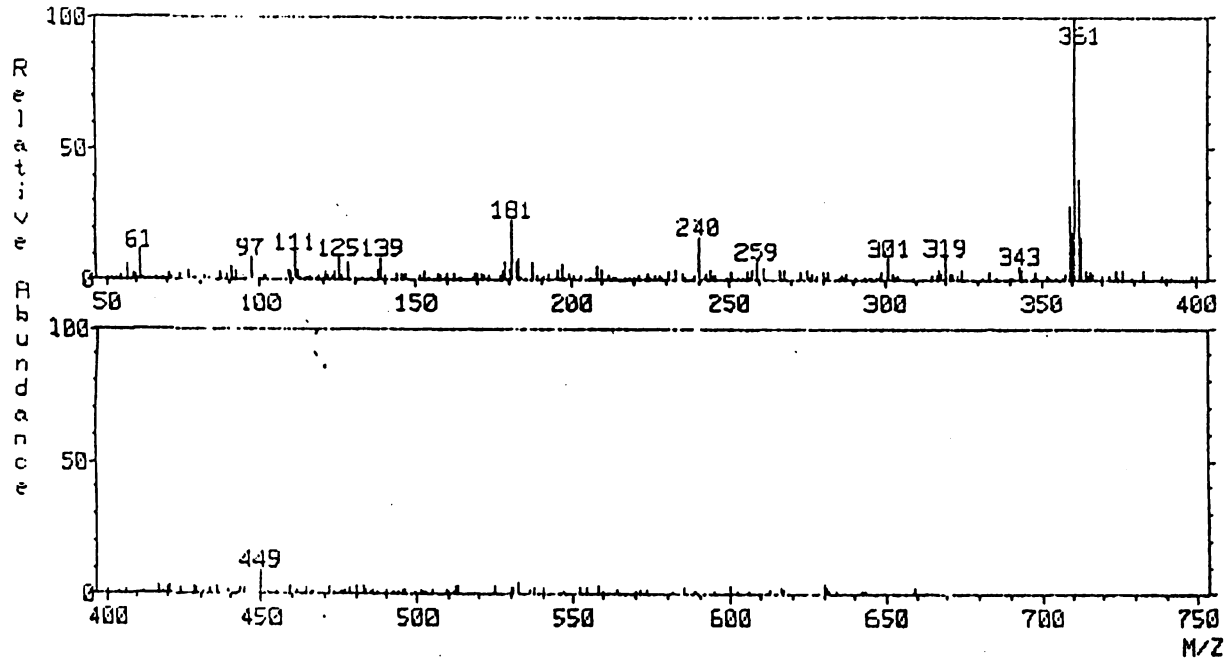


Figure 4.14: CI mass spectrum of peracetylated 3-deoxy-erythro-2,4-heptodiulosonic acid

In the chemical ionization (CI) mass spectrum (Figure 4.14), there is a strong molecular cluster at m/z 361 (m/z + H^+). This molecular cluster corresponds to the acetylated compound minus acetic acid and formaldehyde. When acetic anhydride-d_6 was used, the CI mass spectrum exhibits a strong cluster at m/z 373 (Figure 4.15), showing four deuterated acetyl groups different from the undeuterated one (Figure 4.16).

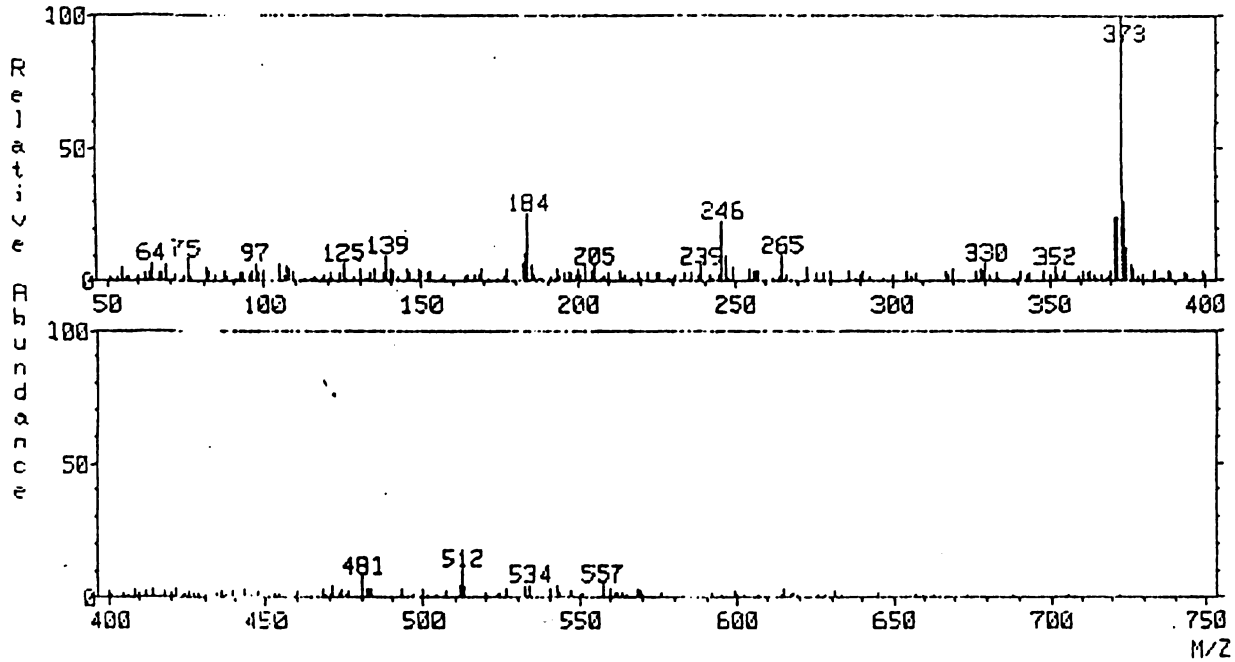


Figure 4.15: CI mass fragmentation of 3-deoxy-erythro-2,4-heptodiulosonic acid deuterated derivatives

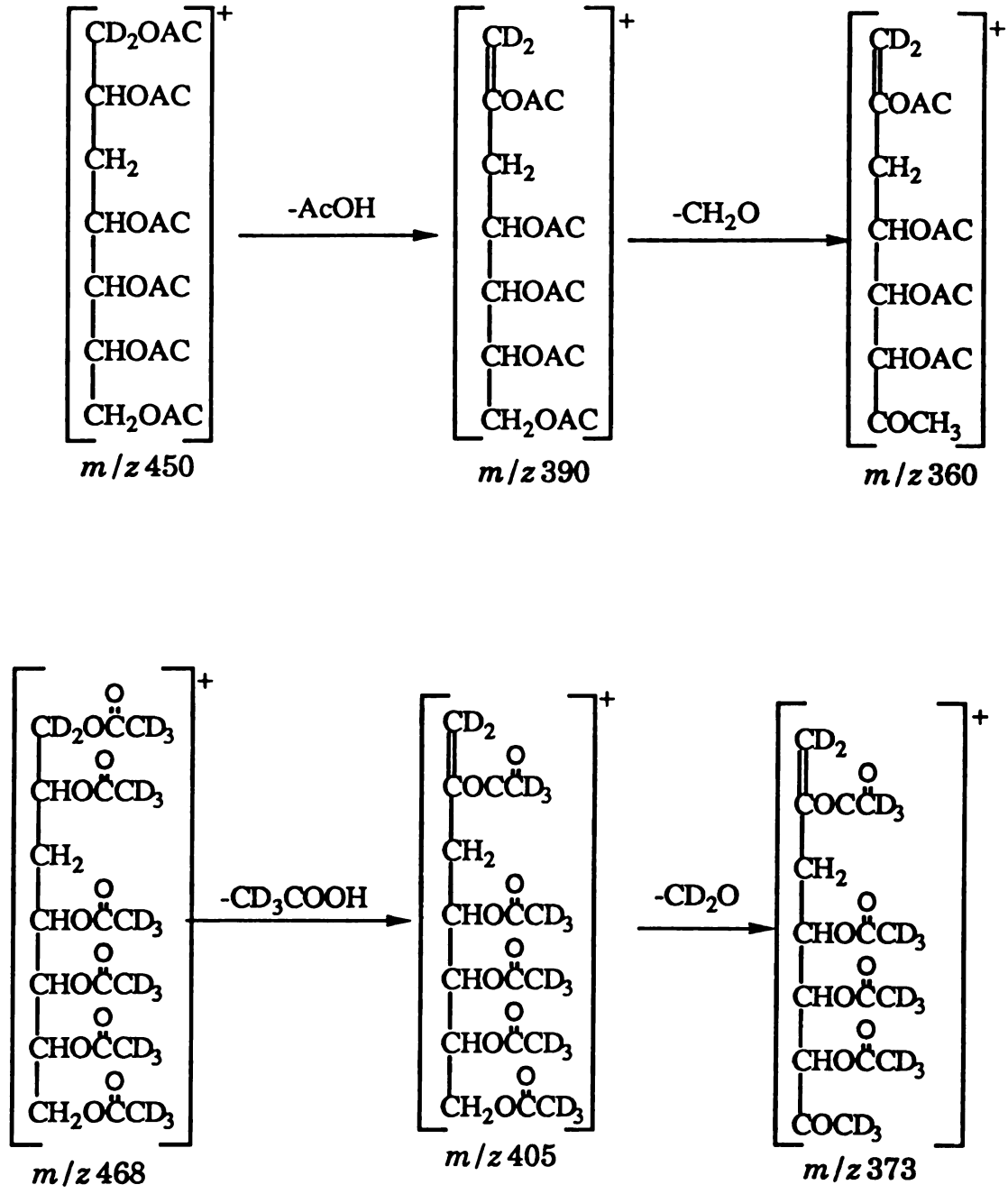


Figure 4.16: Deuterated and undeuterated CI mass fragmentation of 3-deoxy-heptitol derivatives

4.4 Synthesis of 3-Deoxy-Heptitol Derivative

4.4.1 The First Pathway to Synthesize 3-deoxy-heptitol Derivative

To further understand the specific genetic character of *Rhizobium meliloti* 41 (11) and identify the unusual component, a series of syntheses were performed. We started from 2-deoxy-D-glucose (without consideration of the stereochemistry at the four position). The aldehyde group was protected as a 1,3-propylene dithioacetal (12); the structure of the protected product was confirmed by NMR spectroscopy, and peaks were assigned by proton decoupling. All hydroxyl groups were then protected by benzyl bromide, and the anomeric proton was replaced by an iodomethyl group. The iodide group was then transformed to a hydroxyl group. The 1,3-dithiane was cleaved by mercury chloride (13, 14). It can also be cleaved with boron trifluoride (15) and N-bromosuccinimide (16). The reduction of the keto group at the 2 position was achieved by sodium borodeuteride. The benzyl ethers were cleaved by catalytic hydrogenation, and the resulting product was peracetylated (Figure 4.17). The final product was purified by TLC, then analyzed by GC and GC-MS. One component had a similar retention time to the component separated from bacterial LPS, but the resulting EI mass fragmentation from GC-MS (Figure 4.18) was not similar to the spectrum obtained from the bacterial component (Figure 4.1). The CI mass spectrum from the synthesized component had a very strong peak at m/z 419, which corresponds to the acetylated compound with loss of formaldehyde (Figure 4.19). These mass spectrometry data from the synthesized compound indicate that the unusual component we identified in the oligosaccharide from the LPS of *Rhizobium meliloti* 41 possibly is a 3-deoxy-2,4-heptodiulosonic acid.

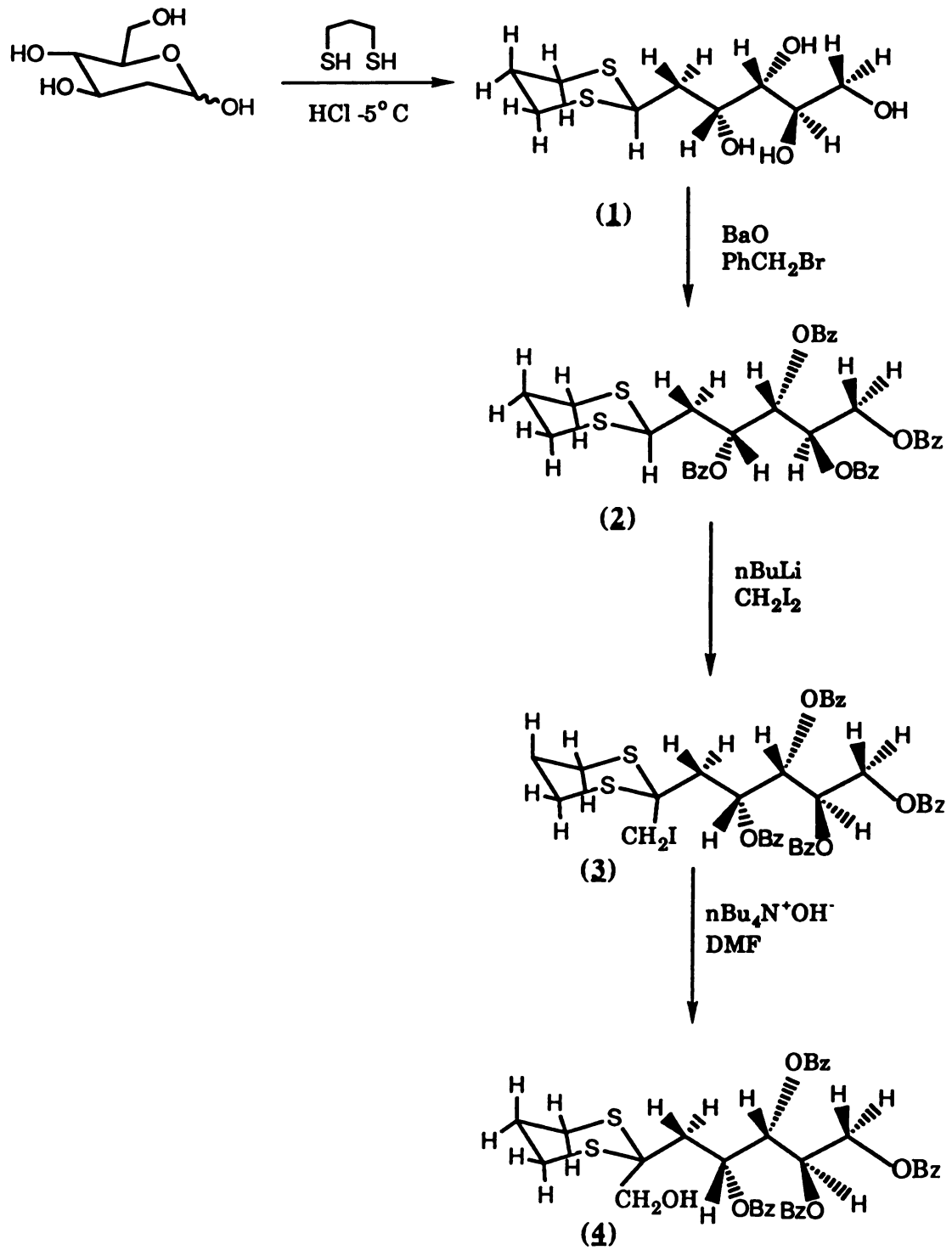
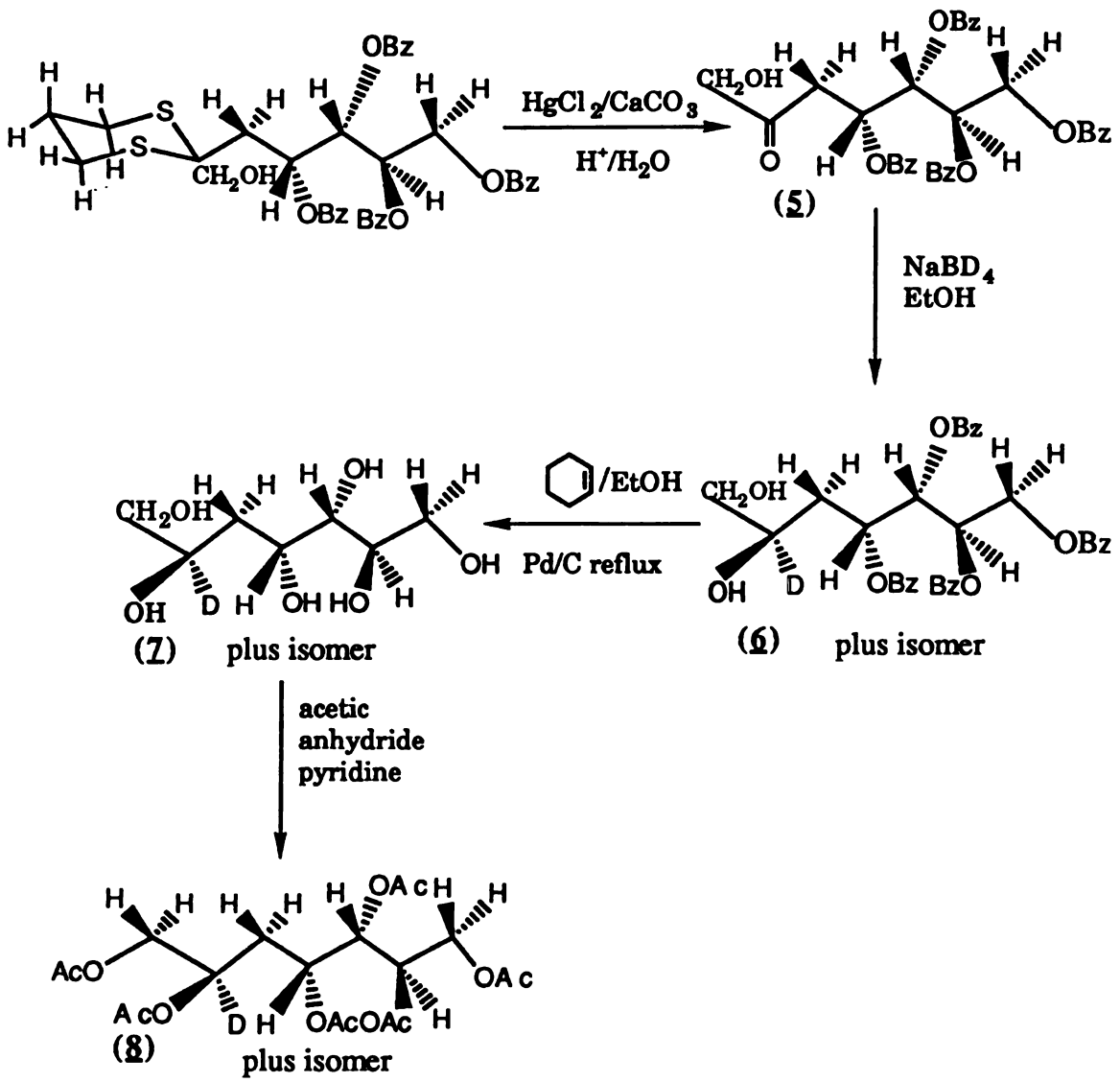


Figure 4.17: The first synthetic pathway for 3-deoxy-heptitol derivatives.



(continued) **Figure 4.17:** The first synthetic pathway for 3-deoxy-heptitol derivatives.

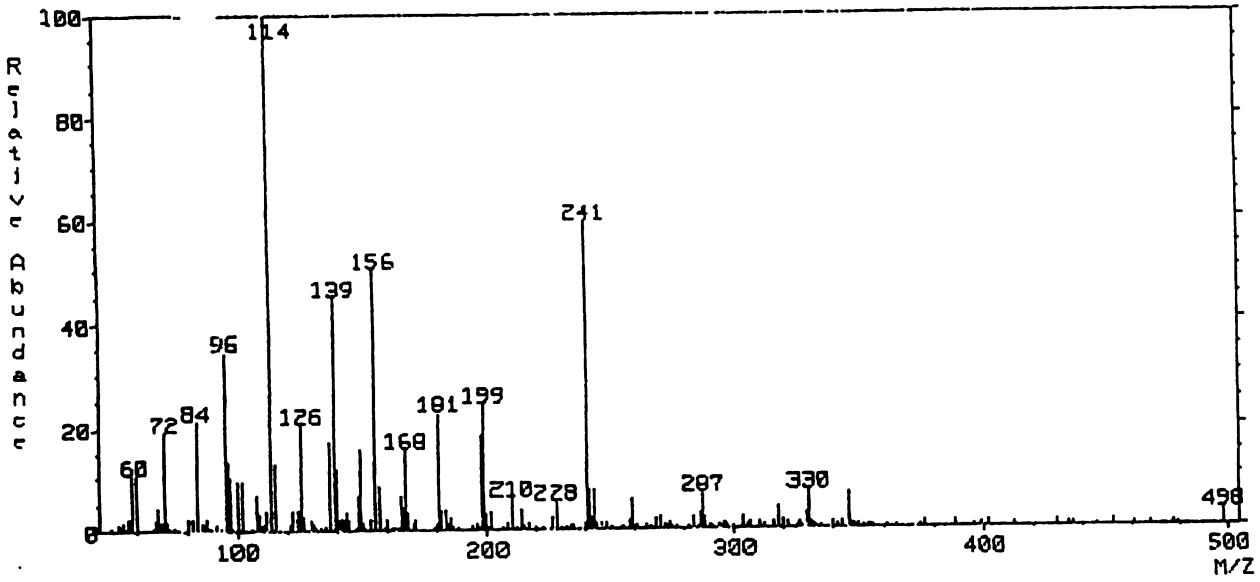


Figure 4.18 EI mass spectrum from synthetic 3-deoxy-heptitol derivatives

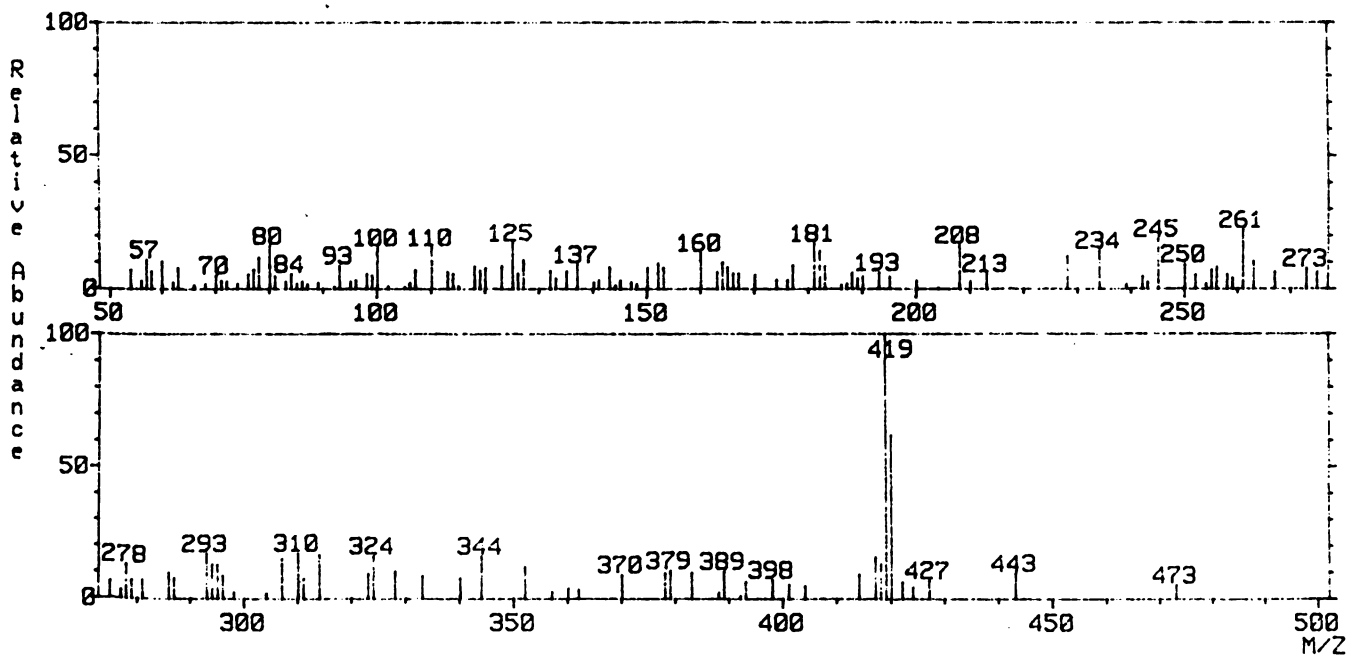


Figure 4.19 CI mass spectrum from synthetic 3-deoxy-heptitol derivatives

4.4.2 Second Pathway to Synthesize 3-Deoxy-Heptitol Derivatives:

A second pathway for synthesis of 3-deoxy-heptitol derivatives is described in Figure 4.20. The general strategy is similar to the first pathway, except that the protecting groups were changed from benzyl to isopropylidene (12). This was intended to make the system more easy to work up. DMF/NaOH (17), $AgNO_3$/acetone/water (18), and tetrabutylammonium hydroxide (19) were tried for the conversion of compound 10 to 11, but the yield was very low.

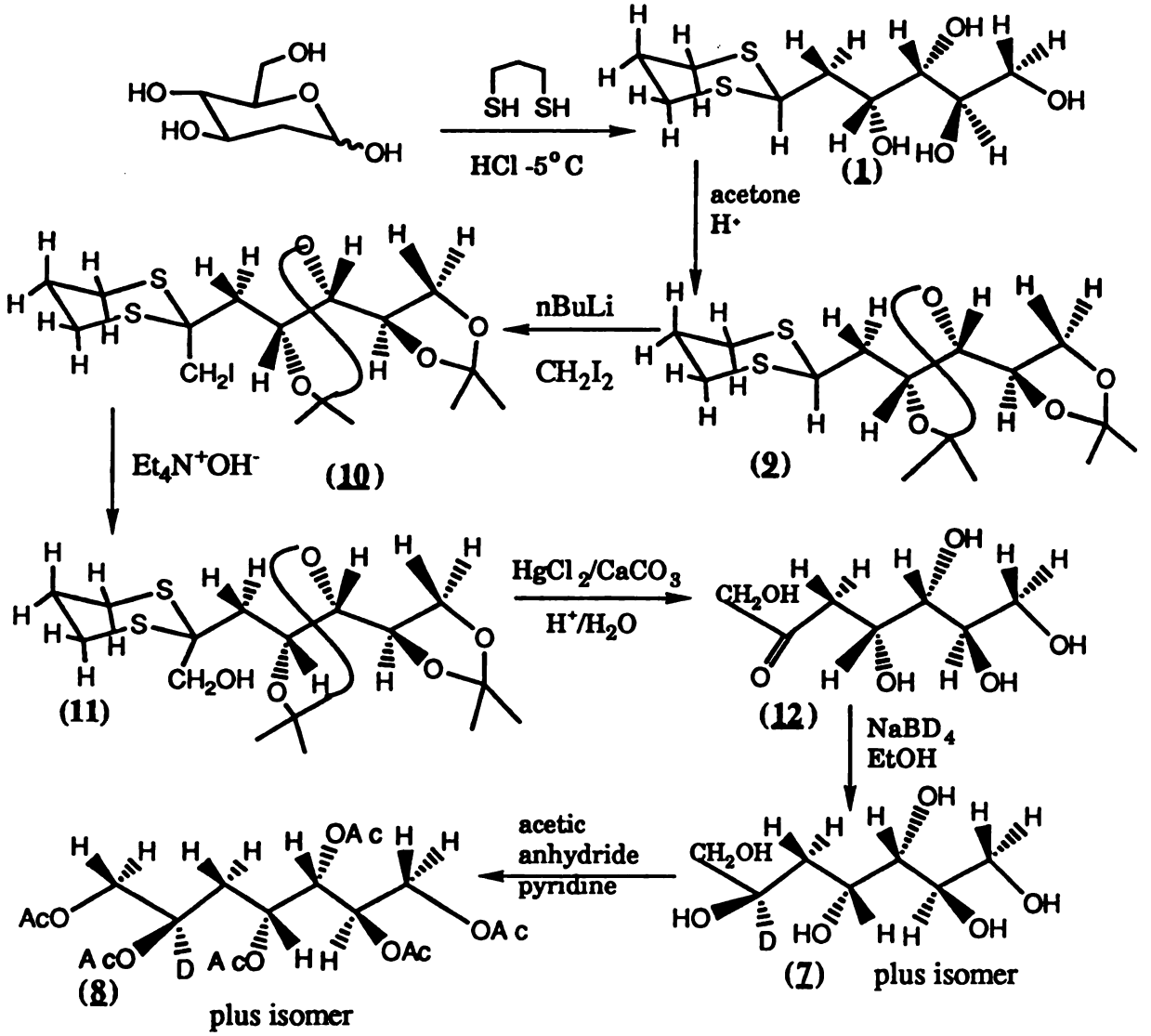


Figure 4.20: The second pathway to synthesize 3-deoxy heptitol derivatives.

4.4.3 The Pathway to Synthesize 3-deoxy-D-erythro-2,4-heptodiulosonic Acid

The fourth strategy to synthesize the target compound is shown in Figure 4.22. In this pathway, some of the reaction failed and are reported here to avoid future workers making same mistakes. First, oxidization of the methyl-4,6-*O*-benzylidene-β-D-galactopyranoside (20) was attempted to make D-erythrose, with the periodic acid/sodium bicarbonate, and lead tetraacetate (21, 22) tried as the oxidation reagent. Methyl-4,6-*O*-benzylidene-β-D-galactopyranoside is not very water-soluble, so it could not be mixed very well with periodic acid/sodium bicarbonate. The conversion of oxidation was very low when using lead tetraacetate. To yield the solubility of the starting material, 4,6-*O*-benzylidene-D-glucose was synthesized. The 4,6-*O*-benzylidene-D-glucose was oxidized by sodium meta-periodate to the desired 2,4-*O*-benzylidene-D-erythrose (23), which existed as a dimer. When protection of the aldehyde group in 2,4-*O*-benzylidene-D-erythrose was attempted with 1,3-propane dithiol, the benzylidene group was cleaved under the acidic condition. To avoid deprotection, a series of different acid conditions was tried. This included HCl etherate solution (24), boron triflouride etherate solution (25), dilute trifluoroacetic acid, and zinc chloride with boron trifluoride. The benzylidene group was always cleaved to form benzaldehyde propylene dithioacetal, no desired product was obtained.

4,6-*O*-propylidene-D-glucose was synthesized to reduce the lability of the protective acetal group. At low temperature, 1 *M* HCl etherate solution was used as the catalysts. The 1,3-propyl dithioacetal was formed without disrupting the protective acetal group.

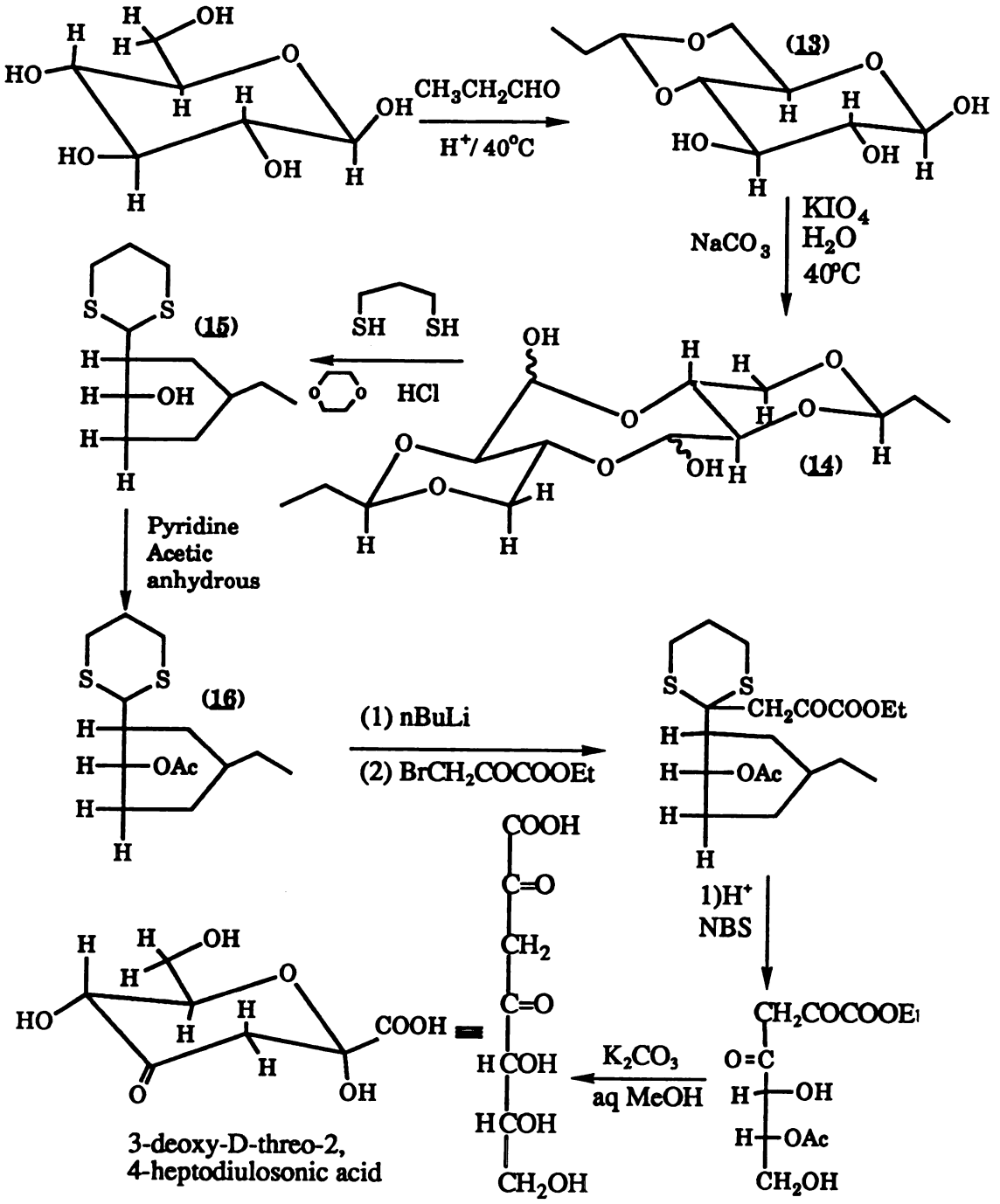


Figure 4.21: The pathway to synthesize 3-deoxy-D-erythro-2,4-heptodiulosonic acid

An acetyl group was attached to the 3-hydroxyl group in the resulting product at the C3 position. In the alkylation stage, the acetyl group was destroyed by the n-butyl lithium; the propylidene group also underwent elimination.

4.5 References

1. Dazzo, F. B. and Hollingsworth, R. I., *Biol. Cell* (1984) 51 267-274.
2. Christian R. H. Raetz, *Ann. Rev. Biochem.* (1990) 59 129-170.
3. Westphal, O., Luderitz, O., Galanos, Ch., Mayer, H. and Rietschel, E. T., *"Advance in Immunopharomaco"* (1985) 3 13-43.
4. Carlson, R., W., Hollingsworth, R. I. and Dazzo, F. B., *Carbohydr. Res.* (1988) 176 127-135.
5. Hollingsworth, R. I., Carlson, R W., Garci, F., and Gage, D. A., *J. Biol. Chem.* (1989) 264 9294-9299.
6. Prehm, P., Strim, S., Jann, B., and Jann, K., *Eur. J. Biochem.* (1975) 56 41-55.
7. Carlson, R. W., *J. Bacteriol.* (1984) 158 1012-1017.
8. Law, I. J., and Strijdom, B. W. *J. Biol. Chem.* (1977) 9 79-84.
9. Smit, G., Kijne, J. W., and Lugtenberg, B. J., *J. Bacteriol.* (1987) 169 4294-4301.
10. Noel, K. D., VandenBosch, K. A., and Kulpaca, B., *J. Bacteriol.* (1986) 168 1392-1401.
11. Williams, M. N. V., Hollingsworth, R. I., Klein, S., and Signer, E. R., *J. Bacteriol.* (1990) 172, 2622-2623.
12. Frost J. W. and Knowles, J. R., *Biochemistry* (1984) 23 4465-4469.
13. Corey, E. J., and Bock, M. G., *Tet. Lett.* (1975) 38 3269-3270.

14. Meyers, A. I., Comins, D. L., Henning, D. M. R., Shimizu, K., *J. Am. Chem. Soc.* (**1979**) 101 7104-7105.

15. Corey, E. J., and Hase T., *Tet. Lett.* (**1979**) 4 335-338.

16. Corey, E. J., and Erickson, B. W., *J. Org. Chem.* (**1971**) 36 3553-3560.

17. Numazawa, M., and Nagaaoka, M., *J. Org. Chem.* (**1982**) 47 4024-4029.

18. Wasserman, H. H., Mariano, P. S., and Keehn, P. M., *J. Org. Chem.* (**1971**) 36 1765-1769.

19. McCourt, D. W., Roller, P. P., and Gelboin, H. V., *J. Org. Chem.* (**1981**) 46 4157-4161.

20. Boffi, C., and Ferrari, L., *Die Starke* (**1969**) 21 100.

21. Perlin, A. S., and Brice, C., *Can. J. Chem.* (**1956**) 34 541-553.

22. Murry, J. E. Mc, adn Blaszczak, L. C., *J. Org. Chem.* (**1974**) 39 2217-2222.

23. Baggett, N., Buck, K. W., Foster, A. B., Rees, B. H., and Webber, J. M., *J. Chem. Soc. (C)*, (**1966**) 212-215.

24. Schmidt, U., and Ohler, E., *Angew. Chem. Int. Ed. Enl.* (**1977**) 16 327.

25. Seebach, D., Jones, N. R., and Cory, E. J., *J. Org. Chem.* (**1968**) 33 300-305.

CHAPTER 5

EXPERIMENTAL FOR CHAPTER 4

5.1 General Methods in Structure Elucidation

5.1.1 Microorganisms

The lipopolysaccharides were obtained from *Rhizobium meliloti 41.*

5.1.2 Dimethylsulphinyl Sodium Preparation (1)

Dry DMSO 30 mL was added from a dry syringe through a serum cap to a three-neck 250 mL flask equipped with a calcium chloride drying tube and containing 1.5 g of sodium hydride (washed twice with 20 mL anhydrous ether). The system was protected under nitrogen. The reaction mixture was stirred at 50° C for 3 hours. After a transparent green color was formed, the reaction was judged to be complete, and was transferred to several one dram vials, flushed with nitrogen, sealed by a Teflon-lined screw-cap, and stored frozen in the dark.

5.1.3 Hydrolysis of Lipopolysaccharides (2)

Acetic acid (1 mL of 1%) was added to lipopolysaccharide (~ 2 mg), and the solution heated at 100° C for 1 hour.

Ten mL water/chloroform (1:1) was added to the sample, and the mixture shaken well. The layers were separated and the extraction was repeated three times adding chloroform (5 mL) each time. The chloroform layers (containing lipid A) were combined together and reduced to dryness. The carbohydrate was left in the water layer which was concentrated and passed through a Bio-Gel P-2 column (200-400 mesh,16 mm x 90 mm), with

1% formic acid as eluant. Fractions of 90 drops (2 mL) were collected in each test tube by a fraction collector.

5.1.4 Colorimetric Methods for Carbohydrate Detection

1. Phenol/Sulfuric acid Assay (total carbohydrate) (3):

After separation on the Bio-gel P-2 column, 75 μL of sample was taken from each fraction, and 0.5 mL 5% aqueous phenol and 2.5 mL sulfuric acid were added. The absorbance at 490 nm was measured on a Milton Royspectrometer using a tungsten-visible lamp. The absorbance for each fraction was plotted versus fraction number. The samples containing carbohydrate were combined and then lyophilized for instrumental analysis.

2. KDO Assay (4):

(1) From each column fraction, 50 μL of sample was transferred to a test tube.

(2) 50 μL of 0.5 *N* H_2SO_4 was added to each test tube, and shaken well.

(3) All samples were placed in a boiling water bath for 8 minutes, then cooled to room temperature.

(4) 50 μL H_5IO_6 (reagent A) was added to each sample, and shaken well.

(5) All samples were left for 10 minutes at room temperature.

(6) 200 μL arisenite (reagent B) was added to each sample, and the tubes shaken well.

(7) 800 μL thiobarbituric acid (reagent C) was added to each sample.

(8) All samples were placed in a boiling water bath for 10 minutes.

(9) All samples were cooled to room temperature under running tap water.

(10) 2.0 mL butanol (reagent D) was added to each sample. After addition of each reagent, all samples were well mixed.

(11) The butanol layer was recovered and its absorbance measured at 552 nm and plotted versus fraction numbers.

(12) All samples were combined based on the distribution of carbohydrate, then frozen and lyophilized for NMR.

Reagent A: 2.28 g H_5IO_6 (*O*-paraperiodic acid) was dissolved to 100 mL water (0.1*N*), and stored in a dark bottle.

Reagent B: 4 g of $NaAsO_2$ was dissolved in 100 mL of 0.5 *N* HCl (4.0%).

Reagent C: 600 mg of 2-thiobarbituric acid dissolved in 100 mL boiling water (0.6%), then cooled to room temperature. This reagent must be prepared prior to use.

Reagent D: Five mL concentrated HCl was added to 95 mL n-butanol.

5.1.5 Compositional Analysis (2)

The reducing terminus was pre-reduced by sodium borohydride: oligosaccharide (0.2 mg) was dissolved in 0.1 mL water, then 0.2 mg of sodium borohydride was added. The mixture was allowed to stand at room temperature for 3 hours, and the excess borohydride was decomposed by adding 10% acetic acid dropwise. The solution was concentrated to dryness under a stream of nitrogen.

The carboxyl groups were methylated and then reduced with sodium borodeuteride. In a Teflon-lined, screw-capped vial, the above sample was dissolved in 1 mL of 1% HCl in methanol. The solution was allowed to stand at 60° C for 3 hours and evaporated to dryness. One mL of methanol was added, and the solution concentrated to dryness to remove trace

amounts of acid. The residue was dissolved in 0.1 mL water, then 0.2 mg of sodium borodeuteride (98% D; obtained from Aldrich Chemical Company, Inc.) was added. The mixture was allowed to stand at room temperature for 16 hours, and the excess borodeuteride was destroyed by drop-wise addition of 10% acetic acid until decomposition was completed. The solution was then evaporated to dryness.

The above sample was hydrolyzed by treatment with 1 mL of 1 *M* TFA at 100° C for 3 hours. The hydrolysate was concentrated to dryness, then 1 mL water was added and the solution concentrated to dryness to remove trace amounts of TFA.

The aldehyde groups of the monomers were reduced by treatment with 0.2 mg sodium borohydride in 0.1 mL water for 1 hour. The solution was then evaporated to dryness. One mL of methanol was added to the sample, then the evaporation step was repeated. This process of addition of methanol and evaporation was repeated 4 times to remove trace amounts of sodium borohydride and boric acid as the volatile trimethyl ester.

The hydroxyl group of hydrolyzed and reduced oligosaccharides was acetylated by acetic anhydride and pyridine (100 μL of a 1:1 mixture). This was accomplished by sonicating for 1 minute and then heating at 100° C for 15 minutes, followed by sonication and heating at 100° C for an additional 15 minutes. The mixture was evaporated to dryness. Finally, the sample was dissolved in a mixture of 1 mL 1% hydrochloric acid and 1 mL of chloroform to extract the acetylated sugars. This extraction was achieved twice and the chloroform layer was collected and reduced to dryness for GC and GC-MS analysis.

5.1.6 ^{1}H and ^{13}C-NMR Analysis

For ^{1}H and ^{13}C-nuclear magnetic resonance (NMR) analysis of oligosaccharides: samples (1 mg) were dissolved in 0.7 mL of deuterium oxide (99.96% D, obtained from Cambridge Isotope Laboratories). The spectra were recorded on Varian VXR 300 MHz and Varian VXR 500 MHz (for ^{1}H) instruments. The chemical shifts were measured relative to the water peak (4.65 ppm at room temperature).

5.1.7 GC Analysis

The GC analyses were performed on a Varian 5800 gas chromatograph equipped with a flame ionization detector and a Hewlett-packard 3390A digital integrator. All GC analyses were done on DB 225 Fused Silica Capillary Columns (30 m length, 0.32 mm diameter, film thickness 0.25 mm, from J & W Scientific).

5.1.8 GC-MS Analysis

GC-MS analyses were carried out on: 1) JEOL JMS-AX505H mass spectrometer interfaced with a Hewlett-Packard 5890A Gas Chromatograph or 2) Hewlett-packard 5970 series mass-selective detector interfaced with a 5890A Gas Chromatograph. Data was analyzed with a Hewlett-packard 59970 MS ChemStation.

5.1.9 Other Facilities

UV spectra were obtained with a Varian DMS 200 UV-Visible Spectrophotometer.

Gel filtration column chromatography was also followed by UV absorbance using an ISCO V^{4} Absorbance Detector.

An HBI Multistaltic Pump was used in column chromatography.

A GILSON Micro Fractionator was used to collect samples from the columns during gel filtration chromatography.

A spectronic 21 made by the MILTON ROY Company was used for measuring the UV absorbance in tubes (colorimetric assays).

UV absorbance was viewed on TLC (thin layer chromatography) using the UVP. Inc. view box.

5.2 Synthetic Transformation

5.2.1 2-Deoxy-D-Arabino-Hexose Propylene Dithioacetal in Figure 4.20 compound 1 (5):

2-deoxy-D-glucose, 5.03 g (0.031 mole), dissolved in 6 mL cold (0° C) hydrochloric acid (37%) and 3.3 mL (0.033 mole) propylene 1,3-dithiol were added to a 50 mL round-bottom flask. The mixture was stirred in an ice-salt bath at -5° C for 20 minutes. Ice-water was added to the solution, all solvent was removed by a rotary evaporator, and the product was obtained as a white solid. The solid was recrystallized from methanol/acetone (1:10), and filtered. The residue was washed with hexane. ^{1}H-NMR (500 MHz, methanol-d_4): δ1.78 (m, 1 H,w/48 Hz, H-8), δ1.96 (dd, 1 H, J = 9.9 and 3.9 Hz, H-2'), δ1.98 (dd, 1 H, J = 8.9 and 3.9 Hz, H-2), δ2.05 (m, 1 H, w/34.9 Hz, H-8'), δ 2.78 (dt, 2 H, J = 2 and 4 Hz, H-7' x 2), δ 2.85 (ddd, 2 H, J = 11.5, J = 29.9 and 2 Hz, H-7 x 2), δ3.27 (dd, 1 H, J = 1 and 9.9 Hz, H-4), δ3.55 (dd, 1 H J = 6.5 and 11.5 Hz, H-6), δ3.6 (ddd, 1 H, J = 6.5, 7.9 and 2.9 Hz, H-5), δ3.74 (dd, 1 H, J = 5.5 and 2.9 Hz, H-6'), δ4.13 (ddd, 1 H, J = 9.9, 3.9 and 1.9 Hz, H-3), δ4.22 (dd, 1 H, J = 9.9 and 4.9 Hz, H-1).

5.2.2 2-Deoxy-3,4,5,6-Tetrabenzyl-D-Arabino-Hexose Propylene Dithioacetal in Figure 4.17 compound 2 (6):

The 0.25 g (9.8×10^{-4} mole), 2-Deoxy-D-arabino-hexose propylene dithioacetal compound 1, 2.5 g barium oxide (0.0163 mole), and 15 mL dimethyl sulfoxide were placed in a 50 mL round bottom flask, then 1.8 mL (0.015 mole) benzyl bromide was added dropwise. The mixture was stirred at room temperature for 48 hours, then centrifuged, and the residue washed with chloroform. The supernatant and the chloroform extracts were combined. The solution was washed with a 15% aqueous solution of sodium thiosulfate, then washed with water, dried over sodium sulfate, and evaporated to dryness .

5.2.3 3-Deoxy-1-Iodo-4,5,6,7-Tetrabenzyl-D-Arabino-Heptose Propylene Dithioacetal in Figure 4.17 compound 3 (7):

A solution of 3×10^{-4} mole 2-deoxy-3,4,5,6-tetrabenzyl-D-arabino-hexose propylene dithioacetal compound 2 and 10 mL THF (distilled over lithium aluminum hydride) was placed in a 50 mL three-necked round-bottom flask. The system was protected under nitrogen and the mixture stirred in a carbon tetrachloride/ dry ice bath at -30° C. One mL 2.5 *M* n-butyllithium in hexane was injected at a rate of 3 mL/min. The mixture was stirred for 1.5 hours, and 0.1 mL (~ 0.012 mole) diiodomethane was injected. The mixture was protected from light with aluminum foil, stirred at -5° C (salted ice) for 18 hours, then quenched with two volumes of water, acidified with HCl (1 *M*) to pH 5-6, and extracted with 3 portions (10 mL each) of ethyl acetate. After washing with 3% aqueous sodium bisulfite and water, the sample was dried with sodium bisulfite.

5.2.4 3-Deoxy-4,5,6,7-Tetrabenzyl-D-Arabino-Heptulose Propylene Dithioacetal in Figure 4.17 compound 4 (8):

Approximately 3 x 10^{-4} mole 3-deoxy-1-iodo-4,5,6,7-tetrabenzyl-D-arabino-heptose propylene dithioacetal compound 3 and 10 mL THF were placed in a 50 mL round-bottom flask, then 1 mL 1 *M* tetrabutyl ammonium hydroxide was added. The mixture was allowed to stay at room temperature for 20 hours, then 7 mL 1% HCl in methanol was added. The mixture was washed with 5% sodium bicarbonate, then extracted with ethyl acetate and dried by passage over absorbant cotton.

5.2.5 3-Deoxy-4,5,6,7-Tetrabenzyl-D-Arabino-Heptulose in Figure 4.17 compound 5 (9, 10):

Approximately 2.5 x 10^{-4} mole 3-Deoxy-4,5,6,7-tetrabenzyl-D-arabino-heptulose propylene dithioacetal compound 4 and 10 mL acetonitrile/water (4/1) were placed in a 50 mL round bottom flask, and 0.07 g calcium carbonate (2.5 eq.), 0.17 g mercuric chloride (2.3 eq.) were added. The mixture was stirred at room temperature for 4 hours, filtered through celite, washed with ammonium acetate, and the aqueous layer extracted with ethyl acetate. The organic layers were combined and washed with odium bicarbonate, filtered with cotton and potassium carbonate, and evaporated under nitrogen to dryness.

5.2.6 3-Deoxy-2-Deutero-4,5,6,7-Tetrabenzyl-D-Gluco(Manno)-Heptitol in Figure 4.17 compound 6:

Approximately 2.5 x 10^{-4} mole 3-deoxy-4,5,6,7-tetrabenzyl-D-arabino-heptulose compound 5 was treated with 0.2 g sodium borodeuteride in

ethanol. The mixture was stirred overnight, quenched with 0.1*N* HCl, extracted with ethyl acetate, and dried by passage over cotton.

5.2.7 3-Deoxy-2-Deutero-D-Gluco(Manno)-Heptitol in Figure 4.17 compound 7 (11):

Approximately 2.0 x 10^{-4} mole 3-deoxy-2-deutero-4,5,6,7-tetrabenzyl-D-gluco(manno)-heptitol compound 6, 0.05 g 10% palladium chloride on active charcoal, 5 mL cyclohexene, and 10 mL ethanol were placed in a 50 mL round-bottom flask which was stirred and heated at 70° C overnight. After cooling to room temperature, the palladium catalyst was filtered off and the filtrate washed with 0.1 *M* ammonium acetate and water, and then concentrated to dryness on the rotary evaporator.

5.2.8 3-Deoxy-2-Deutero-1,2,4,5,6,7-Hexacetyl-D-Gluco(Manno)-Heptitol in Figure 4.17 compound 8 (12):

Approximately 1.5 x 10^{-4} mole 3-deoxy-2-deutero-D-gluco(manno)-heptitol compound 7, 2 mL pyridine, and 3 mL acetic anhydride were placed in a 25 mL round bottom flask. The mixture was stirred and heated at 80° C for 3 hours. After cooling to room temperature, 5 mL 1 *M* HCl was added. The mixture was extracted with chloroform, washed with sodium bicarbonate, and dried with sodium sulfate.

5.2.9 2-Deoxy-3,4 : 5,6-Di-*O*-Isopropylidene-D-Arabino-Hexose Propylene Dithioacetal in Figure 4.20 compound 9 (5):

The 2-Deoxy-D-arabino-hexose propylene dithioacetal compound 1 0.5 g, (0.002 mole) was dissolved in acetone (30 mL) in a 100 mL round-bottom flask, and concentrated sulfuric acid (2 drops) was added. The mixture

was allowed to stand at 50° C for 12 hours, then concentrated to dryness. The yield was 100% of crude product. After recrystallization from methanol, the yield of final product was 81%. MP: 103-104° C, ^{1}H-NMR (500 MHz, $CDCl_3$): δ1.30 (s, 3 H, C-CH_3), δ1.34 (s, 3 H, C-CH_3), δ1.36 (s, 3 H, C-CH_3), δ1.40 (s, 3 H, C-CH_3), δ1.86 (m, 1 H, w/44.9 Hz, H-8), δ1.94 (ddd, 1 H, J = 15, 9.9 and 4.9 Hz, H-2'), δ2.08 (m, 1 H, w/34.9 Hz, H-8'), δ2.17 (ddd, 1 H, J = 15, 9.9 and 4.9 Hz, H-2), δ2.83 (m, 2 H, w/25 Hz, H-7' x 2), δ2.90 (m, 2 H, w/ 40Hz, H-7 x 2), δ3.52 (t, 1 H, J = 15 Hz, H-4), δ3.89 (dd, 1 H, J = 5 and 9.9 Hz, H-6), δ4.0 (dt, 1 H, J = 9.9, and 4.9 Hz, H-5), δ4.09 (dd, 1 H, J = 5 and 9.9 Hz, H-6'), δ4.19 (ddd, 1 H, J = 9.9, 3.9 and 4.9 Hz, H-3), δ4.24 (dd, 1 H, J = 9.9 and 4.9 Hz, H-1). ^{13}C-NMR (300 MHz,$CDCl_3$): δ 25.3, 26.0, 26.4, 26.6, 26.8, 29.8, 30.2, 39.8, 42.0, 66.8, 76.7, 76.9, 81.3, 109.3, 109.6. IR: (CM^{-1}): 2800-3000 (C-H), 1400 (C-CH_3), 1200, 1100 (C-O), 700 (C-S).

5.2.10 3-Deoxy-1-Iodo-4,5: 6,7-Di-*O*-Isopropylidene-D-Arabino-Heptose Propylene Dithioacetal in Figure 4.20 compound 10 (7):

Approximately 3 x 10^{-4} mole 2-deoxy-3, 4: 5, 6-di-*O*-isopropylidene-D-arabino-hexose propylene dithioacetal compound 9 and 10 mL THF (distilled over lithium alumina hydride) was placed in a 50 mL three-neck round-bottom flask and a nitrogen inlet attached. One mL 2.5 *M* n-butyllithium in hexane was injected at -30° C, at a rate of 3 mL/min. The mixture was stirred for 2 hours, before 0.1 mL (~ 0.012 mole) diiodomethane was injected. The mixture was protected from light by aluminum foil and stirred at 0° C for 8 hours. The reaction was then quenched by two volumes of water, acidified with HCl (1 *M*) to pH 5-6 and extracted with 3 portions (10 mL each) of ethyl acetate. The organic layer was washed with 3% aqueous sodium bisulfite and water, and then dried with sodium bisulfite. Product

yield was 41%. ^{1}H- NMR (300 MHz, $CDCl_3$): δ1.30 (s, 3 H, C-CH_3), δ1.34 (s, 3 H, C-CH_3), δ1.36 (s, 3 H, C-CH_3), δ1.40 (s, 3 H, C-CH_3), δ2.13 (quintet, 2 H, J = 9 Hz, H-8, 8'), δ2.39 (dd, 1 H, J = 12 and 18 Hz, H-2'), δ2.52 (dd, 1 H, J = 15 and 3 Hz, H-2), δ3.38 (m, 4 H, w/36 Hz, H-7' x 2, H-7 x 2), δ3.6 (t, 1 H, J = 15 Hz, H-4), δ3.92 (dd, 1 H, J = 5 and 9.9 Hz, H-6), δ4.0-δ4.12 (m, 3 H, w/36 Hz, H-6', H3, H5), δ5.9 (s, 1 H, H-1, CH_2I). ^{13}C NMR (300 MHz, $CDCl_3$): δ25.3, 26.8, 27.2 27.6, 31.2, 31.8, 42.1, 67.2, 76.9, 78.8, 80.6, 109.1, 109.7, 110.8, 120.6, 130.24. IR: (CM^{-1}): 3000 (C-H), 2400 (C-H, alkane monosubstitute), 1400 (C-CH_3), 1500 (C-H, methylene), 1200, 1000 (C-O, C-I), 700 (C-S). MS: (JEOL JMS-AX505H): *m/z* 473 (M^+), *m/z* 346 (lost iodide), *m/z* 331, *m/z* 143.

5.2.11 4,6-*O*-Propylidene-D-Glucose in Figure 4.21 compound 13 (14):

Approximately 30 g (0.17 mole) D-glucose and 70 mL propionaldehyde was added to a 250 mL round-bottom flask equipped with a condenser and drying tube. Concentrated sulfuric acid 0.4 mL was added dropwise with vigorous stirring, and the mixture was kept at 40° C overnight. After cooling to room temperature, 20 mL ethanol was added. The pH was adjusted to 7 by adding saturated sodium carbonate, then an additional 80 mL ethanol was added. The product was obtained as crystals with a 21% yield. ^{1}H-NMR (500 MHz, CD_3OD): δ0.72 (t, 3 H, CH_3), δ1.58 (m, 2 H, CH_2), δ3.22 (t, 1 H, C-2), δ3.46 (m, 1 H, C-3), δ3.57 (m, 1 H, C-4), δ3.75 (m, 1 H, C-5), δ4.05 (dd, 1 H, C-6), δ4.05 (dd, 1 H, C-6), δ4.6 (m, 1 H, propylidene, possibly overlap with β-anomeric proton), δ5.16 (d, 1 H, α-anomeric proton).

5.2.12 2,4-*O*-Propylidene-Erythrose in Figure 4.21 compound 14 (15):

Approximately 1.67 g (0.0076 mole) 4,6-*O*-propylidene-D-glucose compound 13 was dissolved in 50 mL water along with sodium bicarbonate,

1.7 g (0.02 mole, 2.7 eq.) in a 500 mL round-bottom flask. Subsequently, potassium meta-periodate (3.8 eq., 0.017 mole) previously dissolved in 200 mL water was also added. The reaction mixture was stirred at 40° C for 3 hours, extracted with hot ethyl acetate, and after concentrating to dryness, 0.9 g product was obtained. The yield was 75%. TLC condition was ethyl acetate/hexane 4/1. ^{1}H-NMR (500 MHz, $CDCl_3$): δ0.82 (t, 6 H, 2 x CH_3), δ1.55 (m, 4 H, 2 x CH_2), δ3.25 (m, 4 H, C-4), δ3.55 (m, 2 H, 2 x C-3), δ3.95 (dd, 2 H, 2 x C-2), δ4.59 (d, 1 H, C-1), δ4.65 (d, 1 H C-1'). ^{13}C NMR (300 MHz, $CDCl_3$): δ8.8 (2 x CH_3), δ28.5 (2 x CH_2), δ62.8 (C-2), δ62.9 (C-2'), δ71.6 (C-3), δ71.3 (C-3'), δ83.1 (C-4), δ83.4 (C-4'), δ97.5 (CH propylidene), δ97.8 (C'H, propylidene), δ103.8 (C-1), δ104.2 (C-1').

5.2.13 2,4-*O*-Propylidene-D-Erythrose Propylene Dithioacetal in Figure 4.21 compound 15 (16):

Approximately 0.53 g (0.0033 mole) 2,4-propylidene-D-erythrose compound 14 dissolved in 50 mL 1,4-dioxane, and 0.36 mL (3.6 mmol, 1.1 eq.) 1,3-propanedithiol were added to a 100 mL round-bottom flask and cooled to -78° C; 0.7 mL (2 eq.) 1 *M* HCl etherate solution was then added. The mixture was stirred at room temperature for 3 hours, and the solvent was removed via rotary evaporation. The product was purified by silica gel chromatography (ethyl acetate/hexane 1:1). The product yield was 0.74 g (90% yield). ^{1}H-NMR (500 MHz, $CDCl_3$): δ0.94 (t, 3 H, CH_3), δ1.65 (m, 2 H, CH_2), δ2.0 (m, 2 H, -$SCH_2C\underline{H_2}CH_2S$-), δ2.8 (m, 2 H, axial -$SC\underline{H_2}CH_2C\underline{H_2}S$-), δ3.0 (m, 2 H, equatorial -$SC\underline{H_2}CH_2C\underline{H_2}S$-), δ3.4 (dd, 1 H, C-4), δ3.7 (dd, 1 H C-2), δ3.9 (m, 1 H, C-3), δ4.1 (dd, 1 H, C-4'), δ4.25 (d, 1 H, C-1), δ4.45 (t, 1 H, propylidene). ^{13}C NMR (300 MHz, $CDCl_3$): δ8.8 (CH_3), δ27 (CH_2), δ26

($SCH_2\underline{C}H_2CH_2S$-), δ29 (-$S\underline{C}H_2CH_2\underline{C}H_2S$-), δ46 (C-3), δ64 (C-4'), δ70 (C-2), δ84.5 (propylidene), δ103.4 (C-1).

5.2.14 2,4-*O*-Propylidene-3-*O*-Acetyl-D-Erythrose Propylene Dithioacetal

in Figure 4.21 compound 16:

Approximately 0.37 g (1.48 mmole) 2,4-*O*-propylidene D-erythrose propylene dithioacetal was placed in a one dram vial, then 0.5 mL pyridine, 0.5 mL acetic anhydride was added. The solution was left at room temperature for 12 hours before the solvent was blown to dryness under a stream of nitrogen. The TLC eluant was hexane/ethyl acetate 4/1. After purification by silica gel chromatography, the yield was 63%. ^{1}H-NMR (500 MHz, $CDCl_3$): δ0.94 (t, 3 H, CH_3, propylidene), δ1.68 (m, 2 H, CH_2), δ2.0 (s, 3 H, CH_3, acetyl), δ2.6 (m, 1 H, axial -$SCH_2C\underline{H}_2CH_2S$-), δ2.7 (m, 1 H, equatorial -$SCH_2C\underline{H}_2CH_2S$-), δ2.8-2.9 (m, 2 H, axial -$SC\underline{H}_2CH_2C\underline{H}_2S$-), δ3.0-3.1 (m, 2 H equatorial -$SC\underline{H}_2CH_2C\underline{H}_2S$-), δ3.36 (dd, 1 H, C-4), δ3.94 (dd, 1 H C-2), δ3.97 (d, 1 H, C-1), δ4.28 (dd, 1 H, C-4'), δ4.48 (t, 1 H, CH, propylidene), δ5.0 (m, 1 H, C-3). ^{13}C NMR (80 MHz, $CDCl_3$): δ8.8 (CH_3), δ21 (acetyl,CH_3), δ25.6 (), δ27.4 (CH_2, propylidene), δ28.9 ($SCH_2\underline{C}H_2CH_2S$), δ29.1 ($S\underline{C}H_2CH_2\underline{C}H_2S$), δ44.6 (C-3), δ64.5 (C-4'), δ67.2 (C-2), δ84.4 (CH, propylidene), δ103.4 (C-1), δ169.6 (carbonyl).

5.3 References

1. Chaplin M. F. and Kennedy, J. F., *"Carbohydrate Analysis"* (1988) 182, Practical Approach Series, IRL Press, Oxford, Washington, DC.
2. Hollingsworth, R. I., Abe, M., Sherwood, J. E. and Dazzo, F. B. *J. Bacteriol.* (1984) 160 510-516.

3. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F., *Anal. Chem.* (**1956**) 28 350.

4. Waravdekar, V. S. and Saslaw, L. D., *J. Biol. Chem.* (**1959**) 234 1945.

5. Frost J. W. and Knowles, J. R., *Biochemistry* (**1984**) 23 4465-4469.

6. Srivastava, H. C., Harshe, S. N., and Singh, P. P., *Tet. Let.* (**1963**) 27 1869-1873.

7. Corey E. J. and Seebach, D., *Angew. Chem. Internal. Ed. Enl.* (**1965**) 4 1075-1077.

8. Numazawa, M., and Nagaoka, M., *Steroids* (**1982**) 39 345-355.

9. Corey, E. J., and Bock, M. G., *Tet. Let.* (**1975**) 38 3269-3270.

10. Meyers, A. I., Comins, D. L., Henning, D. M. R., Shimizu, K., *J. Am. Chem. Soc.*, (**1979**) 101 7104-7105.

11. Anantharamaiah, G. M., and Sivanandaiah, K. M., *J. Chem. Soc. Perkin I* (**1977**) 490-491.

12. Zhdanov, R. I., and Zhenodarova, S. M., *Synthesis* (**1975**) 222.

13. Ph.D. thesis of Dr. Hollingsworth, R. I., p. 96.

14. Ph.D. thesis of Dr. Hollingsworth, R. I., p. 164.

15. Bonner, T. G., Bourne, E. J., and Lewis, D., *J. Am. Chem. Soc.* (**1965**) 87 7453-7458.

16. Baggett, N., Buck, K. W., Foster, A. B., Rees, B. H., and Webber, J. M., *J. Chem. Soc. (C)*, (**1966**) 212-215.

17. Schmidt, U., and Ohler, E., *Angew. Chem. Int. Ed. Enl.* (**1977**) 16 327.

CHAPTER 6

CHARACTERIZATION OF STRUCTURAL DEFECTS IN THE LIPOPOLYSACCHARIDE OF SYMBIOTICALLY IMPAIRED *RHIZOBIUM LEGUMINOSARUM* BIOVAR *VICIAE* VF-39 MUTANTS

6.1 Abstract

The lipopolysaccharides (LPS) of a wild-type strain of *Rhizobium leguminosarum* biovar *viciae* (strain VF-39) and two symbiotically defective Tn5 mutants (VF-39-32 and VF-39-86) have been studied. The LPS of the mutants reflect impaired synthesis of the *O*-antigen. In the LPS of one mutant, the core tetrasaccharide was lacking and in that of the other it was truncated to a disaccharide containing mannose and 3-deoxy-D-manno-oct-2-ulosonic acid (KDO). The latter mutant also synthesized an unusual carbohydrate component containing mannose, galactose, and an unidentified saccharide. The lipid A composition was similar to that found in other strains of *R. leguminosarum* biovar *viciae*. The *O*-antigen of the wild-type bacterium contained 2-*O*-methylfucose, fucose, 3,6-dideoxy-3-(methylamino)hexose, glucose, 2-amino-2,6-dideoxyhexose and heptose. This study clearly defines a role for the bacterial LPS in the proper functioning of the *Rhizobium* legume symbiosis.

6.2 Introduction

The Rhizobiaceae are a family of Gram-negative bacteria which infect and form a symbiosis with legume plants with consequent conversion or fixation of molecular nitrogen into ammonia. This process is marked by a high level of specificity between the bacteria and the plant, and cell-

surface carbohydrates are thought to be involved in this process at the levels of initial recognition (1, 2) and nodule development (3-6). It has been demonstrated (3) that mutants of *R. japonicum* which appeared to be defective in LPS biosynthesis were incapable of forming nitrogen-fixing nodules in plants. This phenomenon was also demonstrated with *R. phaseoli* mutants (4) and for *R. leguminosarum* biovar *viciae* mutants (5). Whereas these three studies lacked chemical analysis of the LPS structure, this aspect was provided in a study of symbiotically-defective *R. phaseoli* mutants (6). The work now reported extends the results of this study to the *R. leguminosarum* system and underscores the importance of the synthesis of an intact LPS by the bacterium for viable symbiosis.

The rhizobial LPS have unusual features which include transmembrane fatty acids (7,8), unusual sugars including uronic acids in the lipid A (9, 10), and for some LPS, a complete lack of phosphate in the lipid A (10). There is a large degree of structural diversity from strain to strain among the different species (11), and complete structures are known only for two core oligosaccharide components (12-14), namely a trisaccharide containing KDO and two GalA residues and a tetrasaccharide containing KDO, Gal, Man, and GalA residues.

R. leguminosarum biovar *viciae* (strain VF-39), isolated from nodules of *Vicia faba,* can develop an efficient symbiosis with host plants such as *V. hirsuta* and *Pisum sativum.* Mutants of this strain, isolated by Tn5 mutagenesis, were impaired in their ability to elicit normal development of nodules. They were able to induce the formation of nodules and infection threads, but were affected in the release and/or differentiation of the bacteria (5). SDS-PAGE indicates that the symbiotic defect of these mutants is correlated with a failure to produce a normal LPS. This defect

was also found for two mutants (including VF-39-32) by preliminary analysis of the LPS carbohydrate (15). The present study was undertaken in order to determine the modified structures of the LPS of the genetically-different mutants, VF-39-32 and VF-39-86, so as to better understand the origin of the changes and to rationalize their effects on symbiosis events.

6.3 Experimental

6.3.1 Culture of Bacteria and Isolation of LPS and LPS Fragments

Bacteria were grown in a modified Bergensens medium (as described in 16) in shaken-broth cultures. Mutants were grown (5) in the presence of antibiotics. The LPS were isolated by the hot-phenol method (17) and after treatment with DNAse and RNAse, purified on Sepharose 4B using an ammonium formate buffer (9, 10). Column eluates were analyzed by the phenol-sulphuric acid method (18). Fragments were released from the LPS by hydrolysis for 2 h with aqueous 1% acetic acid at 100° C. The hydrolysate was extracted with 5:1 chloroform-methanol in order to remove lipid A, the aqueous layer was lyophilized, and the residue was chromatographed on a column of Biogel P2 using aqueous 0.1% formic acid. Fractions which appeared impure by ^{1}H-NMR spectroscopy and other analyses were subjected to ion-exchange chromatography on DEAE-Sephadex G25 in aqueous 0.01% formic acid adjusted to pH 6.5 with ammonia, eluting with a linear gradient of 0-0.2 *M* ammonium chloride over two hundred mL.

6.3.2 Compositional Analysis

Analyses for fatty acids and carbohydrates in lipid A were carried out as described (7-10) after conversion into methyl esters and alditol acetates, respectively. Deuterium-labeled acetate groups were introduced by

acylation with acetic anhydride-d_6. Gas chromatography of carbohydrates was carried out on a Hewlett Packard 5890 gas chromatograph using a capillary DB225 column and a flame-ionization detector. Mass spectrometry detection was effected with a JEOL 505 mass spectrometer in the EI or CI (ammonia) mode.

6.3.3 Structural Analyses

Oligosaccharides were methylated after reduction of carboxyl and other carbonyl functions with sodium borodeuteride (as described in 12). ^{1}H-NMR spectroscopy (500-MHz) was performed on solutions in D_2O with a Varian VXR500 spectrometer. FAB-MS was performed with a JEOL HX110-HF instrument in the negative- or positive-ion mode and using a Cs or Xe source.

6.4 Results and Discussion

Chromatography on Sepharose 4B of the LPS from the wild-type strain *Rhizobium leguminosarum* biovar *viciae* VF-39 (Figure 6.1) showed four major fractions of which A-C corresponded to LPS, and the late-eluting peak was composed almost entirely of a glucan. Hydrolysis of each of the fractions A-C with 1% acetic acid followed by extraction of the lipid A with chloroform-methanol (as described under general methods) and chromatography of the components in the aqueous phase on Biogel P2 gave three sub-fractions, and these three profiles were similar (Figure 6.1).

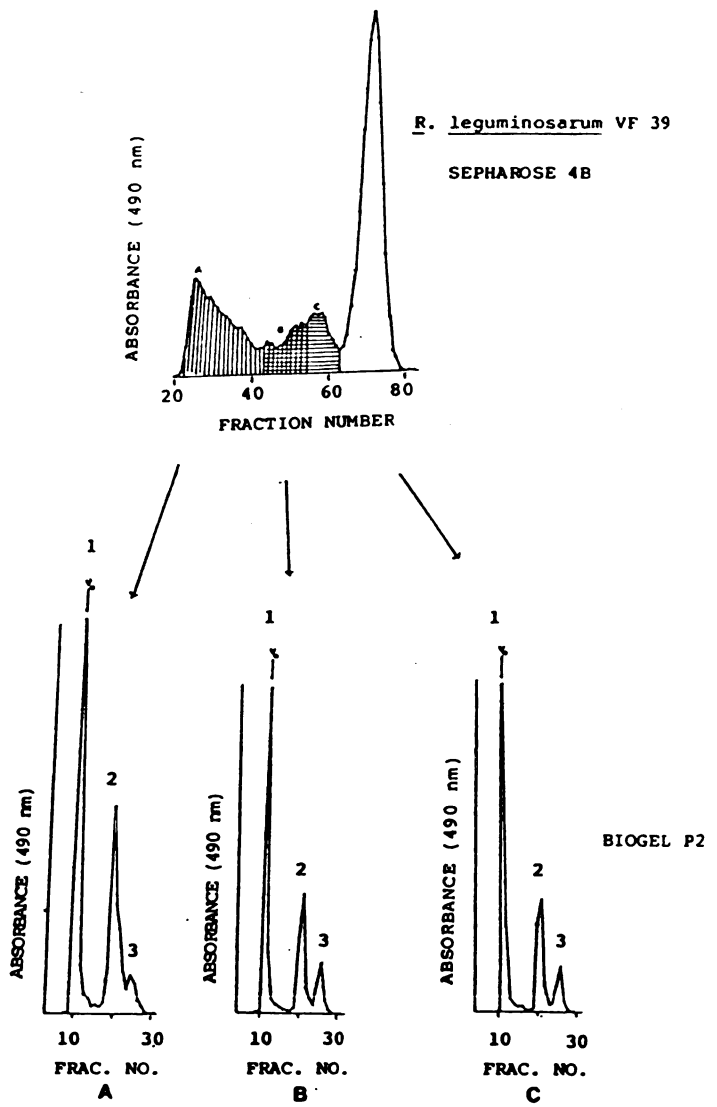


Figure 6.1 Gel filtration on Sepharose 4B of the crude LPS isolated from *Rhizobium leguminosarum* biovar *viciae* VF-39. Fractions A-C were each subjected to mild hydrolysis with 1% acetic acid at 100° C for 2 hours and then extracted with 5:1 chloroform-methanol and the aqueous layer then subjected to gel filtration chromatography on Biogel P2. Peaks with the same numbers had similar compositions and NMR spectra, and contained the *O*-antigen, a tetrasaccharide 1 and a trisaccharide 2. The vertical axes represent absorbance in the phenol-sulphuric acid assay.

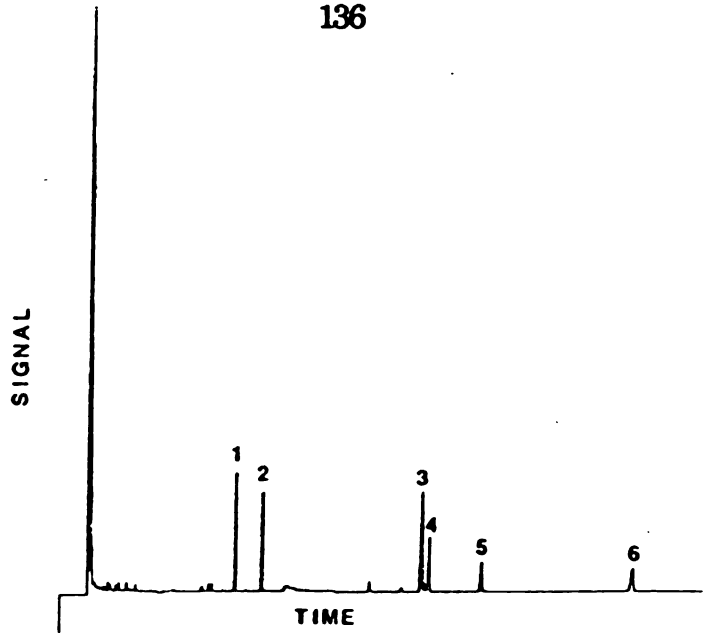


Figure 6.2 Gas chromatogram (see Experimental) of *O*-antigen carbohydrate components of *Rhizobium leguminosarum* biovar *viciae* VF-39 as the alditol acetates derived from 2-*O*-methylfucose, fucose, glucose, 2-amino-2,6-dideoxyglucose, 2-amino-2,6-dideoxyhexose, and heptose in the order 1-6, respectively.

GC-MS indicated that peak 1 contained 2-*O*-methyl fucose, fucose, a 3,6-dideoxy-3-(methylamino) hexose (probably the *galacto* isomer (19)), glucose, amino-2,6-dideoxyhexose and heptose (Figure 6.2). The presence of most of these components was also indicated by ^{1}H-NMR spectroscopy which showed that the *O*-antigen was heavily acetylated (signals between 1.8 and 2.2 ppm; Figure 6. 3). There were signals at 0.8-1.4 (CHMe), 2.75 (NMe), ~3.28 ppm (OMe). Peak 2 contained the deacetylated version of a tetrasaccharide made up of Man, Gal, GalA and KDO first characterized in *R trifolii* ANU843 (12) (the linkage of mannose to KDO was later revised (13)). The composition and linkage positions were confirmed by methylation analysis and the molecular weight by FAB-MS, The structure is shown in Figure 6.4a, and the ^{1}H-NMR spectrum, which is identical to that of the *R trifolii* tetrasaccharide after deacetylation, is shown in Figure

6.4b Peak 3 contained the trisaccharide **2** identified (6, 13, 14) in *R trifolii* ANU843.

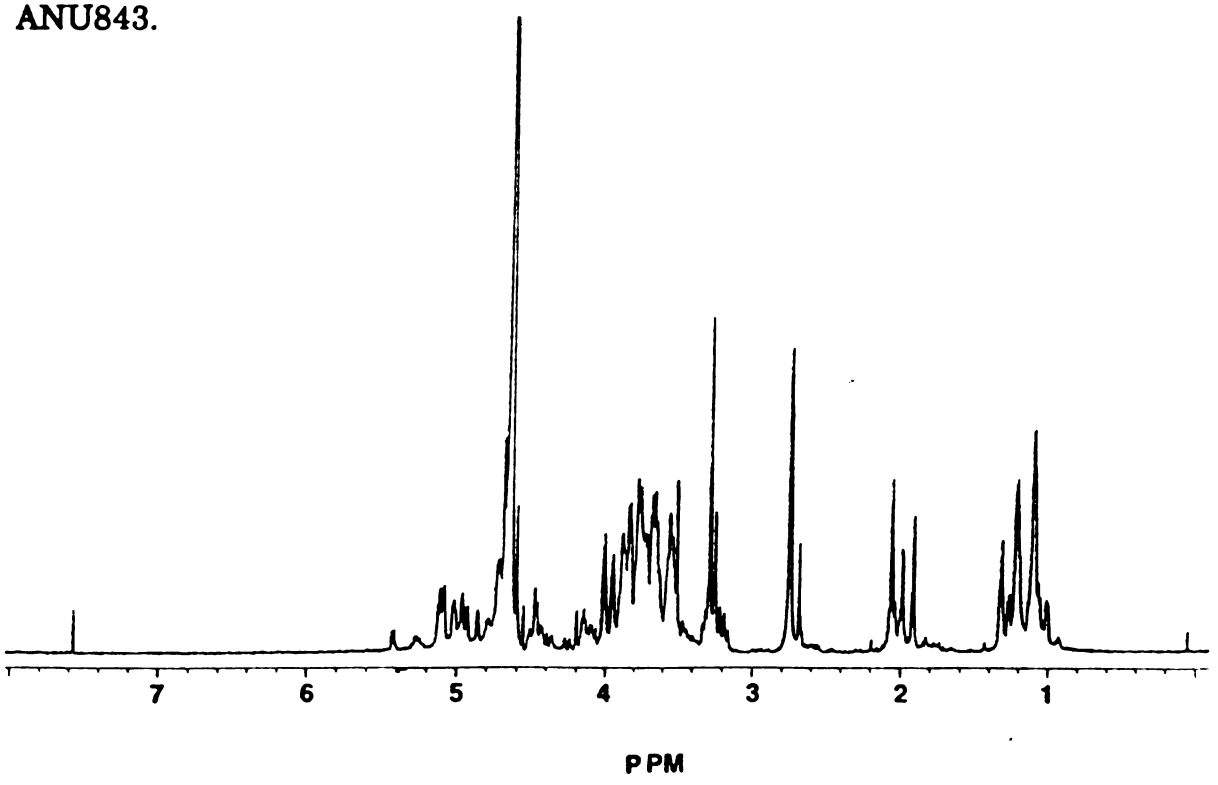


Figure 6.3 500-MHz ^{1}H-NMR spectrum of *O*-antigen (peak 1 in Figure 6.1).

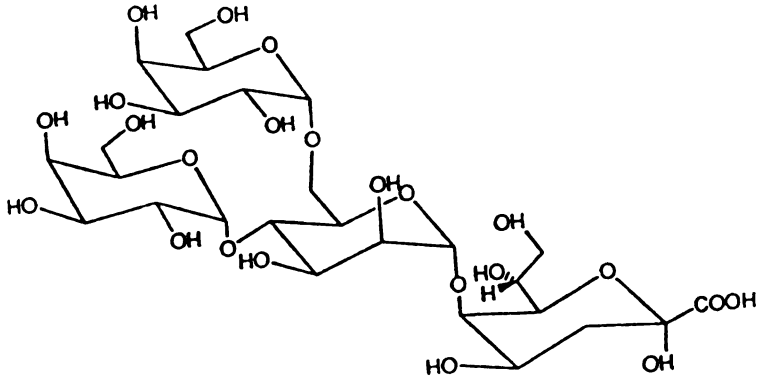


Figure 6.4a: The structure of the tetrasaccharide **1**.

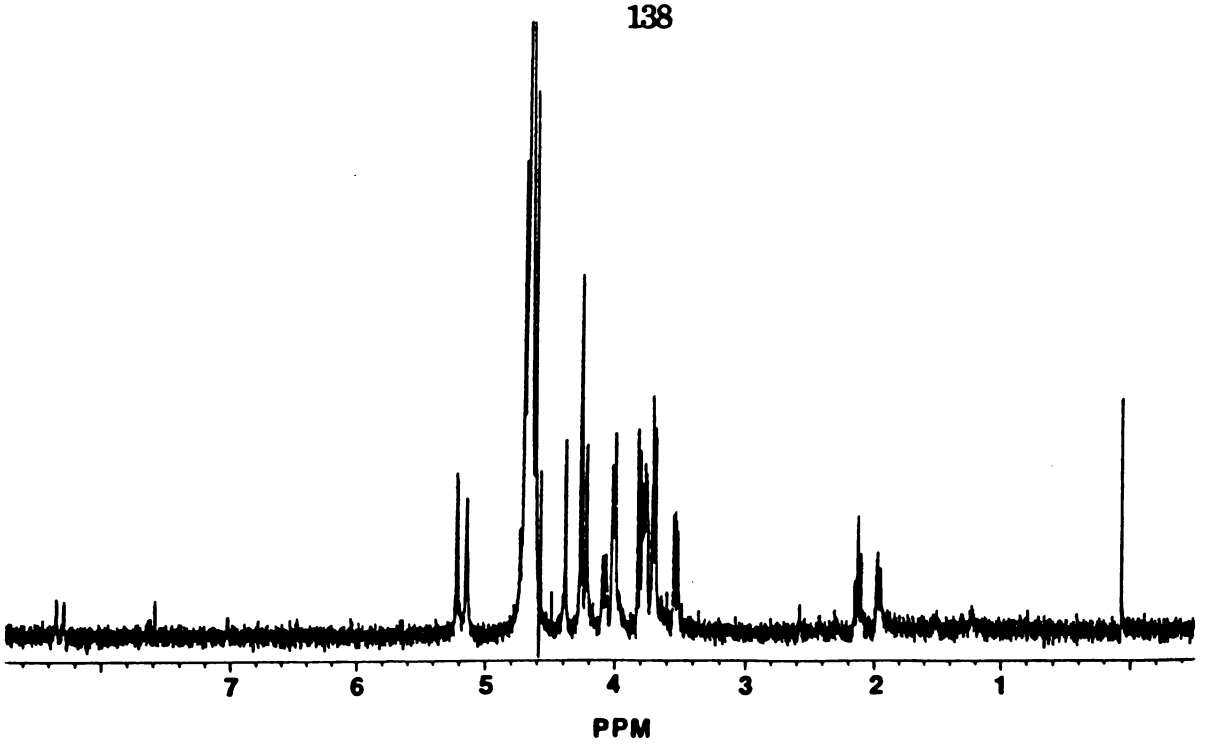


Figure 6.4b ^{1}H-NMR spectrum of the tetrasaccharide shown in Figure 6.4a. The three H-1 signals in the range 4.8-5.4 ppm and are due to α-Gal, Man and Gal, respectively, in decreasing order of chemical shift. Signals for the poorly resolved deoxy function appear at ~2 ppm.

The gel-filtration profile on Sepharose 4B of the total LPS of the mutant VF- 39-32 was radically different from that of the wild type (Figure 6.5). Two major LPS fractions (A and B) which after hydrolysis (with 1% acetic acid as described for the parent strain) gave profiles on Biogel P-2 that were different from those obtained (Figure 6.1) from the wild type. Thus, peak 1 from fraction A contained none of the amino sugars found in the wild type and appeared to contain only a small amount of a capsular antigen comprising mainly glucose. Peaks 2 and 3 from fractions A and B contained similar labeled major components. Peak 1 from fraction B contained small amounts of the *O*-antigen found in the parent strain. Peak 3 contained the trisaccharide **2**. The major peak contained a component,

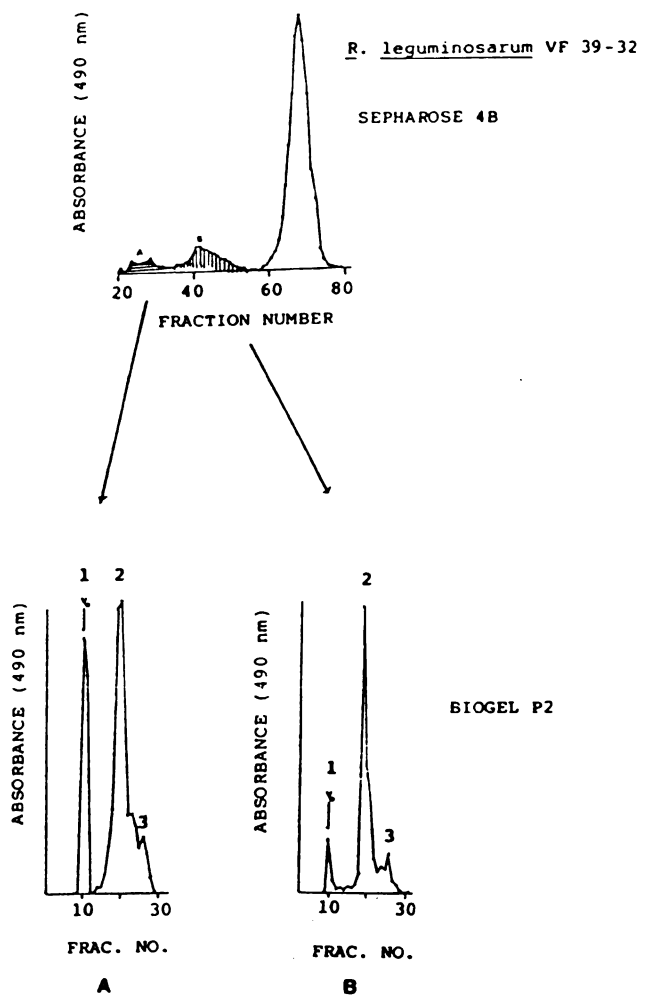


Figure 6.5 Gel filtration on Sepharose 4B of the crude LPS of *Rhizobium leguminosarum* biovar *viciae* VF-39-32. Fractions A and B were each subjected to mild-acid hydrolysis with 1% acetic acid at 100° C for 2 hours and then extracted with 5:1 chloroform-methanol and the aqueous layer then loaded to gel filtration chromatography on Biogel P2. Peak 1 from fraction A, which contained mainly glucose, indicating it might be due to small amounts of capsular material, was devoid of amino sugars. Peak 1 from fraction B contained a small amount of *O*-antigen. Peak 3 from fraction A and B contained the trisaccharide 2, and peak 2 contained an unidentified product.

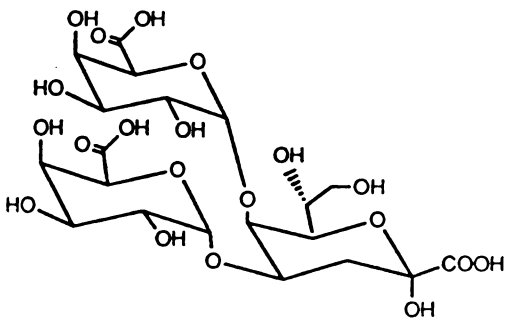


Figure 6.6a: The structure of the trisaccharide **2**.

^{1}H-NMR spectrum is shown in Figure 6.6b. Two sets of up-field resonances dd at 2.02 (J = 14.5 and 1 Hz) and 2.60 ppm (J = 14.5 and 7.6Hz) indicated the presence of a deoxy sugar. GC (EI and CI)-MS of alditol acetate and trideuteroacetate from this residue, confirmed the presence of Gal and Man and indicated the presence of a 2-deoxyaldose. The signal at 4.53 ppm (dd) was coupled to the methylene group, and the chemical shift, being due to an anomeric proton, is far down-field for H-4 of KDO. The structure of this molecule is still under investigation, but it does not correspond (6) to a truncated form of the tetrasaccharide **1** in Figure 6.4a.

The gel-filtration elution profile on Sepharose 4B of the LPS of the mutant, VF-39-86 was also different from that of the wild type (Figure 6.7). The major fractions A-C, when subjected to mild acid hydrolysis followed by gel filtration on Biogel P2, yielded profiles (Figure 6.7) which were different to those obtained from the parent strain. Thus, peak A yielded three peaks of which peak 1 corresponded to traces of capsular material but none of the components present in the *O*-antigen of the parent strain (Figure 6.2). The quantity of material isolated was too small for ^{1}H-NMR

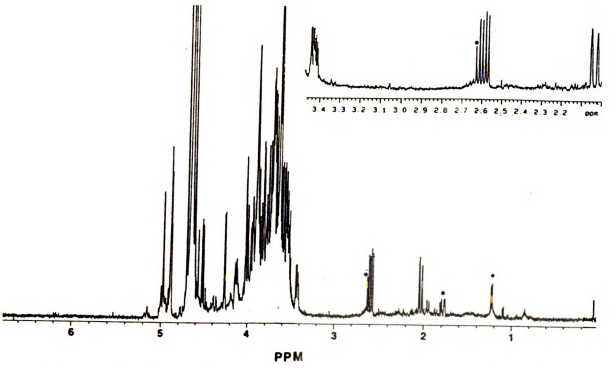


Figure 6.6b ^{1}H-NMR spectrum of unidentified component from strain VF-39-32 (peak 2 in Figure 6. 5). The peaks labelled with asterisks are due to contaminants.

spectroscopy. Peaks 2 and 3 also contained very small amounts of material, and their ^{1}H-NMR spectra contained peaks down-field of 6 ppm indicating that they may contain nucleic acid. Fraction B gave several small peaks each of which contained carbohydrate in quantities too small for study. However, the major peak (4) from fractions B and C contained three aldopyranoside components, as indicated by the presence of the three H-1 resonances between 4.8 and 5.2 ppm in the ^{1}H-NMR spectrum (Figure 6.8). The ion-exchange chromatography of this material on DEAE-Sephadex yielded two components; one of them had a ^{1}H-NMR spectrum (Figure 6.9a) that contained only one H-1 resonance and which corresponded to a

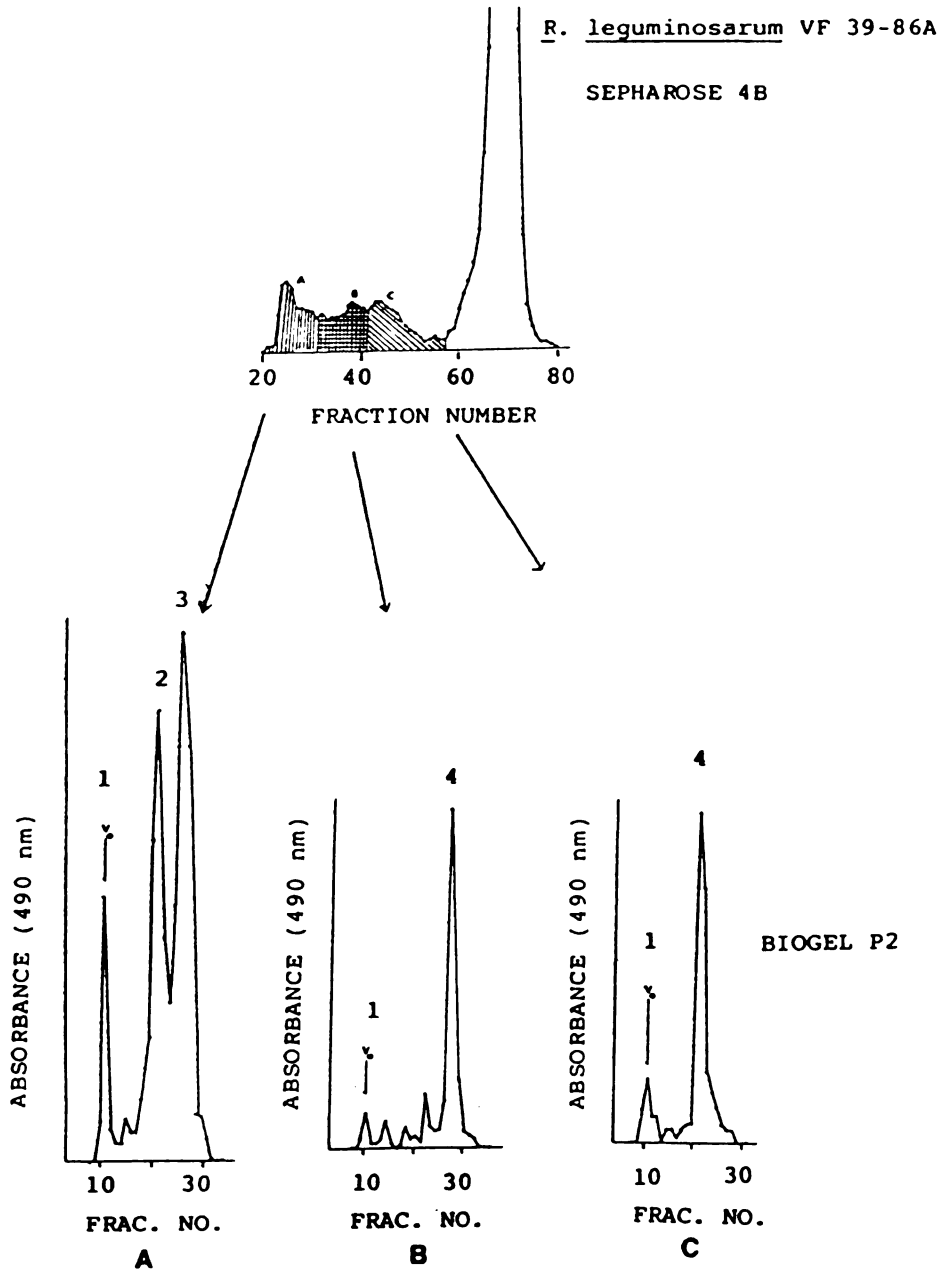


Figure 6.7 Gel filtration elution profile on Sepharose 4B of the crude LPS isolated from *Rhizobium leguminosarum* biovar *viciae* VF-39-86. Fractions A-C were each subjected to mild-acid hydrolysis then to gel filtration on Biogel P2 as described for the parent strain and strain VF-39-32. Fraction A yielded small amounts of capsular material (peak 1) and possibly nucleic acid material (peaks 2 and 3). Fractions B and C gave one major component (peak 4).

disaccharide (Figure 6.9b) of Man and KDO. The same molecule was obtained (6) from the LPS of a *R. phaseoli* mutant and shown by NMR spectroscopy and methylation analysis to be α-Man-(1-5)-KDO [3]. The molecular weight of this component was confirmed by FAB-MS of m/z 445 for the sodium adduct of sodium carbonated molecule $[M+Na]^+$ (Figure 6.10). The fatty acids found in the lipid A of the parent strain and both mutants were the same and in the same proportions as those found in *Rhizobium leguminosarum* biovar *trifolii* ANU843 (7, 8), namely C_{14}, C_{15}, C_{16} and C_{18} 3-hydroxy acids and 27-hydroxyoctacosanoic acid.

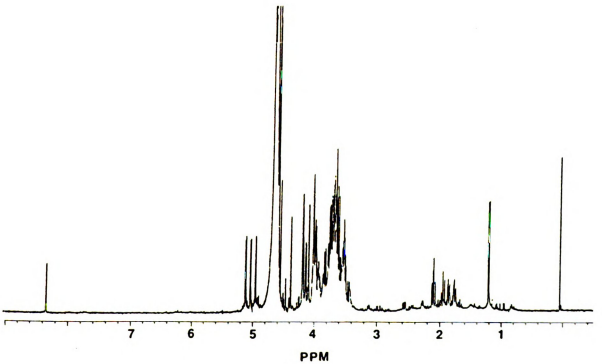


Figure 6.8 ^{1}H-NMR spectrum of the material in peak 4 of Figure 6.7.

The results obtained for the symbiotically deficient mutants VF-39-32 and VF-39-86A of *Rhizobium leguminosarum* indicate that (a) the ability of the bacterium to synthesize an intact LPS, containing the *O*-antigen is necessary for proper symbiosis, (b) the synthesis of the core trisaccharide **2** component is independent of the synthesis of the tetrasaccharide **1**, and (c) the bacterium can synthesize more than one type of LPS and impairment or suppression of the synthesis of the usual LPS can lead to the synthesis of an alternative LPS which is not normally expressed. This work underlines the importance of bacterial cell surface polysaccharides in infection.

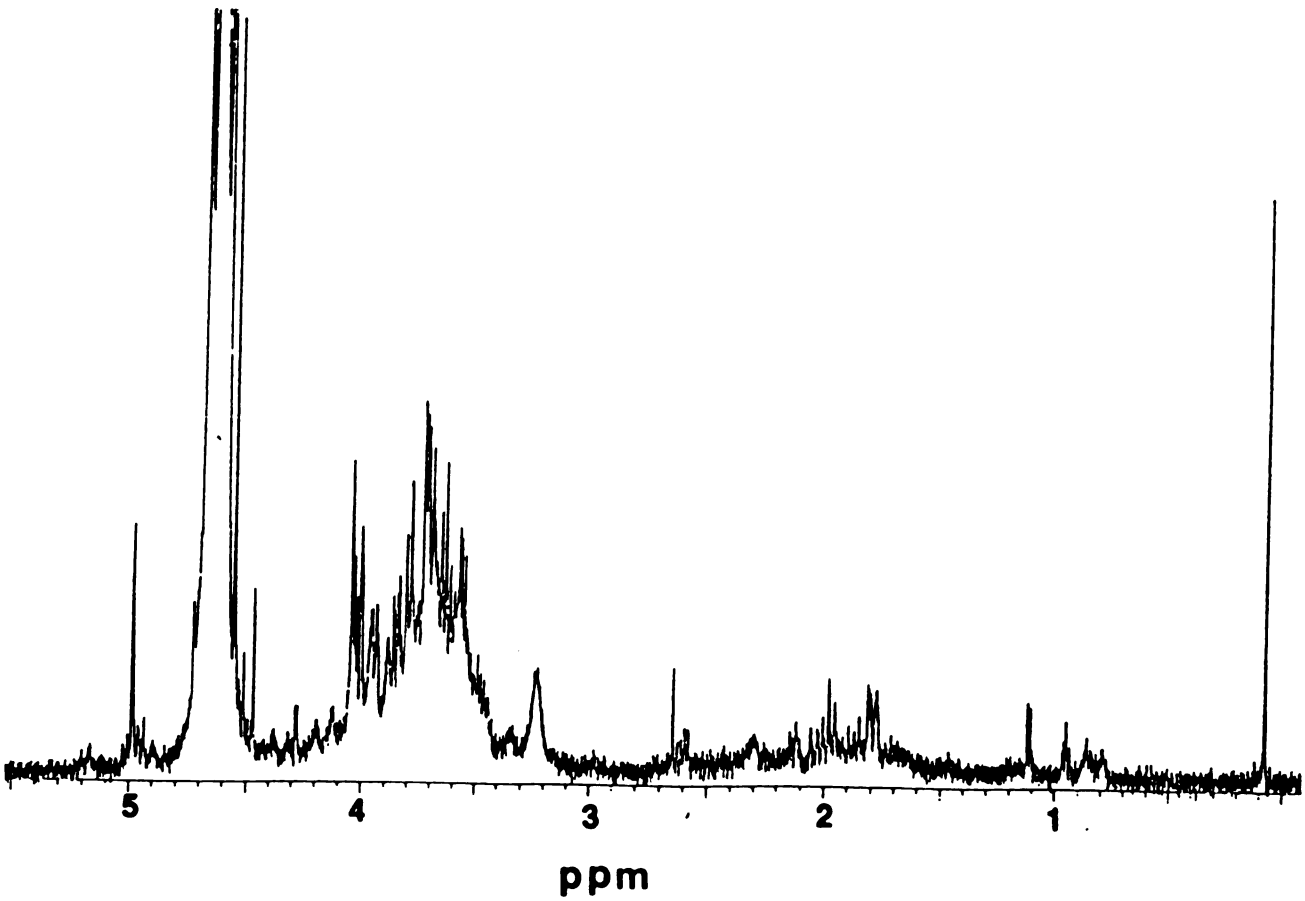


Figure 6.9a ^{1}H-NMR spectrum of the disaccharide from peak 4 of Figure 6.7 purified by ion-exchange chromatography.

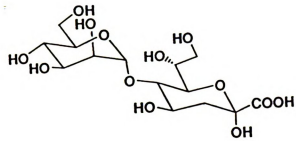


Figure 6.9b: The structure of the disaccharide of mannose and KDO.

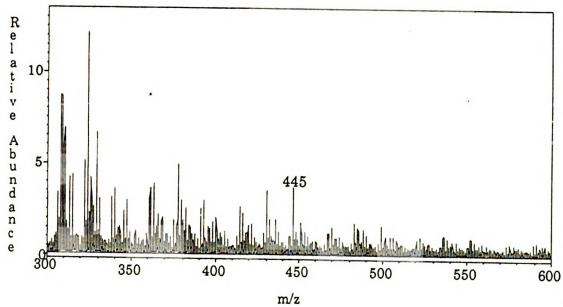


Figure 6.10: The mass spectrum of the sodium adduct to the sodium carbonated disaccharide α-Man-(1-5)-KDO.

6.5 References

1. Bohlool, B. B., and Schmidt, E. L., *Science,* (1974) 185 269-271.
2. Dazzo, F. B., and Hubbell, D., *Appl. Microbiol.*, (1975) 30 1071-1023.
3. Maier, R., and Brill, W., *J. Bacteriol.*, (1978) 133 1295-1299
4. Noel, K., Vandenbosch, K. and Kulpaca, B., *J. Bacteriol.*, (1986) 168 1392-1401.
5. Priefer, U., *J. Bacteriol.*, (1989) 171 6161-6168.
6. Carlson, R. W., Garcia, F., Noel, D. and R. I. Hollingsworth, *Carbohydr. Res.*, (1989) 195 101-110.
7. Hollingsworth, R. I. and Carlson, R. W., *J. Biol. Chem.*, (1989) 264 9300-9303.
8. Bhat, U. R., Mayer, H., Yokuta, A., Hollingsworth, R. I. and Carlson, R. W., *J. Bacteriol.*, (1991) 173 2155-2159.
9. Hollingsworth, R. I., and Elghanian, D. L., *J. Biol. Chem.* (1989) 264 14039-14042.
10. Hollingsworth, R. I. and Elghanian, D. L., *"Endotoxin Research"* (1990) 1 73-84, Spitzer, N. J. J. and Ziegler, E. J. (Eds.), Elsevier, Amsterdaam.
11. Carlson, R. W., Sanders, R., Napoli, C. and Albersheim, D., *Plant Physiol.*, (1978) 62 912-917.
12. Hollingsworth, R. I., Carlson, R. W., Garci, F. and Gage, D. *J. Biol. Chem.*, (1989) 264 9294-9299.
13. Hollingsworth, R. I., Carlson, R. W., Garci, F. and Gage, D., *J. Biol.Chem.*, (1989) 65 12752.
14. Carlson, R. W., Hollingsworth, R. I. and Dazzo, F. B., *Carbohydr. Res.*, (1988) 176 127-135.
15. Carlson, R. W., unpublished data.

16. Hollingsworth, R. I., Abe, M. and Dazzo, F. B., *Carbohydr. Res.*, **(1984)** 133 C1-C4.

17. Westphal, O. and Luderitz, O., *Angew. Chem. Int. Ed. Enl.* **(1954)** 66 407-417.

18. Dubois, M., Gilles, R. A., Hamilton, J. K., Roberts, P. A. and Smith, F., *Anal. Chem.*, **(1956)** 25 350-354.

19. Hollingsworth, R. I., Hrabak, E. M. and Dazzo, F. B., *Carbohydr. Res.*, **(1986)** 154 103-113.

CHAPTER 7

ISOLATION AND TENTATIVE CHARACTERIZATION OF A TRISACCHARIDE CONTAINING A NEW 2-DEOXY-3-C-(HYDROXYMETHYL)PENTOFURANOSYL COMPONENT FROM THE LIPOPOLYSACCHARIDE OF A *RHIZOBIUM LEGUMINOSARUM* BIOVAR *VICIAE* LPS MUTANT

7.1 Introduction

In a recent study (1), we described structural defects of the lipopolysaccharide of two mutants of *Rhizobium leguminosarum* biovar *viciae* VF-39 which are impaired in their symbiotic phenotype with respect to infection of their plant host. This impairment is characterized by failure of the bacteria to be released from infection threads into the nodules (2). One of these mutants, VF-39-32, was found to be incapable of synthesizing a tetrasaccharide component normally made by the wild-type organism. This tetrasaccharide contains mannose, galactose, galacturonic acid and 3-deoxy-2-octulosonic acid. Instead, an oligosaccharide containing an unusual deoxyglycosyl component was isolated. In this report we describe the structure of this unusual glycosyl component and complete the characterization of the effect of the mutation on lipopolysaccharide structure.

7.2 Experimental

7.2.1 Culture of Bacteria and Isolation of LPS and LPS Fractions

Bacteria were grown in a modified Bergensens medium (3) in shaken-broth cultures. Mutants were grown (4) in the presence of antibiotics. The LPS were isolated by the hot-phenol method (5) and after treatment with DNAse and RNAse purified on Sepharose 4B using an ammonium formate

buffer as described (6, 7). Column eluates were analyzed by the phenol-sulfuric acid method (8). Fragments were released from the LPS by hydrolysis for two hours with aqueous 1% acetic acid at 100° C. The hydrolysate was extracted with 5:1 chloroform-methanol in order to remove lipid A, the aqueous layer was lyophilized, and the residue was chromatographed on a column of Biogel P-2 using aqueous 0.1% formic acid. Fractions which appeared impure by ^{1}H-NMR spectroscopy and other analyses were subjected to ion-exchange chromatography on DEAE-Sephadex G25 in aqueous 0.01% formic acid adjusted to pH 6.5 with ammonia, eluting with a linear gradient of 0-0.2 *M* ammonium chloride over 200 mL.

7.2.2 Compositional Analyses

Compositional analyses for fatty acids and carbohydrates in lipid A were carried out as described (6, 7, 9, 10) after conversion into methyl esters and alditol acetates, respectively. GC analysis of carbohydrates was carried out on a Hewlett Packard 5890 gas chromatograph using a capillary DB225 column and a flame-ionization detector. Mass spectrometry detection was effected with a JEOL 505 mass spectrometer in the EI or CI (ammonia) mode.

7.2.3 Methylation Analysis

After treating with methanol containing a trace of HCl at 55° C followed by evaporation to dryness and reduction by sodium borodeuteride, the product was permethylated with dimsyl anion and methyl iodide, concentrated to dryness and adsorbed on a C-18 reverse-phase cartridge in water. The cartridge was eluted with water, 2:1 methanol/water, and finally with pure methanol. Each eluate was collected separately and subjected to acid hydrolysis with aqueous 2 *M* TFA at 120° C for 1 hour. The hydrolysates

were then reduced with sodium borohydride, acetylated, and subjected to GC/MS analysis.

7.2.4 Structural Analyses

Oligosaccharides were methylated after reduction of carboxyl and other carbonyl functions with sodium borodeuteride as described. ^1H NMR spectroscopy (500-MHz) was performed on solutions in D_2O with a Varian VXR500 spectrometer. FAB-MS was performed with a JEOL HX110-HF instrument in the negative- or positive-ion mode and using a Cs or Xe source.

7.3 Results and Discussion

Chromatography on Sepharose 4B of the crude LPS isolated from the wild-type strain *Rhizobium leguminosarum* biovar *viciae* VF-39 (top part of Figure 7.1) showed four major fractions. A-C corresponded to LPS, and the late-eluting peak was composed almost entirely of a glucan. Hydrolysis of each of the fractions A-C with 1% acetic acid followed by extraction of the lipid A with chloroform-methanol (as described under general methods) and chromatography of the components in the aqueous phase on Biogel P-2 gave three sub-fractions, the three elution profiles were similar (bottom part of Figure 7.1).

The gel-filtration profile on Sepharose 4B of the total LPS of the mutant VF- 39-32 was radically different from that of the wild-type (top part of Figure 7.2). Two major LPS fractions (A and B) which, after hydrolysis (with 1% acetic acid as described in general methods) gave elution profiles on Biogel P-2 (bottom part of Figure 7.2) that were different from those (bottom part of Figure 7.1) obtained from the wild-type. Thus, peak 1 from fraction A contained none of the amino sugars found in the wild type and appeared to

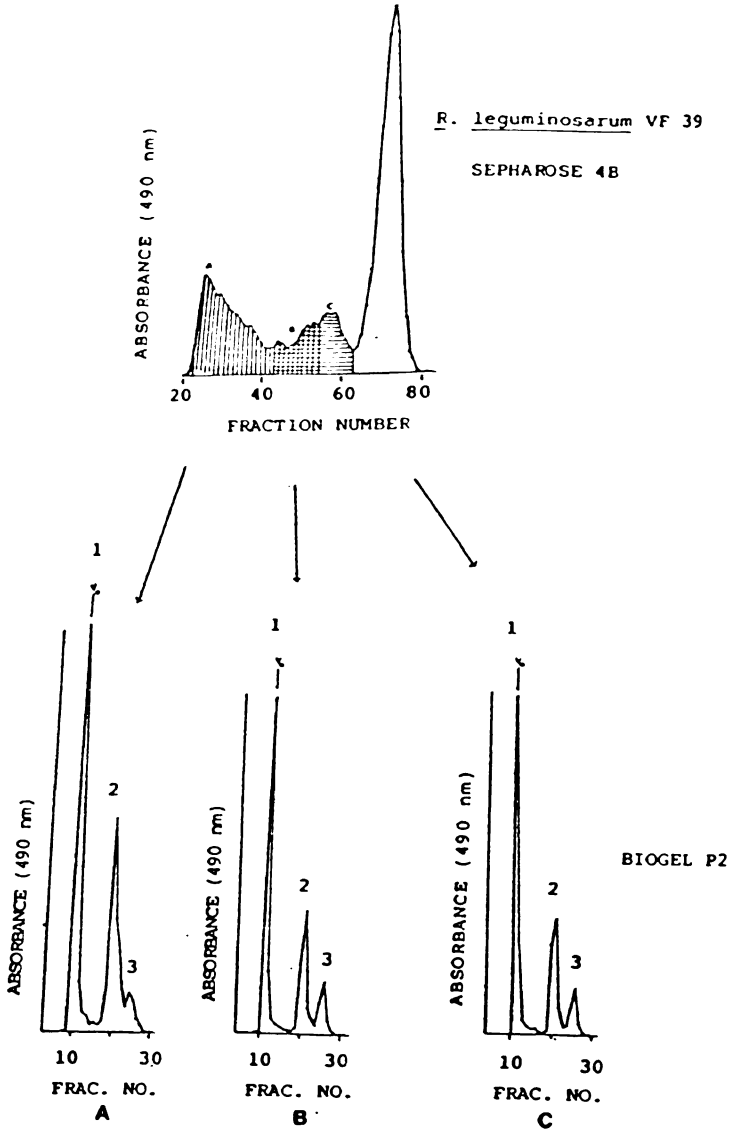


Figure 7.1: The gel filtration profile on Sepharose 4B of the crude LPS isolated from *Rhizobium leguminosarum* biovar *viciae* VF-39 is shown at the top of the Figure. Fractions A-C were each subjected to mild hydrolysis with 1% acetic acid at 100° C for two hours and then extracted with 5:1 chloroform-methanol. The aqueous layer was then subjected to gel filtration chromatography on Biogel P-2. Peaks at bottom of the figure with the same numbers had similar compositions and NMR spectra, and contained the *O*-antigen, a tetrasaccharide **1** and a trisaccharide **2**. The vertical axes represent absorbance in the phenol-sulfuric acid assay.

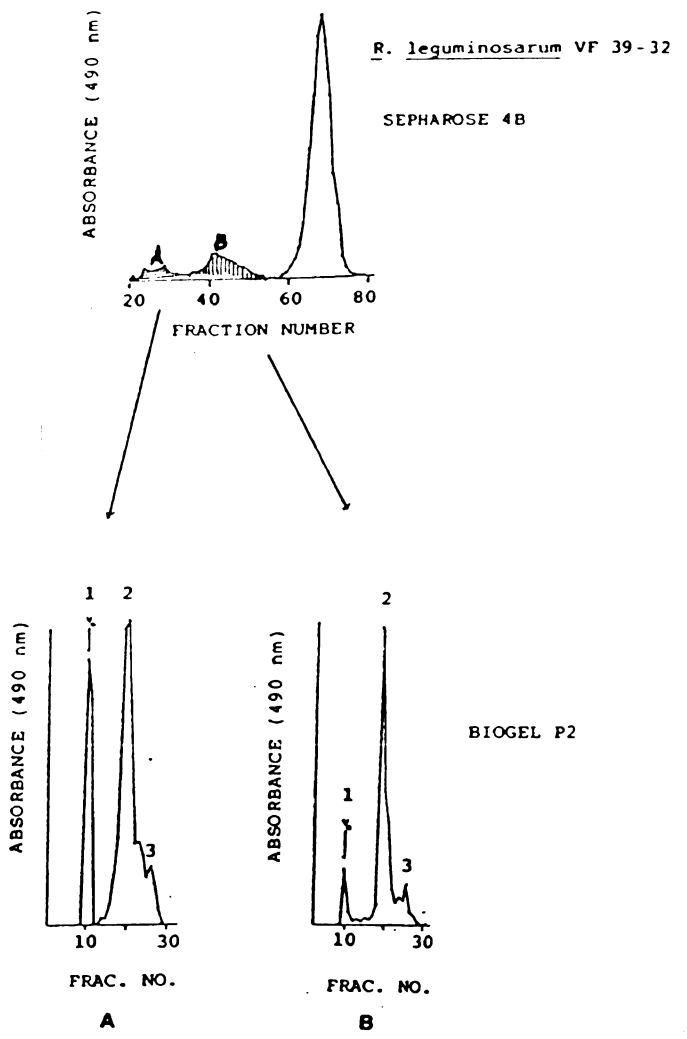


Figure 7.2: Gel filtration profile on Sepharose 4B of the crude LPS of *Rhizobium leguminosarum* biovar *viciae* VF-39-32 is shown at the top of the figure. Fractions A and B were each subjected to mild-acid hydrolysis with 1% acetic acid at 100° C for two hours and then extracted with 5:1 chloroform-methanol. The aqueous layer was subjected to gel filtration chromatography on Biogel P2, and the resulting elution profile is shown at the bottom of the figure. Peak 1 from fraction A, which contained mainly glucose indicating it might be due to small amounts of capsular material, was devoid of amino sugars. Peak 1 from fraction B contained a small amount of *O*-antigen. Peak 3 from fractions A and B contained the trisaccharide 2, and peak 2 contained an unidentified product.

contain only a small amount of a capsular antigen comprising mainly glucose. Peaks 2 and 3 from fractions A and B contained similar labeled major components. Peak 1 from fraction B contained small amounts of the *O*-antigen found in the parent strain. Peak 3 from fractions A and B contained the trisaccharide 2, and Peak 2 contained an unidentified product.

The ^{1}H NMR spectrum (Figure 7.3) of peak 2 (at the bottom of Figure 7.2), the major component from the ion-exchange column, suggested the presence of a deoxy function at one of the ring positions by the presence of an up-field doublet of doublets (J = 14.5 + 1 Hz) at 2.02 ppm and another doublet of doublets (J = 14.5 + 7.6 Hz) at 2.60 ppm. The unusually large chemical shift at 2.60 ppm suggested that the ring might have the furanose form, and the large splitting indicated geminal protons. Irradiation of the more up-field doublet of doublets at 2.02 ppm (Figure 7.4) led to the collapse of the second doublet of doublets at 2.60 ppm as well as loss of the smaller splitting of another doublet of doublets (J = 7.6 + 1 Hz) at 4.53 ppm. This latter signal was assigned to an anomeric proton. This experiment also confirmed that the pair of up-field doublet of doublets were indeed mutually coupled. Irradiation of the doublet of doublets at 2.60 ppm (Figure 7.5) led to the collapse of the more up-field signal at 2.02 ppm and to the collapse of the anomeric proton at 4.53 ppm. This indicated the presence of a 2-deoxy function in the molecule. The lack of perturbation of any other signals besides the anomeric proton when either of the 2-deoxy methylene protons was irradiated indicated that no proton was present on the 3-position. The signal with the large splittings at 2.60 ppm, are very far down-field indicates that is not due to an axial proton with an axial neighbor on an adjacent carbon in a pyranose chair form. The possibility of a pyranose chair form was also ruled out by the very small coupling constant displayed between the most up-field doublet of doublets at

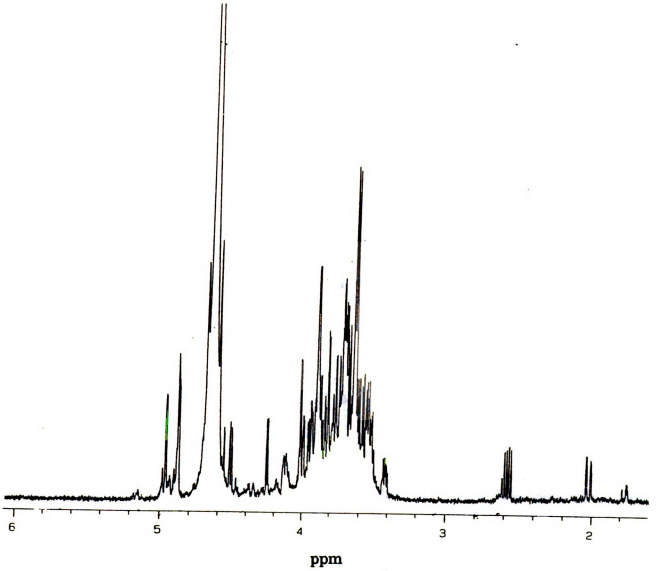


Figure 7.3: 500MHz ^{1}H-NMR spectrum of peak 2 (at the bottom part of figure 7.2), which was the aqueous fraction of the mild, partial acid hydrolysate of peak B (at the top part of Figure 7.2) analyzed by the Biogel P-2 column. Note the presence of the upfield doublet of doublets (J = 14.5 Hz + 1 Hz) at 2.02 ppm and a second doublet of doublets (J = 14.5 Hz + 7.6 Hz) at 2.60 ppm. The 14.5 Hz mutual coupling indicates the presence of a methylene group and the downfield chemical shift suggests a deoxy function. The unusual chemical shift with the 7.6 Hz coupling indicates that the ring might not be in a pyranose form.

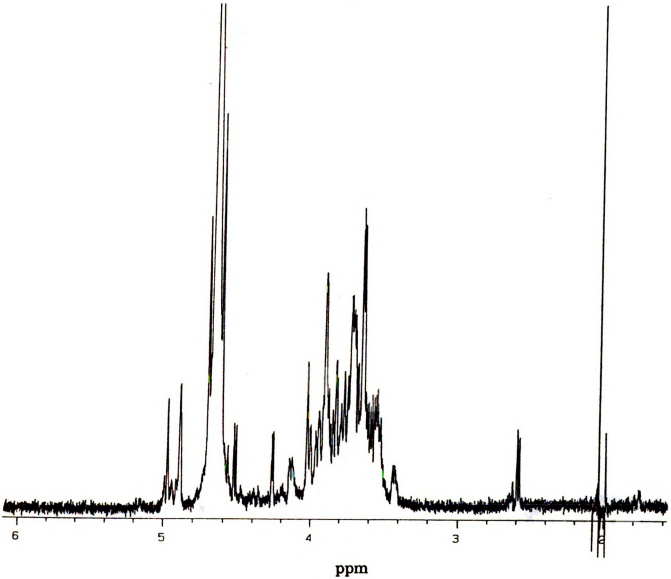


Figure 7.4: 500MHz ^{1}H-NMR spectrum of peak 2 (at the bottom of Figure 7.2) from the Biogel P-2 column with irradiation of the doublet of doublets at 2.02 ppm. Note the collapse of the other doublet of doublets to a simple doublet with a splitting of 7.6 Hz. The loss of the 14.5 Hz coupling confirms that the two upfield signals of 2.02 and 2.60 ppm are due to geminal protons. Note also that another doublet of doublets (J = 7.6 Hz + 1 Hz) at 4.53 ppm (Figure 7.3) has lost the smaller splitting. The chemical shift of this signal suggests that it is due to an anomeric proton, thus making the methylene group a 2-deoxy function.

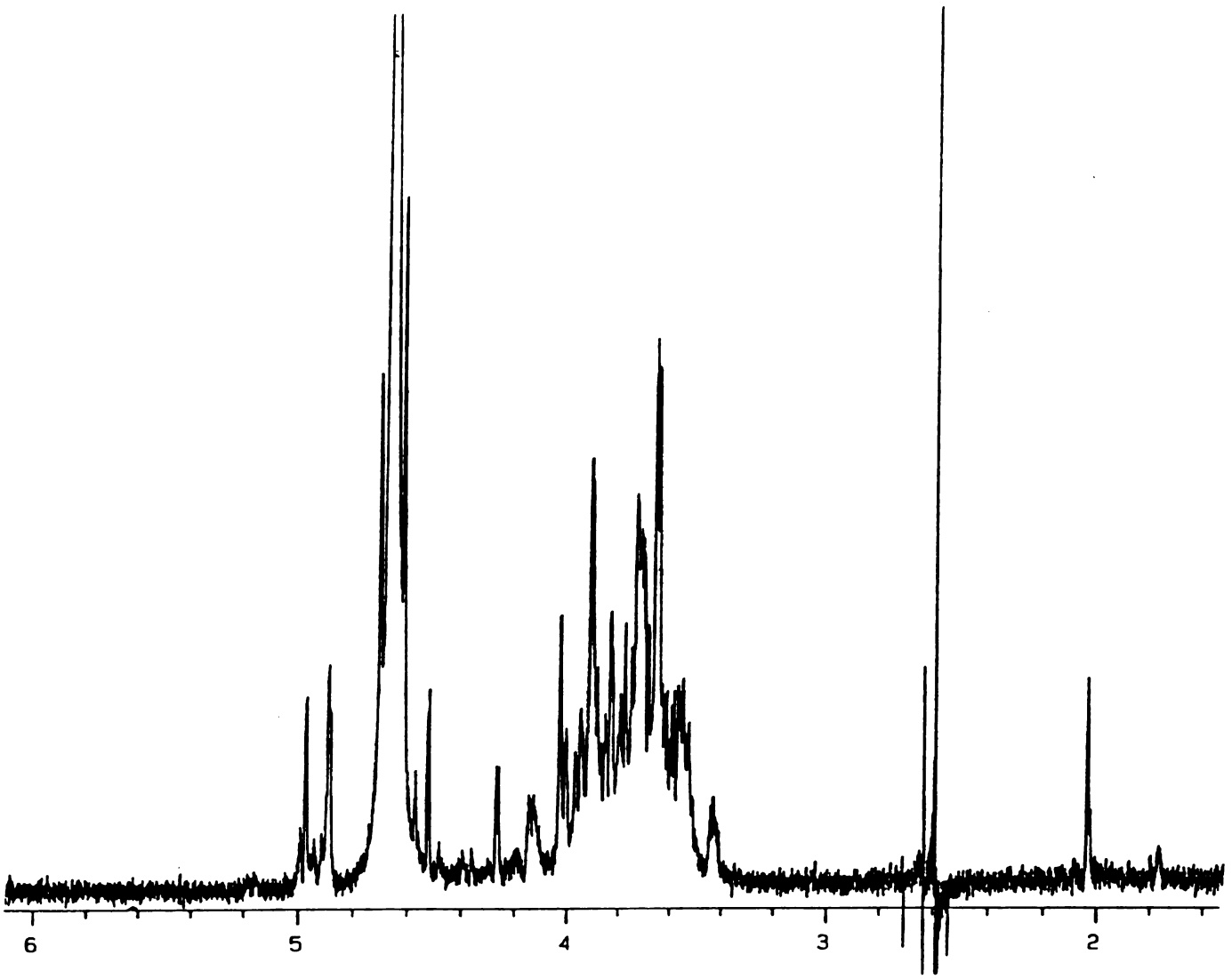


Figure 7.5: 500MHz ^{1}H-NMR spectrum as in Figure 7.3, after irradiating the signal at 2.60 ppm. Note that the signal at 2.02 ppm has collapsed to a very narrow doublet. Note also that the signal at 4.53 ppm has collapsed to a doublet with a very small splitting. This result confirms that the two upfield signals and the one at 4.53 ppm form a closed-spin system and confirms that the glycosyl residue responsible for them has a 2-deoxy function. The lack of any other coupling to the 2-deoxy-methylene protons indicates that there are no protons on the 3-position.

2.02 ppm and the anomeric proton at 4.53 ppm. In a pyranose chair configuration, the most up-field proton would be equatorial and *cis-* with the anomeric proton, and the expected coupling constant for these *cis-* protons should be 3-4 Hz instead of the observed 1 Hz. In the Figure 7.3, the anomeric proton must be axial in order to show the 7.6 Hz coupling with one of the neighboring C-2 hydrogen. In addition, this most up-field doublet of doublets shifts too high compared to the other 2-deoxy-methylene signal to be the equatorial component. These facts strongly suggested that a pyranose form did not exist.

There are two signals in the anomeric region of the spectrum (Figure 7.3 and expanded in Figure 7.6A). One signal at 4.98 ppm was assigned to an α-anomeric proton of mannose and the other, a doublet at 4.89 ppm, with 1.5 Hz coupling, was assigned to the α-anomeric proton of galactose. Upon irradiation of the doublet of the doublets at 2.60 ppm, an increase in the intensity of the mannosyl residue relative to the more upfield galacosyl residue is noted (Figure 7.5 and expanded in Figure 7.6B). Also on irradiation of the doublet of doublets at 2.02 ppm, enhancement of the mannose signal is seen (Figure 7.4 and expanded in Figure 7.6C). This indicates that the mannosyl residue is in the position adjacent to the 2-deoxy-methylene group. The n. O. e. to anomeric proton of mannose in the 3 position of the furanose ring can be observed; this is evidence for a five-membered ring, because generally such phenomena are not observed in the pyranose form.

After reduction by sodium borohydride, acid hydrolysis, reduction and acetylation, gas chromatography produced the chromatogram of the alditol acetates of the oligosaccharide components shown in Figure 7.7. There were

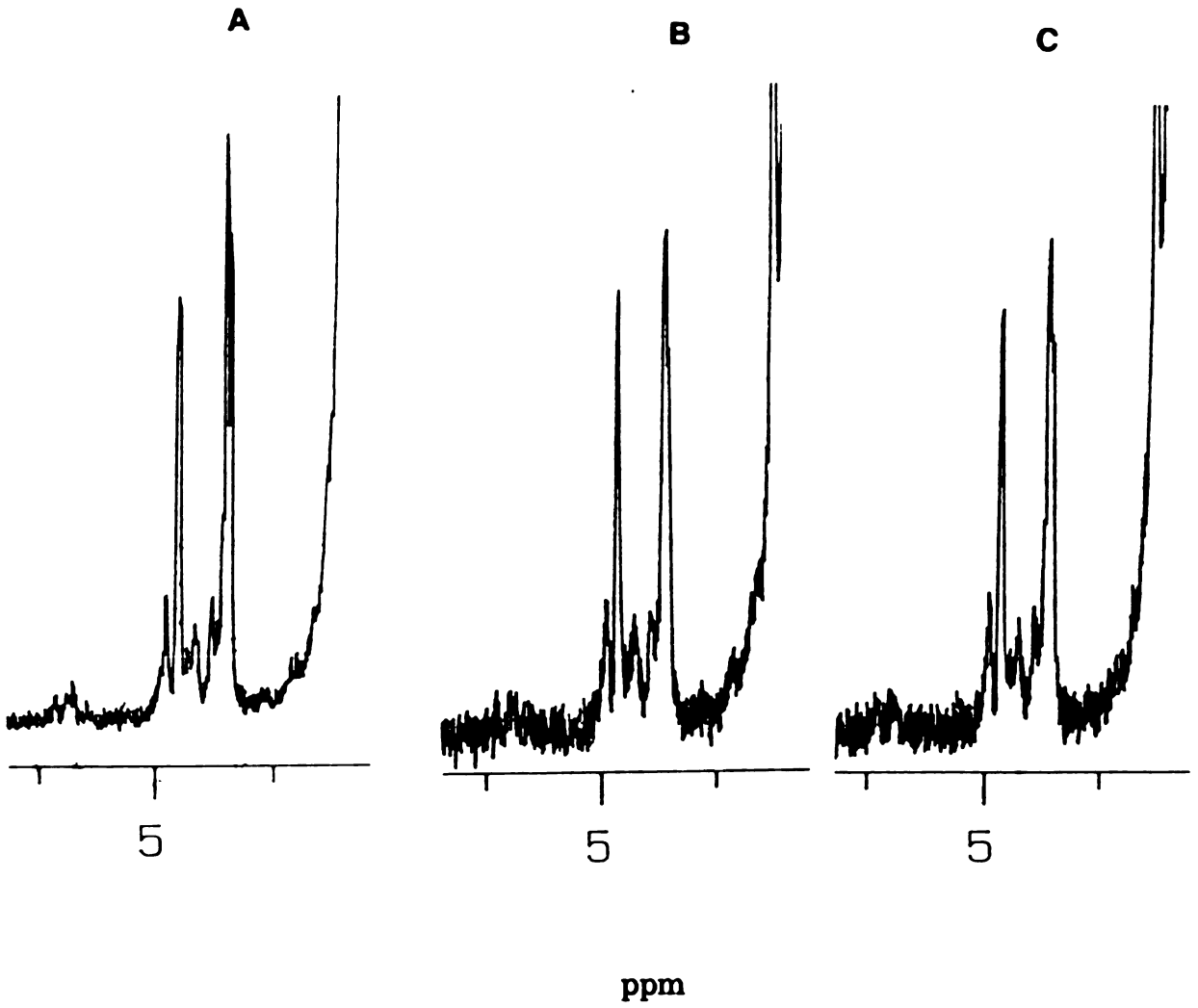


Figure 7.6:
A) Partial ^{1}H-NMR spectrum showing anomeric protons of mannose and galactose residues.

B) Spectrum after irradiating of the doublet of doublets at 2.60 ppm. Note the increase in intensity of the mannosyl residue relative to the more upfield galactosyl residue. This indicates that the mannosyl residue is in the position adjacent to the methylene group.

C) Spectrum after irradiation of the doublet of doublets at 2.02 ppm. Again note the enhancement of the mannose signal.

four peaks appearing as a pair of doublets. Peaks 1 and 2 had identical mass spectra. The chemical ionization mass spectrum, using ammonium as a carrier gas, showed that the first peak of the first pair had a pseudo-molecular ion at m/z 352 (Figure 7.8) which corresponds to a true molecular weight of 334, since analytes form an ammonium adduct under the CI conditions. The structure of the first alditol acetate is shown in Figure 7.9; peracetylation of the first alditol acetate was not complete, because the tertiary hydroxyl group is difficult to acetylate. The pseudo-molecular ion for the second peak of the first pair appeared at m/z 366 (Figure 7.10), which suggested a true molecular weight of 348, and indicated that the two peaks differed by 14 mass units or one methyl group. The structure of the second alditol acetate is shown in Figure 7.11. Methylation at the tertiary hydroxyl group explained the 14 mass units different between the two structures in Figures 7.9 and 7.11; this was due to 1% HCl in methanol applied during the derivatization procedure. Peak 3 in Figure 7.7 is due to mannose, and peak 4 is due to galactose. These later two components were detected in equal proportions; these assignments were confirmed by GC/MS. The GC and GC/MS assignment of the mannose and galactose also confirmed the ^{1}H-NMR spectrum (Figure 7.3), in which two signals appeared: one singlet at 4.98 ppm and the other signal at 4.84 ppm with 1.5 Hz coupling, characteristic of these residues.

Retention time

Figure 7.7: Gas chromatogram of the alditol acetate of the unusual glycosyl components. Peaks 1 and 2 are due to the component being identified. Peak 3 is mannose, and peak 4 is galactose.

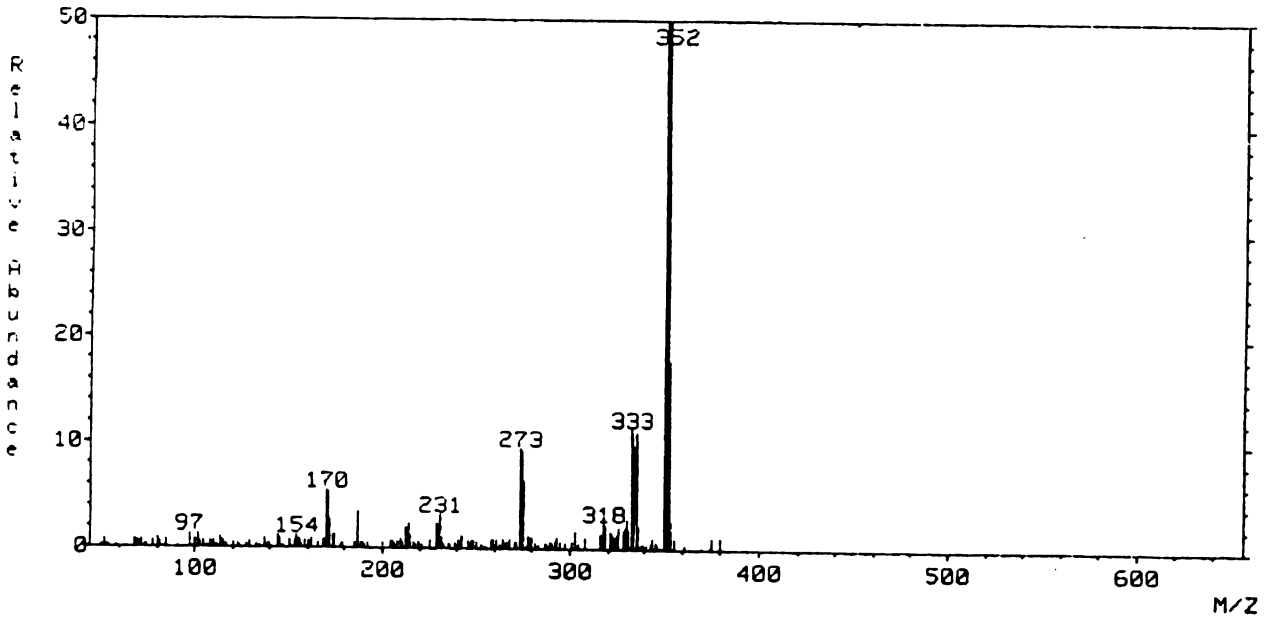
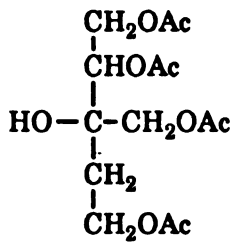


Figure 7.8: The CI mass spectrum of the first peak from the first pair of doublets in the gas chromatogram of Figure 7.7.



m/z 334

Figure 7.9: The structure of the alditol acetate in Figure 7.8.

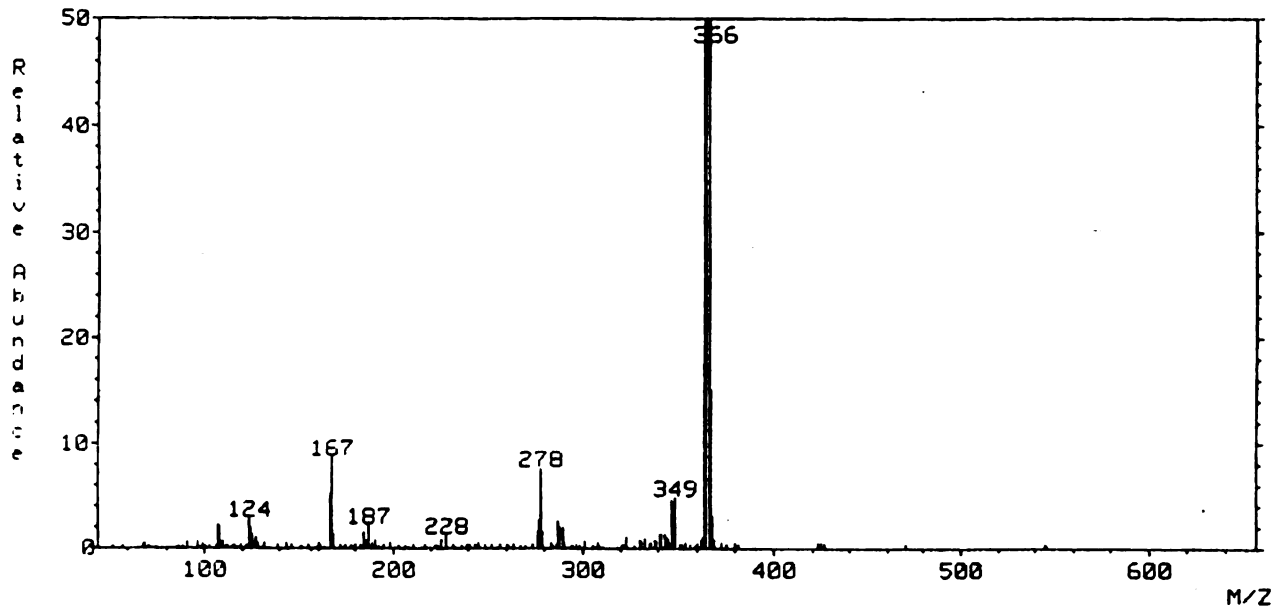
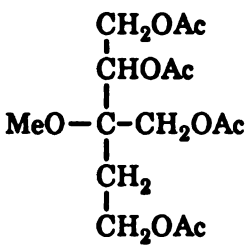


Figure 7.10: The CI mass spectrum of the second peak from the first pair of doublets in the gas chromatogram of Figure 7.7.



m/z 348

Figure 7.11: The structure of the alditol acetate for Figure 7.10.

The oligosaccharide was subjected to methylation analysis. After treating with methanol containing a trace of HCl at 55° C followed by evaporation to dryness and reduction by sodium borodeuteride, the product was then permethylated with dimsyl anion and methyl iodide, concentrated to dryness and adsorbed to a C-18 reverse phase cartridge in water. The cartridge was eluted with water, 2:1 methanol/water, and finally with pure methanol. Each eluant was collected separately and subjected to acid hydrolysis with aqueous 2 M TFA at 120° C for 1 hour. The hydrolysates were then reduced with sodium borohydride, acetylated, and subjected to GC/MS analysis. The total ion chromatogram (TIC) for the methylation analysis appears in Figure 7.12. The fraction contained three major peaks, the small peaks at the very front were due to minor contaminate. Peaks 1 and 2 had identical mass spectra, and were due to terminal mannose and galactose (Figure 7.13), the structure and fragmentations were showed in Figure 7.14. The peak 3 had a mass spectrum which is unlike that of a normal partially methylated, peracetylated hexose (Figure 7.15). The branching residue in the unusual compound is confirmed by the m/z 261 and 203 fragmentations, and this also supported the NMR evidence that there is no proton at C-3. The proposed structure and fragmentations are shown in Figure 7.16. The 2:1 methanol/water fraction was a major component which contained terminal mannose, terminal galactose and a KDO derivative.

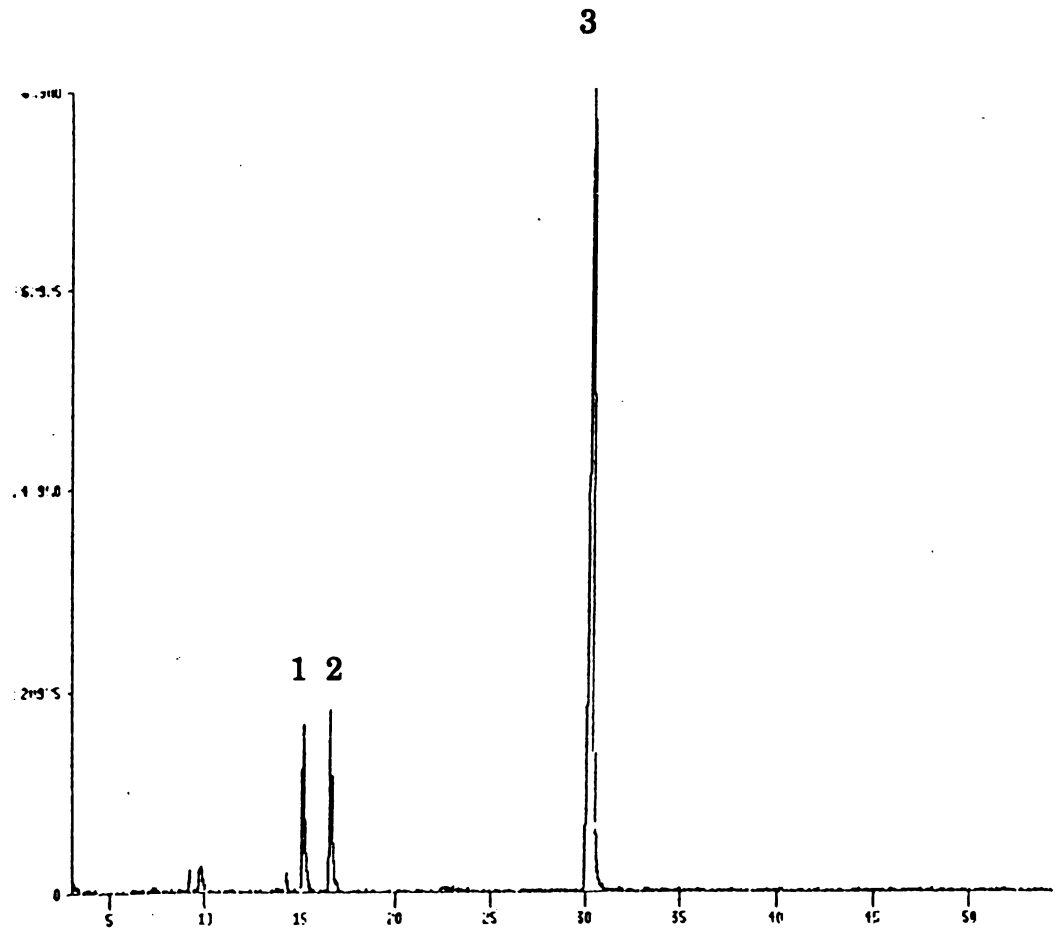


Figure 7.12: The total ion chromatogram for the methylation analysis and fraction contained three major peaks, the small peaks at the very front were due to minor contaminants.

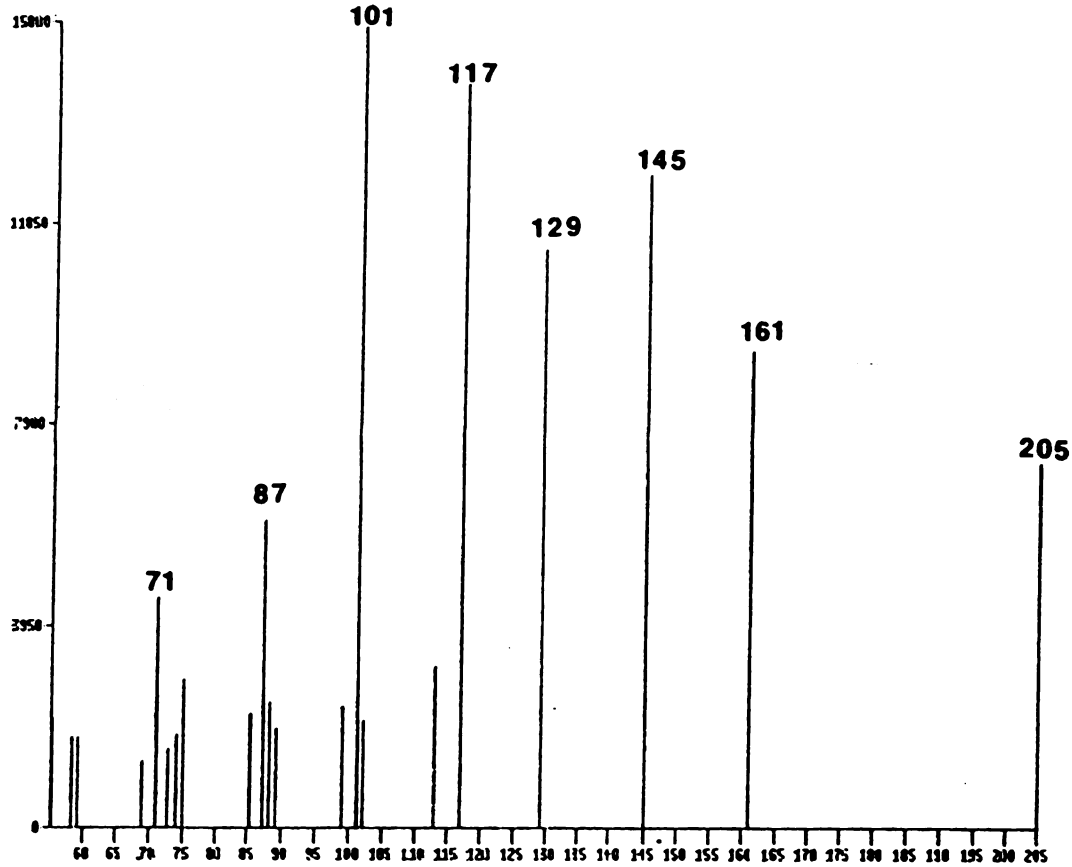


Figure 7.13: Mass spectrum of peak 1 and 2 from Figure 7.12.

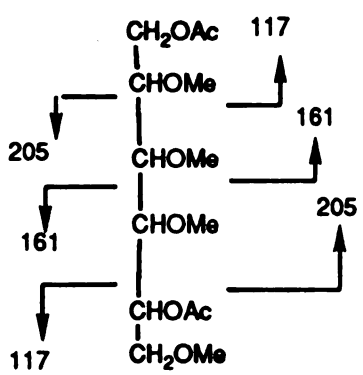


Figure 7.14: The structure and fragmentations for Figure 7.13.

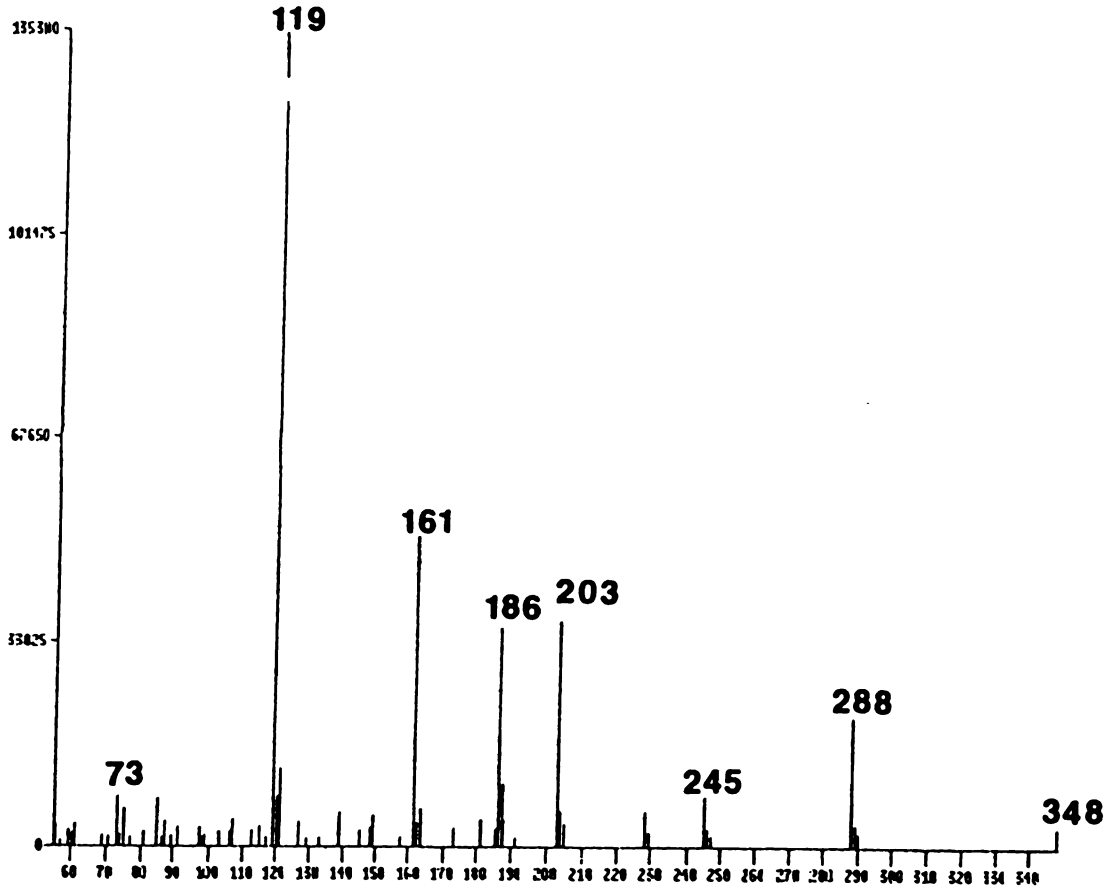


Figure 7.15: Mass spectrum of peak 3 from Figure 7.12.

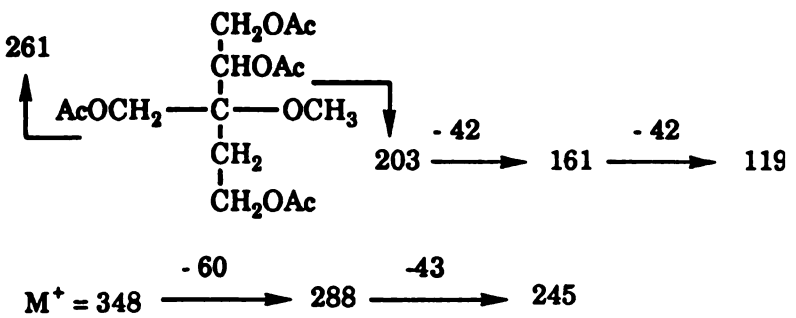


Figure 7.16: The structure and fragmentations for Figure 7.15, the branching residue is suggested by the 261 and 203 fragments and is supported by the NMR evidence that there is no proton at C-3.

There are two possible configurations of the furanose ring: one is envelope form shown in Figure 7.17, and another is the twisted ring form shown in Figure 7.18.

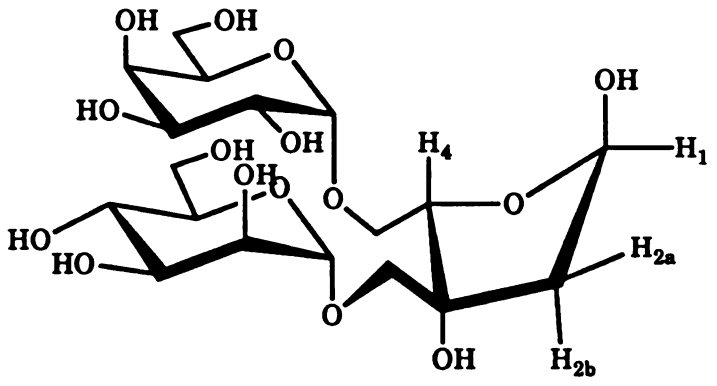


Figure 7.17: Envelope configuration of the furanose ring.

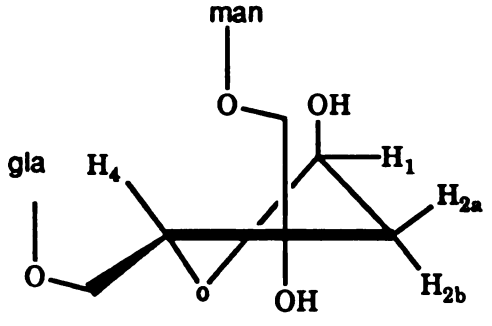


Figure 7.18: Twist configuration of the furanose ring.

According to the semiempirical calculations (11, 12) of five member ring compounds of butyrolactone and 3-hydroxybutyrolactone, the calculated coupling constants of the envelope form and the twisted ring are very close, and the difference in energy of these two forms is that the envelope form is 0.3 kJ mol^{-1} higher than the twisted ring. In both Figures 7.17 and 7.18, all of the hydroxyl groups prefer to occupy the pseudo-axial positions in order to

reduce the electrostatic effect between the two pseudo equatorial hydroxyl groups. From the current data obtained in this research it is difficult to determine the exact configuration of the furanose ring. Further computational work is necessary to get this part of the work done.

7.4 References

1. Zhang, Y., Hollingsworth, R. I., and Priefer, U. B. *Carbhydr. Res.* (1992) 231 261-271.
2. Maier, R. and Brill, W., *J. Bacteriol.*, (1978) 133 1295-1299.
3. Hollingsworth, R. I., Abe, M. and Dazzo, F. B., *Carbohydr. Res.*, (1984) 133 C1-C4.
4. Priefer, U., *J. Bacteriol.*, (1989) 171 6161-6168.
5. Westphal, O. and Luderitz, O., *Angew. Chem.*, (1954) 66 407-417.
6. Hollingsworth, R. I., and Elghanian, D. L., *J. Biol. Chem.* (1989) 264 14039-14042.
7. Hollingsworth, R. I. and Elghanian, D. Lill, *"Endotoxin Research,"* (1990) 1 73-84, Spitzer, N. J. J. and Ziegler, E. J. (Eds.), Elsevier, Amsterdam.
8. Dubois, M., Gilles, R. A., Hamilton, J. K., Roberts, P. A. and Smith, F., *Anal. Chem.*, (1956) 25 350-354.
9. Hollingsworth, R. I. and Carlson, R., *J. Biol. Chem.*, (1989) 264 9300-9303.
10. Bhat, U. R., Mayer, H., Yokuta, A., Hollingsworth, R. I. and Carlson, R. W., *J. Bacteriol.*, (1991) 173 2155-2159.
11. Riggs, N. V., *Aust. J. Chem.*, (1985) 38 1575-1583.
12. Riggs, N. V., *Aust. J. Chem.*, (1985) 38 1585-1589.