This is to certify that the

dissertation entitled

**Analysis of the Molecular Basis for the
Supramolecular Structure of Lipopolysaccharides**

presented by

Ying Wang

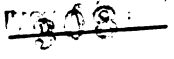has been accepted towards fulfillment
of the requirements for

Ph.D _____ degree in ___Chemistry___

_____
Major professor

Date _03/23/95_

DATE DUE	DATE DUE	DATE DUE

ANALYSIS OF THE MOLECULAR BASIS FOR THE

SUPRAMOLECULAR STRUCTURE OF LIPOPOLYSACCHARIDES

By

Ying Wang

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Chemistry

1995

ABSTRACT

ANALYSIS OF THE MOLECULAR BASIS FOR THE SUPRAMOLECULAR STRUCTURE OF LIPOPOLYSACCHARIDES

By

Ying Wang

Bacterial lipopolysaccharides trigger a cascade of biological activities in mammalian systems. The supramolecular structure of lipopolysaccharide aggregates is a highly ordered self-assembly which is known to contribute to its biological activities, and is also an attractive candidate for a variety of applications. To gain insight into the molecular basis for the formation of three-dimensional structures, characterization of the chemical structure is crucial. The lipopolysaccharide of *Rhizobium trifolii* has been studied. The structures of all three components, the O-antigen, the core and the lipid A regions, have been determined using a combination of chemical and spectroscopic methods.

The O-antigenic polysaccharide has been determined to contain a pentasaccharide repeating unit with the structure of $\rightarrow$3)-α-L-Rha-(1$\rightarrow$3)-[α-D-ManNAc-(1$\rightarrow$2)]-α-L-Rha-(1$\rightarrow$4)-β-D-GlcNAc-(1$\rightarrow$3)-α-L-Rha-(1$\rightarrow$. The glycosyl compositions were obtained by GC and GC-MS of alditol acetate derivatives. The linkage positions and configurations of the glycosyl residues were obtained by 1D and 2D NMR spectroscopy. The sequence was deduced by 1D and 2D n.O.e. experiments and by partial hydrolysis studies. The core component has been isolated and contains a trisaccharide identical to the one found in another strain.

A solvent system has been developed which is composed of 1:1:2:10 pyridine-D_5 / DCl in D_2O / CD_3OD /$CDCl_3$. The use of this solvent system gives uniformly well-resolved ^{1}H-NMR spectra over a wide selection of lipid classes including lipid A molecules, phospholipids and large gangliosides.

The structure of the lipid A region has been elucidated employing 1D and 2D NMR spectroscopy with the above solvent and mass spectrometry using isotope labeling with electron impact and electrospray ionization. These studies were complemented by linkage-specific and site-specific chemical cleavage methods. The lipid A contains an α-Gal-(1→6)-β-GlcN disaccharide headgroup with a lactyl ether at the glucosamine reducing end, and bears a total of five fatty acyl chains.

A computer model of the lipid A moiety of *E. coli* has been proposed based on a combination of molecular mechanics calculations and NMR spectroscopy. The calculated structure has a simple geometrical symmetry and can easily form extended hexagonal structures by Van der Waals interactions between their fatty acid chains and electrostatic interactions by magnesium or calcium / phosphate salt bridges, and provides a clear molecular basis for the formation of the observed hexagonal arrays.

To My Mother

To My Father

To My Brother and Sisters

ACKNOWLEDGMENTS

During my stay here at Michigan State University, I have met many, many nice people. They have supported me and helped me. They have made my life easier and happier. To them I owe so much. To them I wish to say, "Thank you!"

Among all these people, I wish to thank my mentor, Dr. Rawle I. Hollingsworth, the most! I could not be any luckier than being his student. He has taught me many different aspects of science, not just chemistry, NMR and other techniques, but also a way of pursuing science. He has shared his knowledge and experience unselfishly with me. He has always been so encouraging, supportive, helpful, patient, and optimistic which makes everything worthwhile and enjoyable. Rawle is just great and I thank him for everything!

I thank my committee members: Dr. John Allison, Dr. Gary Blanchard, Dr. Daniel Nocera for all their help and support. I especially thank Dr. Allison and Dr. Blanchard for writing me recommendation letters. I would also like to thank the NMR facility, especially Dr. Long Le and Dr. Johnson Kermit. I thank Michigan State University and the Department of Energy for financial support.

I wish to thank all Hollingsworth group members: Maria Beconi-Barker, Luc Berube, Jim Bradford, Rob Cedergren, Xiaoyang Du, Gang Huang, Guangfei Huang, Ben Hummel, Seunho Jung, Kung-Il Kim, Steve Lamb, Jeongrim Lee, Carol Mindock, and

Yin Tang. I thank them for all their help and support. I thank them for all the joy and fun! I wish to say, "This Hollingsworth lab is a wonderful place!"

I wish to thank my friends: Carolyn Hsu, Zhengrong Kong, Dongfeng Gu, Jingpin Jia, Wanmei Yang, Ying Jiang, Mai Luu, Shiho Fukoi, Zehui Zhang, Huiyun Luo, Ying Liang, and Shana Stock...... I especially thank Carolyn for being such a good-hearted person and for helping me through the dark days. I thank Kong for being unconditionally supportive and nice to me. I wish to thank all the Chinese students of the Chemistry Department, and I also wish to thank the Chinese community at MSU.

Lastly and deepestly, I thank my family! Without them I would not have accomplished anything. I thank my family for all their unconditional love and support. I thank them for all their encouragement and all the confidence they have in me. Specially, I thank my mom, Mongguang Cheng, who has taught me numerous stories and lessons about life from which I will benefit all my life. I thank her for being a loving, unselfish mother, and a wise, strong woman. I thank my father, Qingyi Wang, for all his financial support. I thank my brother for always helping and "protecting" me. I thank my sister-in-law for being a good friend to me. I thank my little sister for all the dear letters she wrote me.

Thank you, all my family!

Thank you, all my teachers!

Thank you, all my friends!

Thank you, all the nice people!

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

Chapter 3

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendix

LIST OF ABBREVIATIONS

BIII   Bergensen's medium

DEPT   distortionless enhancement by polarization transfer

DMAP   4-dimethylaminopyridine

DQF-COSY   double quantum filtered correlation spectroscopy

E. coli   *Escherichia coli*

ES-MS   electrospray ionization mass spectrometry

FAB-MS   fast atom bombardment mass spectrometry

Gal   galactose

GalA   galacturonic acid

GC   gas chromatography

GC-MS   gas chromatography - mass spectrometry

Glc   glucose

GlcN   glucosamine

GlcNAc   N-acetyl glucosamine

HMQC   heteronuclear multiple-quantum coherence

HOHAHA   homonuclear Hartmann-Hahn

KDO   3-deoxy-D-manno-2-octulosonic acid

LPS   lipopolysaccharide

Man   mannose

ManNAc   N-acetyl mannosamine

NMR	nuclear magnetic resonance
n.O.e.	nuclear Overhauser effect
NOESY	nuclear Overhauser and exchange spectroscopy
PC	L-α-phosphatidylcholine
PE	L-α-phosphatidylethanolamine
PS	L-α-phosphatidyl-L-serine
Rha	rhamnose
R. trifolii	*Rhizobium trifolii*
S. minnesota	*Salmonella minnesota*
TFA	triflouroacetic acid
TFAA	triflouroacetic anhydride
TOCSY	total correlation spectroscopy

CHAPTER 1

INTRODUCTION

Self assembling systems, molecular recognition and molecular intelligence, these terms and descriptions refer to molecular systems in which molecules contain structural information which direct specific intermolecular interactions to result in aggregate, supramolecular structures with defined properties. The information required to form these supramolecular structures are present in the isolated molecules and assembly is self-directed. The potential use of such structures in technology is currently limited only by our imagination and their availability. Biological systems are filled with examples of molecular structures which are capable of self assembling to form supramolecular, periodic structures. One such important class of molecules is bacterial lipopolysaccharides (LPSs). These molecules are composed of hydrophilic charged domains and hydrophobic domains which act as major structural domains which direct their self assembly into well defined periodic two dimensional lattices (Figure 1). In order to obtain insight into the basis for the supramolecular structure of LPS, detailed information on its chemical structure is crucial.

Lipopolysaccharides are amphiphilic macromolecules consisting of two regions with contrasting chemical and physical properties, a hydrophilic region composed of varying polysaccharide, and a hydrophobic lipid region, termed lipid A. Lipopolysaccharides are found exclusively on the outer membrane of gram-negative bacteria (Figure 2) (1). The outer membrane of gram-negative bacteria is chemically distinct from the usual biological membranes yet is also a bilayered structure, and resembles phospholipid bilayer membrane in its general architecture (1-5). Its inner leaflet resembles in composition of the usual phospholipid membrane. Its outer leaflet,
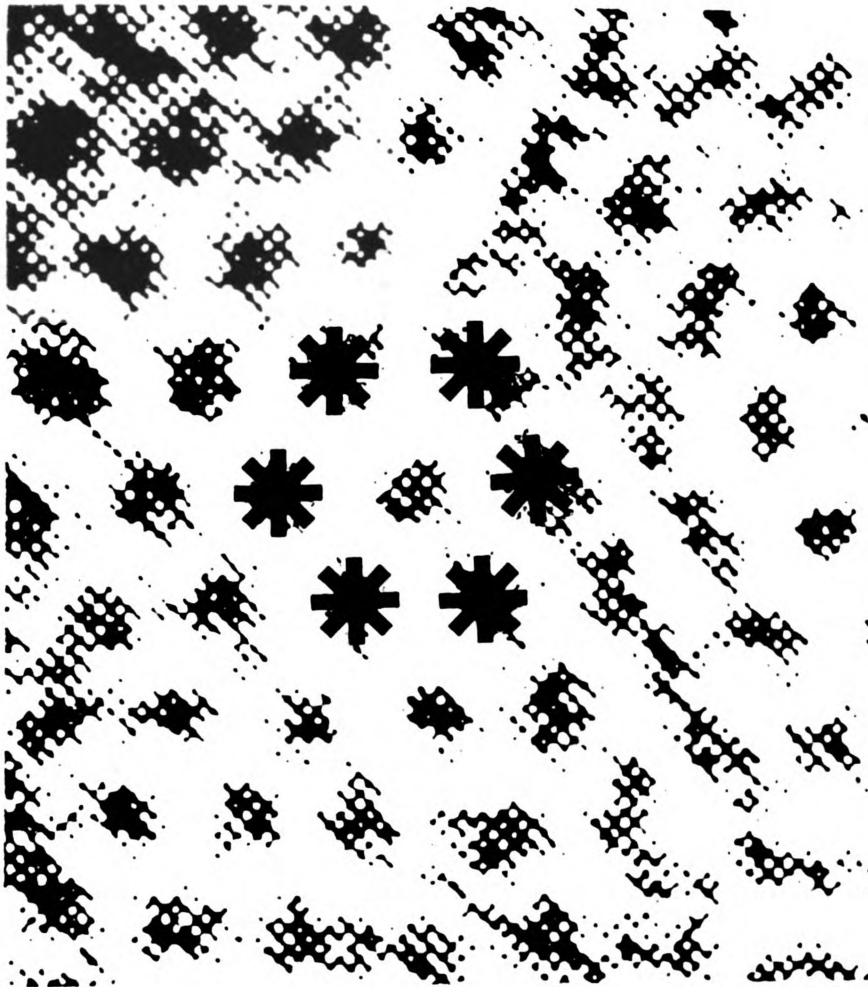
Figure 1. Electron micrograph of an aqueous film of *E. coli* LPS in the form of its calcium salt. Note the hexagonal array of aggregates reflecting the molecular order.
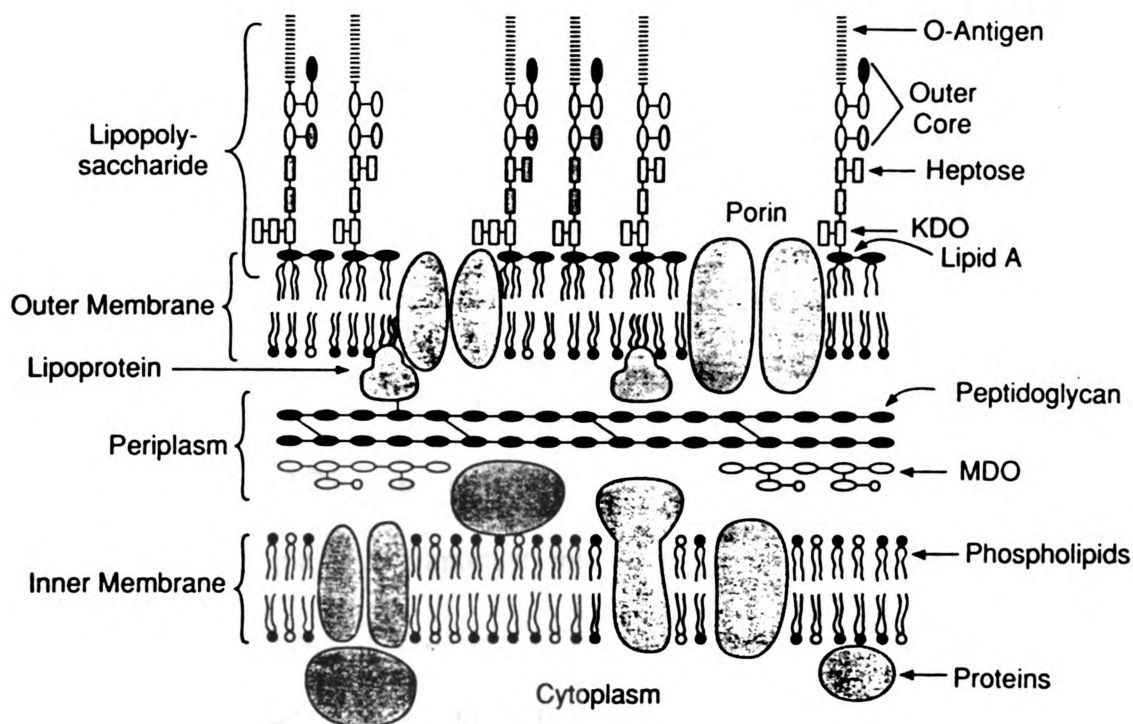
Figure 2. Molecular representation of the envelope of a gram-negative bacterium. Ovals and rectangles represent sugar residues, whereas circles depict the polar head groups of glycerophospholipids. MDO represents membrane-derived oligosaccharides, and Kdo is 3-deoxy-D-manno-octulosonic acid. (After C. R. Raetz (1990) *Annu. Rev. Biochem.* 59, 129-70.)

on the other hand, has a unique constituent which is a complex glycolipid, i.e., lipopolysaccharide. As a result, the leaflets of the outer membrane are extremely asymmetrical. The membrane forms a highly effective permeation barrier around the cell, which prevents leakage of proteins, as well as protecting the content of the cell against attack by harmful agents like bile salts and hydrophobic antibiotics (5, 6). The physical organization and unusual function of this membrane is attributed to the presence of LPS. Exclusion of hydrophobic compounds in gram-negative bacteria is accomplished by surrounding the cells with hydrophilic polysaccharides, conversely, due to the lipid nature of LPS, the outer membrane can be expected to exclude hydrophilic compounds as well.

The interest in LPS arose approximately one hundred years ago when it was recognized that LPS plays an important role in the expression of antigenicity and pathogenicity of gram-negative bacteria (7, 8). Today LPS is known as the main antigen of gram-negative bacteria, carrying the antigenic determinants of O-specificity (9). In their membrane-associated location, LPSs are targets for bacteriophages (10), they harbor binding sites for antibodies and nonimmunoglobulin serum factors, and, thus, they are involved in the specific recognition and elimination of bacteria by the host organism's defense systems (11). On the other hand, LPS may function to prevent the activation of complement and uptake of bacteria by phagocytic cells and, by shielding pathogens from cellular host defenses, play an important role in bacterial virulence (12). In addition, LPS isolated or released from bacteria are endowed with a broad spectrum of biological activities such as pyrogenicity and lethal toxicity (13). When introduced into higher

organisms, LPS elicits fever and activates a series of immunological and biochemical events that lead to the mobilization of the host defense mechanisms; in large doses, LPS can cause irreversible shock and even death. To emphasize these activities, LPSs have been termed endotoxins (14), that is, they are the factors of pathogenicity responsible for many pathophysiological activities accompanying gram-negative infections. Also, because of their endotoxic properties, LPS contribute to the pathogenic potential of gram-negative bacteria. Paradoxically, the same endotoxins that threaten human health can enhance the body's overall immune resistance to bacterial or viral infections and cancer (15-17).

Chemically, LPSs are composed of three major structural domains, an O-specific polysaccharide, a core oligosaccharide with the inner and outer region, and a lipid A component (Figure 3) (2). These domains differ from one species to another and even among the strains of the same species. Much work has been devoted to the structural characterization of lipopolysaccharides. All three components, the O-specific polysaccharide, the core oligosaccharide, and the lipid A, have been exclusively studied.

The O-specific chain or O-antigen represents a carbohydrate polymer consisting of up to 50 oligosaccharide repeating units. The chemical structure of the O-antigen is determined by the repeating unit constituents and shows extreme diversity (21). The repeating units contain up to eight different, or in some cases identical, sugar residues that are generally interlinked by glycosidic bonds. The nature, ring form, type of linkage, and substitution of the individual monosaccharide residues, as well as their sequence within a
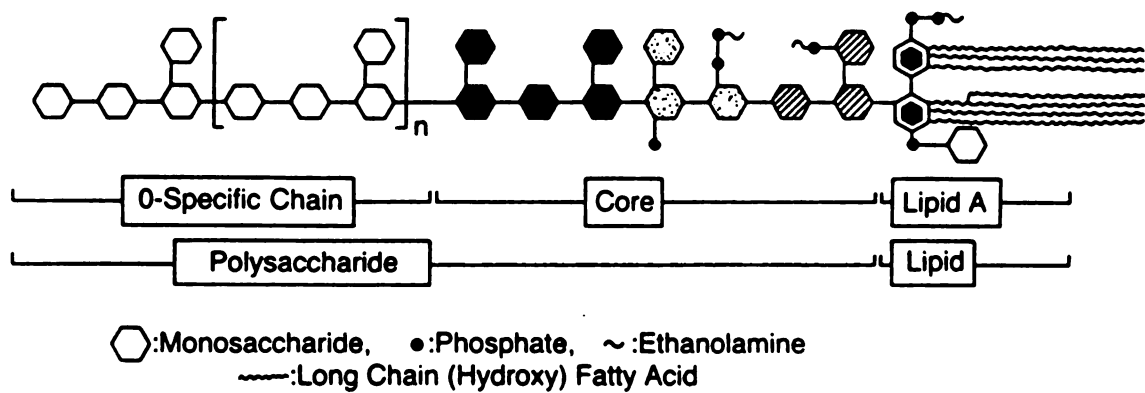
Figure 3. Schematic showing the structure of the LPS from *Salmonella typhimurium*. (After C. R. Raetz (1990) *Annu. Rev. Biochem.* 59, 129-701.)

repeating unit, is characteristic and unique for a given LPS and the parental bacterial serotype. A large diversity of sugars have been found in the repeating units of the O-antigen. These include neutral sugars, amino sugars and uronic acids. Moreover sugars can be substituted with charged function groups such as phosphate, amino acid, carboxylate, and ethanolamine triphosphate groups. Therefore, the O-antigen can contribute considerably to the net surface charge of the bacterial cell. The O-specific chain is the most variable segment. The number of repeating units can vary, even in a culture of one strain (18-20), from none to more than 40. This provides the cell with the opportunity for subtle variations in the molecular make up at different sites at its surface. The O-specific chain determines the serological specificity of LPS and the parental bacterial strains (22), that is, it carries the so-called O-factors that are located within the repeating units, which determines O-immunogenic and O-antigenic properties of LPS. Antibodies, enzymes, and complement proteins specifically interact with defined regions of the O-specific chain.

In contrast to the extreme diversity of the O-specific chains, the core region shows higher taxonomical restriction. The core oligosaccharide of LPS can be divided into an inner subdomain and an outer subdomain. The inner core includes two unusual sugars, 3-deoxy-D-manno-2-octulosonic acid (Kdo), and a heptopyranose occurring in both the L-glycero-D-manno and D-glycero-D-manno configurations (23, 24). The outer core contains a number of more common hexoses like D-glucose, D-galactose and N-acetyl-D-glucosamine in the pyranosidic ring form. The core provides attachment sites for lipid A and O-antigen.

Lipid A represents the covalently linked lipid component of LPS and anchors the LPS to the outer membrane. Lipid A is a highly conserved structure. The hexaacyl lipid A of *E. coli* (Figure 4) is a classical representative of enterobacterial lipid A. Its structure has been elucidated in great detail during last decade (1, 25, 26). The typical lipid A structure contains a β-D-glucosaminyl-(1-6)-α-D-glucosamine disaccharide which is phosphorylated in positions 1 (of the reducing glucosamine residue) and 4' (of the distal glucosamine residue). The 2-, 3-, 2'-, 3'-hydroxyl groups are usually substituted by 3-hydroxyltetradecanoyl groups with the R-configuration. Some of these 3-hydroxyltetradecanoyl groups are acylated at the 3-hydroxyl group by other fatty acids. There is some variation from species to species with relatively small variations in number, identity and location of the acyl substitutents. This also varies from species to species.

The structural role of lipid A as a hydrophobic anchor is provided by the abundant fatty acyl groups attached to the sugar. The presence of acyloxyacyl group further enhances the hydrophobicity of the lipid A. The presence of hydroxyl fatty acids might be involved in the intermolecular hydrogen binding. The phosphate groups would be expected to stabilize the outer membrane by cross-linking adjacent lipid A molecules via divalent cation. The unique structure of lipid A presumably reflects its specific roles in outer membrane assembly and function. As postulated since the early 1950s and proven during the last decade, lipid A contains the endotoxic principle of LPS and is responsible for the induction of various classical endotoxin effects (26, 27).
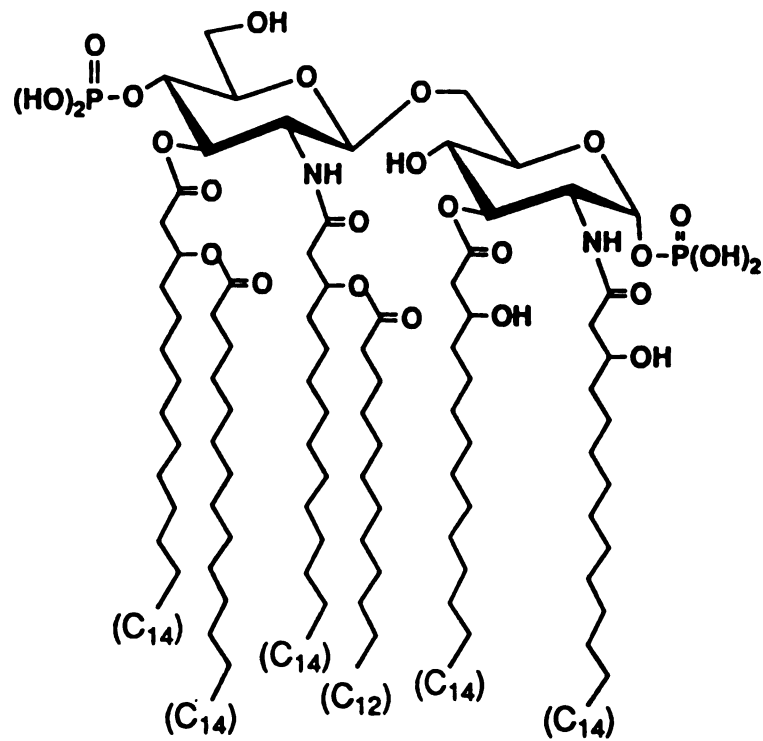
Figure 4. Structure of the typical lipid A obtained from the LPS *of Salmonella typhimurium* and *E. coli.* (After D. C. Morrison and J. L. Ryan (1987) *Ann. Rev. Med.* 38, 417-32.)

Like other amphiphilic molecules, e.g. phospholipids, LPS and free lipid A aggregate into supramolecular structures above a certain concentration, the so-called critical micellar concentration (CMC). Amphiphilic molecules, in general, consist of a hydrophilic polar headgroup and hydrophobic apolar hydrocarbon chains. Aggregation is promoted by hydrophobic interaction or, more precisely, the minimization of the Gibbs free energy of the water-amphiphile system, which is accomplished by the tendency of free water to increase its entropy. The three-dimensional supramolecular structures depend on the primary chemical structure of the amphiphile, and on its secondary structure, which corresponds to the molecular shape of the individual molecules. The correlation between the molecular shape of the individual amphiphilic molecules and the supramolecular structure of aggregates of a large number of identical molecules is schematically outlined in Figure 5 (28, 29).

LPSs may assume various three-dimensional structures depending on their molecular structures and on ambient conditions such as water content, temperature, and counterion concentration (30, 31). The supramolecular structures of LPS and free lipid A have been investigated by employing various techniques, i.e., electron spin resonance (ESR), 31P-NMR, freeze-fracture electron microscopy, infrared spectroscopy, fluorescence spectroscopy, calorimetry, neutron diffraction and small- and wide-angle X-ray diffraction (32-36). A broad spectrum of three-dimensional structures of LPS and lipid A has been proposed, ranging from micellar over lamellar to inverted states. As summarized in a review by Seydel and Brandenberg (28), isolated LPS and free lipid A of *S. minnesota* and *E. coli* form predominantly inverted hexagonal aggregates in the liquid

Figure 5. Correlation between the molecular shape of lipids and the three-dimensional supramolecular structures they form. (After U. Seydel and K. Brandenbyrg (1992) in *Bacterial Endotoxic Lipopolysaccharides Vol. 1*, D. C. Morrison and J. L. Ryan, Eds., CRC Press, Inc., pp 225-250.)

A   Micellar

B   Normal
         hexagonal (H$_I$)

C   Lamellar (L)

D   Inverted
         micellar

E   Cubic (Q)

F   Inverted
         hexagonal (H$_{II}$)
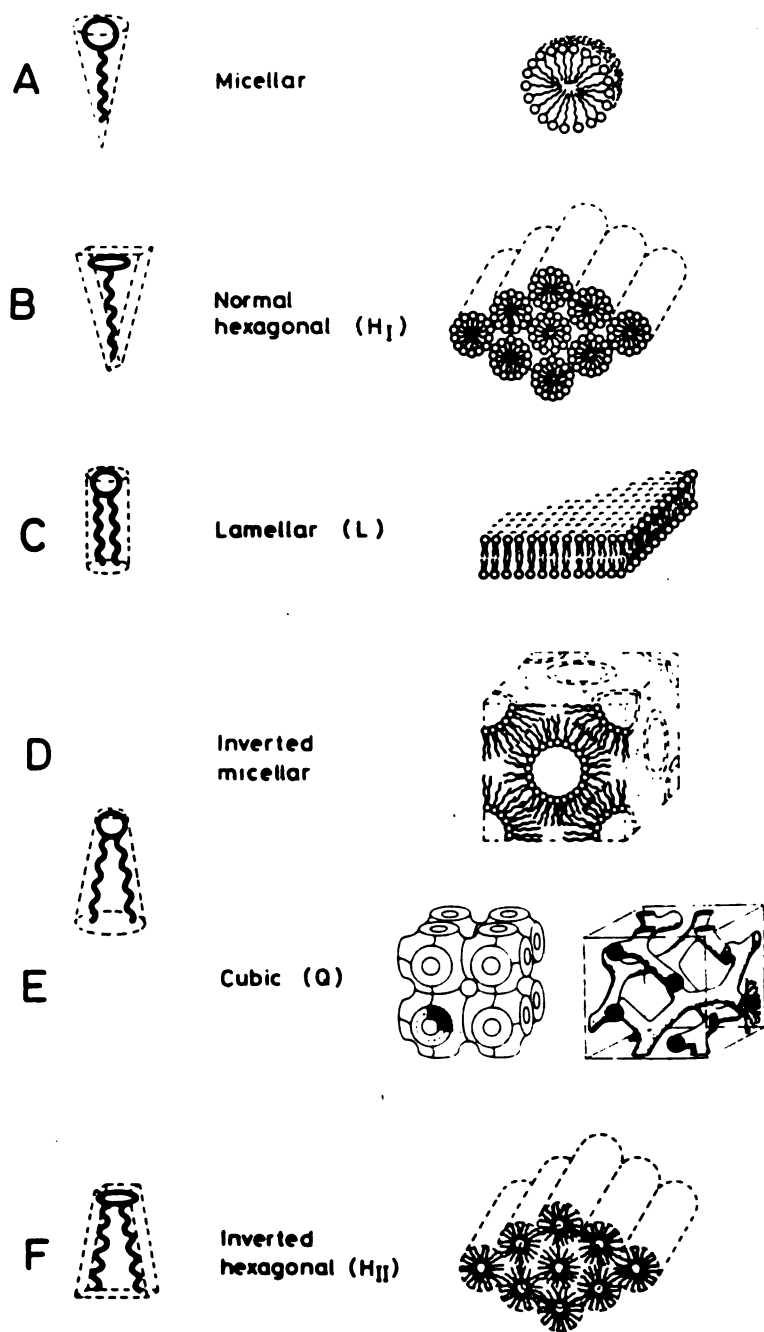
Figure 5

crystalline state. In the gel state, they may adopt cubic phases or lamellar phases depending on the water content and cation concentrations. Under near physiological conditions of water content, divalent cation concentration, and temperature, they form exclusively cubic structures. These results do not seem to convey a unique picture of the supramolecular structure of LPS. This is not only due to the different chemical structures investigated, but may also be explained by the different experimental conditions under which the samples were investigated.

However, dried films of LPS and lipid A samples have been observed forming a very simple phase behavior with only one ordered phase, which was a lamellar structure with hexagonal dense packing of the acyl chains. This conclusion was obtained by Labischinski et al. and several other groups (33, 37-39) by applying X-ray diffraction techniques. Small-angle X-ray and neutron diffraction have been employed for the determination of the long-range order (supramolecular conformation) and wide-angle X-ray diffraction for the determination of the short-range order (arrangement of the acyl chains). The investigated compounds include LPS and lipid A samples from *several E. coli* and *S. minnesota* strains, and synthetic compounds corresponding in their chemical structures to those of lipid A from *E. coil* and from *S. minnesota*. In all cases, they form two dimensional hexagonal arrays when laid down on a suitable substrate as a film.

It is well known that the supramolecular architecture of LPSs and lipid A are critical determinants of their biological activities. Despite this, the highly ordered LPS and lipid A self-assemblies are attractive biomaterials for various potential applications.

The construction of large molecular systems containing repetitive elements and certain symmetry properties have grasped the attention of organic chemists and material scientists. However, such arrays are difficult to construct in an *ab initio* method using even the most advanced synthetic tools, the most productive path seems to be to look for nature and try to understand them with the ultimate aim of either copying some aspects or using them directly.

Most previous research has been focused on enterobacterial or typical lipid A and LPS. These studies, however, did not shed much light on the molecular basis for the supramolecular structures, since lipid A regions of the investigated LPSs are very similar. It is, therefore, important to study the structures, both molecular and supramolecular, of unusual LPS for two major reasons. Firstly, if the chemical structures are very different, do they have the same supramolecular structures of the typical bacteria? If they do, this will give us insight into the molecular basis for the arrangement. The second reason is that atypical bacteria might have very different symmetry in their supramolecular lattices, e.g., cubic instead of hexagonal. This could be a very important finding since it would give us more flexibility in the design of periodic arrays. The final important point to be addressed is: what influence does the structure of the polysaccharide region have on the properties and nature of the LPS supramolecular array? Answers to these questions will determine the extent to which we can modify the surface of the LPS array by attaching structures such as molecular probes, fluorescence probes and other pendant groups.

The lipopolysaccharides, especially the lipid A regions of *Rhizobium trifolii*, which are nitrogen-fixation bacteria forming symbiotic relationship with legume plants, have radically different structure from the typical one of enterobacteriaceae. *R. trifolii* lipid A was found to be completely devoid of phosphate (40-42) and it has been suggested that the lipid A portion of *Rhizobium* is essentially different from that of the enterobacteriaceae and that it has a different backbone. Previous work in our lab has suggested that the *R. trifolii* lipid A might contain a disaccharide of galacturonic acid and glucosamine. Further chemical analysis of the fatty acids on *R. trifolii* 0403 lipid A by Hollingsworth et al. (43) showed the presence of 3-hydroxylated fatty acids exclusively, and an unusual 27-hydroxyoctacosanoic acid in the LPS (44).

Characterization of the chemical structure of *Rhizobium trifolii* LPS has been the major task of this research project. Detailed structural analyses have been performed on LPS from two related strains, *R. trifolii* 4S and *R. trifolii* ANU843. The chemical structures of all three components, the O-specific polysaccharide, the core oligosaccharide, and the lipid A region, have been determined by a combination of chemical and spectroscopic methods. The knowledge of the chemical structure of this unusual LPS, with further analysis on its three-dimensional self-assembly structure, will provide us a clear insight into the molecular basis for the supramolecular structure of LPS.

REFERENCES

1. Raetz, C. R. *Annu. Rev. Biochem.* **1990**, 59, 129-70.

2. Luderitz, D. S. *Current Topics in Membranes and Transport* **1992**, 17, 79-151.

3. Gennis, R. B. *Biomembranes: Molecular Structure and Function*; Springer-Verlag New York Inc., 1989; pp 154-158.

4. Griffs, J. M.; Schneider, H.; Mandrell, R. E.; Yamasaki, R.; jarvis, G. A.; Kim, J. J.; Gibson, B. W.; Hamadeh, R.; Apicella, M. A. *Rev. Infect. Dis.* **1988**, 10, 287.

5. Lugtenbeg, B.; Alphen, L. V. *Biochimica et Biophysica Acta* **1983**, 737, 51-115.

6. Rietschel, E. T.*Cellular and Molecular Aspects of Endotoxin Reactions*; Elseiver Science Publishers: Amsterdam, 1990; pp. 15-32.

7. Morrison, D. C.; Ryan, J. L. In *Bacterial Endotoxic Lipopolysaccharides, Vol. I. Molecular Biochemistry and Cellular Biology*; Morrison, D. C., Ryan, J. L. Eds.; CRC press: Boca Raton, 1992.

8. Rieschel, E. T.; Brade, H. *Scientific American* **1992**, 267, 54-61.

9. Westphal, O.; Jann, K.; Himmelspach, K. *Prog. Allergy* **1983**, 33, 9-39.

10. Lindberg, A. A.; Wollin, R.; Bruse, G.; Ekwall, E.; Svenson, S. B. *Immunology and immunochemistry of synthetic and semisynthetic Salmonella O-antige specific glycoconjugates*; ACS Sym, Ser. **1983**, 231, 83.

11. Sharon, N. *Complex Carbohydrates: Their Chemistry, Biosythesis, and Functions*; Addison-Wesley Publishing Company, Inc. 1975; pp. 328-336.

12. Makela, P. H.; Hovi, M.; Saxen, H.; Muotiala, A.; Rhen, M. In *Cellular and Molecular Aspects of Endotoxin Reactions Vol. 1*; Nowotny, A., Spitzer, Y., Ziegler, E. J., Eds.; Elsevier Science: Amsterdam, 1990; pp. 537.

13. Morrison, D. C.; Ryan, J. L. *Ann. Rev. Med.* **1987**, 38, 417-32.

14. Hinshaw, L. B. *Handbook of Endotoxin*, Vol 2: *Pathophysiology of Endotoxin*; Amsterdam: Elservier; 1984.

15. Nowotny, A. *Rev. Infect. Dis.* **1987**, 9, 503.

16. Ribi, E.; Amano, K.; Cantrell, J.; Schwartzman, S.; Parker, R.; Takayama, K. *Cancer Immunol. Immuother.* **1982**, 12, 91-96.

17.   Takayama, K.; Ribi, E.; Cantrell, J. L. *Cancer Res.* **1981**, 41, 2645.

18.   Goldman, R. C.; Leive, L. *Eur. J. Biochem.* **1980**, 107, 145-153.

19.   Munford, R. S.; Hall, C. L.; Rick, P. D. *J. Bacteriol.* **1980**, 144, 630-640.

20.   Palva, E. T.; Makela, P. H. *Eur. J. Biochem.* **1980**, 107, 137-143.

21.   Orskov, I.; Orskov, F.; Jann, B.; Jann, K. *Bacteriol. Rev.* **1977**, 41, 667-710.

22.   Luderitz, O.; Freudenferg, M. A.; Galanos, C.; Lehman, V.; Rietschel, E. T.;
       Shaw, D. H. In *Microbial Membrane Lipids Vol. 17*; Razin, S., Rottem, S., Eds.;
       Academic Press: New York, 1982; pp. 79.

23.   Rietschel, E. T.; Brade, L.; Lindner, B.; Zahringer, U. In *Bacterial Endotoxic
       Lipopolysaccharides, Vol. I. Molecular Biochemistry and Cellular Biology*;
       Morrison, D. C., Ryan, J. L., Eds.; CRC Press: Boca Raton, 1992; pp.3-41.

24.   Holst, O.; Brade, H. In *Bacterial Endotoxic Lipopolysaccharides, Vol. I.
       Molecular Biochemistry and Cellular Biology*; Morrison, D. C., Ryan, J. L., Eds.;
       CRC Press: Boca Raton, 1992; pp. 135-170.

25.   Rietschel, E. T. In *Handbook of Endotoxin, Vol. 1: Chemistry of Endotoxin*;
       Rietschel, E. T., Ed.; Elsevier Science: Amsterdam, 1984; pp 187-220.

26.   Imoto, M.; Kusumoto, S.; Shiba, T.; Naoiki, H.; Iwashita, T.; Rietschel, E. T.;
       Wollenweber, H. W.; Galanos, C.; Luderitz, O. *Tetrahedron Lett.* **1983**, 24, 4017-
       4020.

27.   McCartney, A. C.; Wardlaw, A. C. In *Immunology of Bacterial Cell Envelope*;
       Stewart, D. E., Davis, M., Eds.; John Willey and Sons Ltd.:Chichester, United
       Kindom, 1985; pp. 208-238.

28.   Seydel, U.; Brandenburg, K. In *Bacterial Endotoxic Lipopolysaccharides Vol I.
       Molecular Biochemistry and Cellular Biology*; Morrison, D. C., Ryan, J. L., Eds.;
       CRC Press Inc.: Boca Raton, 1992; pp. 225-250.

29.   Israelachvili, J.; Marcelja, S.; Horn, R. G. *Q. Rev. Biophys.* **1980**, 13, 121.

30.   Alphen, L. V.; Verkleij, A.; Burnell, E.; Lugtenberg, B. *Biochimica et Biophysica
       Acta* **1980**, 597, 502-517.

31.   Shands Jr., J. W.; Graham, J. A; Nath, K. *J. Mol. Biol.* **1967**, 25, 15-21.

32.   Seydel, U.; Brandenburg, K.; Koch, M. H.; Rietschel, E. T. *Eur. J. Biochem.*
       1989, 186, 325-332.

33.   Labischiski, H.; Barnickel, G.; Bradaczek, H.; Naumann, D.; Rietschel, E. T., Giesbrecht, P. *J. Bacteriol.* **1985**, 162, 9-20.

34.   Brandenburg, K.; Seydel, U. *Biochimica et Biophysica Acta* **1984**, 775, 225-238.

35.   Brandenberg, K.; Koch, M. H.; Seydel, U. J. *J. Struct. Biol.* **1990**, 105, 11.

36.   Coughlin, R. T.; Peterson, A. A.; Haug, A.; Pownall, H. J.; McGroarty, E. J. *Biochem. Biophys. Acta* **1985**, 821, 404-412.

37.   Emmerling, G.; Henning, U.; Gulik-Krzywicki, T. *Eur. J. Biochem.* **1977**, 78, 503.

38.   Hayter, J. B.; Rivera, M.; McGroarty, E. J. *J. Biol. Chem.* **1987**, 262, 5100.

39.   Labischinski, H.; Nauman, D.; Schultz, C.; Kusumoto, S.; Shiba, T.; Rietschel, E. T.; Giesbrecht, P. *Eur. J. Biochem.* **1989**, 179, 659.

40.   Hollingsworth, R. I.; D. Lill-Elghanian, D. in *Cellular and Molecular Aspects of Endotoxin Reactions*; Nowotny, A., Spitzer, J. J., Ziegler, E. J., Eds.; Elsevier Publishers: Amsterdam, 1990; pp. 73-84.

41.   Zevenhuizen, L.; Scholten-Koerselman, I.; Posthumus, M. *Arch. Microbiol.* **1980**, 125, 1-8.

42.   Humphrey, B.; Vincent, J. *J. Gen. Microbiol.* **1969**; 59, 411-425.

43.   Hollingsworth, R. I.; Dazzo, F. B. *Anal. Microbiol.* **1988**, 7, 295-302.

44.   Hollingsworth, R. I.; Carlson, R. W. *J. Biol. Chem.* **1989**, 264, 9300-9304.

CHAPTER 2

THE STRUCTURE OF THE O-ANTIGENIC CHAIN OF THE

LIPOPOLYSACCHARIDE OF *RHIZOBIUM TRIFOLII* 4S

ABSTRACT

The structure of the O-antigen chain of the lipopolysaccharide (LPS) of *Rhizobium trifolii* 4S has been determined by a combination of chemical and spectroscopic methods. The glycosyl components were found to be L-rhamnose, N-acetyl-D-glucosamine and N-acetyl-D-mannosamine in 3 : 1 : 1 molar proportion, as determined by gas chromatography and gas chromatography-mass spectrometry of alditol acetate and persilylated (R)-2-hydrxoxytbutyl glycoside derivatives. The linkage positions and configurations of the glycosyl residues were obtained by 1D and 2D NMR spectroscopy. The polymer has a pentasaccharide repeating unit containing rhamnose and N-acetyl glucosamine in the main chain and N-acetyl mannosamine as the sole side chain component. This latter residue is linked to a main chain rhamnose residue. This result was suggested by NMR spectroscopy and confirmed by periodate oxidation. The sequence was deduced by 1D and 2D NMR NOE experiments and by partial hydrolysis studies. The repeating unit of the polysaccharide is shown. This constitutes the first complete structure of an O-antigenic chain of the lipopolysaccharide of any strain of *Rhizobium trifolii*.

α-D-ManNAc

1

$\downarrow$

2

→3)-α-L-Rha-(1→3)-α-L-Rha-(1→4)-β-D-GlcNAc-(1→3)-α-L-Rha-(1→
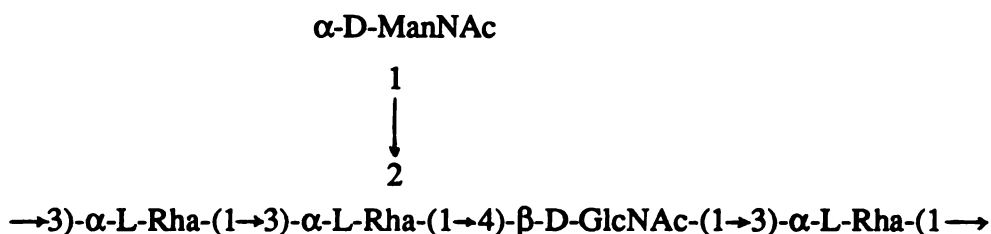
INTRODUCTION

The lipopolysaccharides of *Rhizobium* are known to be involved in critical aspects of bacteroid development and nodule occupancy. Several years ago it was demonstrated (1-3) that bacteria having defects in lipopolysaccharide structure (as judged by compositional analyses or electrophoretic properties) also have defects in their ability to succesfully complete the earlier aspects of nodule invasion. Such bacteria were not released from infection threads and the nodules formed from their infection were invariably unoccupied. More recently, a more defined structural basis for the inability of these "rough" bacterial mutants to form viable nodules was advanced when it was demonstrated that such mutants contained lipopolysaccharides that had no O-antigen and had truncated or different core components (4-6).

Although there has been much progress in elucidating the structures of the R-core components of the lipopolysaccharides of the *Rhizobiacae* (7-9), there has been considerably less success in elucidating the structres of the O-antigen components. In fact, no complete structures have yet been proposed for the O-antigen of any of the "fast growing" strains. Here we describe the complete structure of the O-antigen chain of *R. trifolii* strain 4S. This is a much studied strain for which there is a considerable amount of biochemical and physiological data (10, 11). These studies will facilitate biochemical studies of the roles of the O-antigen of this strain in bacteroid development. Questions about whether this specific O-antigen structure is formed *in planta* can now be answered.

MATERIALS AND METHODS

O-antigen Isolation. *Rhizobium trifolii* 4S were grown at 30°C in modified Bergensen's (BIII) medium. The LPS was isolated as described previously (12). A sample (10 mg) of LPS was hydrolyzed with 2 ml of 1% acetic acid for 3 hours at 100°C and then extracted several times with an equal volume of 5:1 chloroform/methanol. The aqueous fraction containing the polysaccharide was dried under a stream of nitrogen. The low-molecular-weight components derived from the R-core were selectively removed by acetylating the carbohydrate mixture with 1:1 pyridine/acetic anhydride. This effectively removed the low-molecular-weight core oligosaccharides without altering the O-antigenic polysaccharide, as the latter is insoluble in this acetylating mixture. This polymer was found by chromatographic analyses to be one homogeneous component and was used for all of the subsequent studies.

Compositional Analysis. The polysaccharide was converted to alditol acetate derivatives. A sample of polysaccharide (0.1 mg) was treated with 3 ml of 1% hydrochloric acid in methanol. The solution was sonicated for 1 min, heated for 15 min at 68°C, and then kept overnight at room temperature. The solution was then dried under a stream of nitrogen. To the dried sample was added 5 mg of sodium borodeuteride and 0.3 ml of 1:1 methanol-water. The solution was kept overnight at room temperature, and then dried under nitrogen and treated with 2 ml of 2 M trifluroacetic acid for 1.5 hours at 115°C. After cooling to room temperature, the solution was concentrated to dryness under nitrogen. A few drops of water were added and the solution was dried again under nitrogen with heating in a water bath (50°C). This addition of water and drying was repeated twice in order to remove traces of acid. The sample was then dissolved in 0.2 ml of water, and a solution of 2 mg of sodium borohydride in 0.1 ml of water was added dropwise. The reaction mixture was kept for 2 hours at room temperature, and then 2.4 M hydrochloric acid was added to the solution dropwise until effervescence ceased. The

solution was dried under nitrogen and to the residue was then added 1 ml of 2% acetic acid/methanol. This solution was then evaporated under nitrogen again. This last step was repeated 5 times. The dried product was per-acetylated by treatment with 1 ml of pyridine, followed by 2 min of sonication and then 1 ml of acetic anhydride. The solution was heated for 1 hour at 70°C with occasional sonication and then dried under nitrogen. The acetylated product was partitioned between 1 ml of saturated aqueous sodium chloride solution and 3 ml of chloroform. The chloroform fraction was removed and dried under nitrogen with no heat and subjected to GC analysis on a Hewlet Packard 5890 gas chromatograph using a DB225 capillary column with helium as the carrier gas. A temperature program with an initial temperature of 180°C holding for 2 min, then increasing to 230°C at the rate of 2°C/min with a final hold of 60 min was employed. GC-MS analysis was then performed on a JEOL 505 mass spectrometer system using electron impact ionization (70 eV) and detecting in the positive mode.

Determination of D and L Configuration. The monosaccharides were converted into their trimethylsilylated (-)-2-butyl glycoside derivatives (13), and these were subjected to GC analysis. A sample of polysaccharide (0.2 mg) was treated with 2 ml of 2M trifluroacidic acid for 1.5 hours at 120°C. The hydrolyzed product was then dried under a stream of nitrogen and acetylated as already described. The acetylated product was then dried under nitrogen and treated with 0.2 ml of R-2-(-)butanolic hydrochloride, which was made with R-(-)-2-butanol and acetyl chloride in a ratio of 10:1 (v / v). After butanolysis for 9 hours at 80°C, the solution was neutralized with silver oxide. The mixture was kept for half at room temperarure an hour to allow the silver salts to precipitate. The supernatant solution was removed and dried under nitrogen, and then treated with 5.0 μl of 1.5 mequiv/ml N-trimethylsilylimidazole in silylation grade pyridine for 15 min at 60°C. The trimethylsilylated mixture was injected into a GC column directly. The GC analysis was performed on a Hewlet Packard 5890 gas chromatograph equipped with a DB1 capillary column with helium as the carrier gas. The temperature was

programed from 150°C to 300°C at 3 °C/min. GC-MS analysis was performed at the same instrument as already described.

Periodate Oxidation. A sample of polysaccharide (0.1 mg) was treated with 0.5 ml of 0.2 M sodium periodate solution in the dark for 48 hours at 10°C. The solution was then desalted on a TSK 2000 gel filtration HPLC column, using water as eluant and refractive index detection. Fractions corresponding to peaks were collected and the sole carbohydrate containing fraction was identified by the phenol-sulphuric acid assay (14). The oxidized polysaccharide was hydrolyzed with 0.2 M trifluoroacetic acid at 55°C for 10 min. A sample of the polysaccharide thus oxidized and purified was oxidized again as already described. Both oxidized ploysaccharide products were converted to alditol acetate derivatives and analyzed on GC as already describled.

Partial Hydrolysis. A sample of polysaccharide (0.1 mg) was treated with 2 ml of 44% formic acid overnight at 70°C. The solution was dried under nitrogen and then about 50 ml of water was added. The solution was dried under nitrogen again to ensure complete removal of formic acid. The partially hydrolyzed products were then per-acetylated as already described and subjected to fast atom bombardment mass spectrometric (FAB MS) analysis. 4-Nitrobenzyl alcohol was used as the matrix.

NMR Spectroscopy. All NMR spectra were measured in D_2O at 500 MHz for ^{1}H or 125 MHz for ^{13}C with a VARIAN VXR500 spectrometer. For the HMQC experiment (15), a spectral width of 25740 Hz was employed for the ^{13}C dimension. A total of 64 transients were acquired at 1024 points each. A total of 256 data sets were acquired. Double quantum filtered J-correlated 2-dimensional spectrum (phase sensitive mode) (16) was obtained using a total of 256 data sets (32 transients at 2048 data points each). Data for the HOHAHA experiments (17) were obtained using similar acquisition and processing conditions. A mixing time of 100 ms was used for both the HOHAHA and NOESY experiments. The NOESY spectrum (18) was obtained using a total of 600 data

sets with 32 transients at 2048 data points each. The spectrum obtained over a spectral width of 4329 Hz. The NMR sample of the polysaccharide was purged with helium for half an hour before the NOESY experiment was performed. The water line was supressed by presaturation. All 1D and 2D proton NMR spectra were recorded at 70°C and the proton chemical shifts were referenced to the water line at 4.25 ppm. This was established using an external reference. The HMQC spectrum was recorded at 50°C and the ^{13}C and ^{13}C-DEPT spectra were recorded at room temperature. The carbon chemical shifts were referred to external standard $CDCl_3$ signal at 77.0 ppm.

RESULTS AND DISCUSSION

GC (Figure 1) and GC-MS analyses on the alditol acetate derivatives of the polysaccharide revealed the presence of 3 glycosyl components: rhamnose, N-acetyl glucosamine and N-acetyl mannosamine in 3 : 1 : 1 molar ratio. Their absolute configurations were determined to be L-rhamnose, D-glucosamine and D-mannosamine (Figure 2, 3). The ^{1}H-NMR spectrum of the polysaccharide (Figure 4) showed five anomeric proton resonances in the region of 4.7-5.2 ppm. The doublet at 4.74 ppm (J = 8.0 Hz) was assigned to a β- anomeric proton of N-acetyl glucosamine in the pyranose form. The large coupling constant arises from the $J_{1,2}$ axial-axial coupling. The other four anomeric proton signals were singlets, indicating that these residues had the α-configuration. Three doublets (J = 6.5 Hz) between 1.15 ppm and 1.25 ppm were assigned to the 6-deoxy groups of three rhamnosyl residues. Two singlets at 2.00 ppm and 2.04 ppm were assigned to the protons of N-acetyl groups that arise from N-acetyl glucosamine and N-acetyl mannosamine residues. Similar analysis on samples that were not subjected to acetic anhydride-pyridine treatment were N-acetylated to the same extent. Signals between 3.2-4.6 ppm were assigned to the rest of the sugar protons.
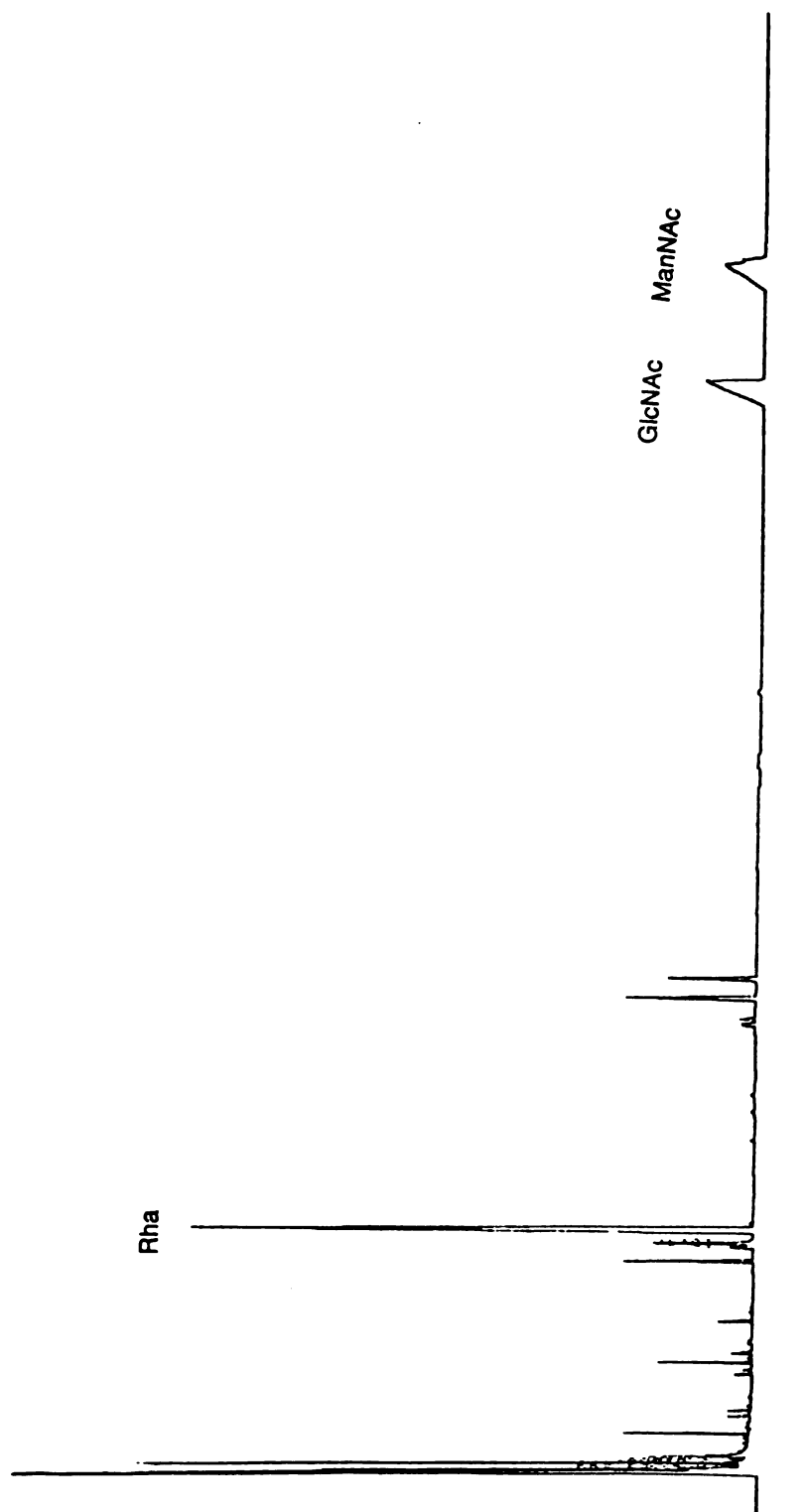
Figure 1. GC profile of the alditol acetate derivatives of the polysaccharide. Rhamnose (Rha), N-acetylglucoamine (GlcNAc), and N-acetylmannosamine (ManNAc) are present in 3 : 1: 1 molar ratio.

Figure 2. (A) ^{1}H NMR spectrum of the hydrolyzed products of the polysaccharide after treated with TFA for 1.5 hrs at 120°C. (B) ^{1}H NMR spectrum of a standard mixture of L-rhamnose, D-glucosamine, and D-mannosamine in 3 : 1 : 1 molar ratio. Note that the two spectra are nearly identical, and this confirm the composition of the polysaccharide and the absolute configurations of the monosaccharide components.
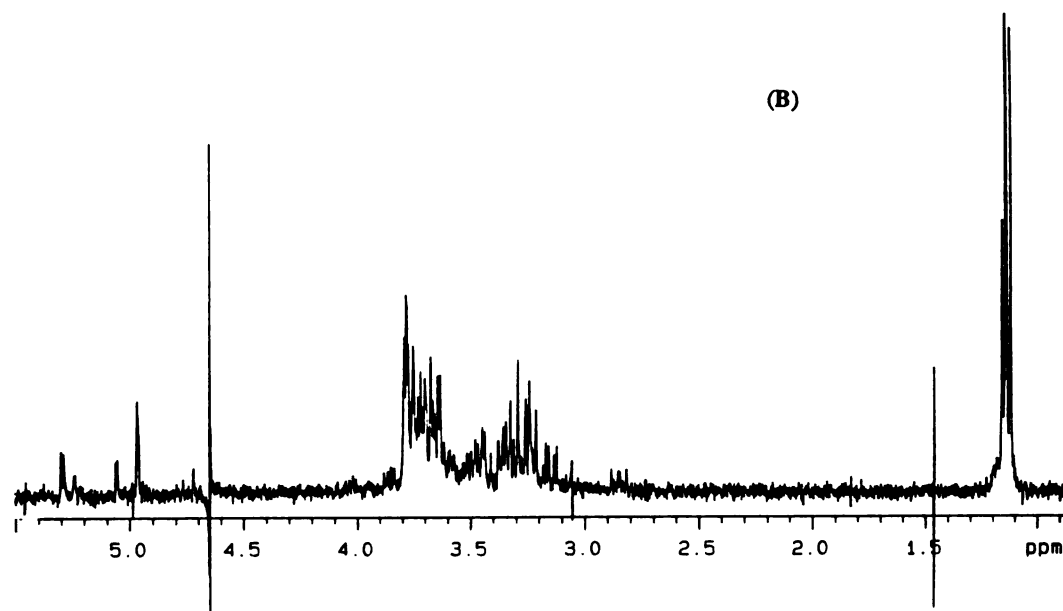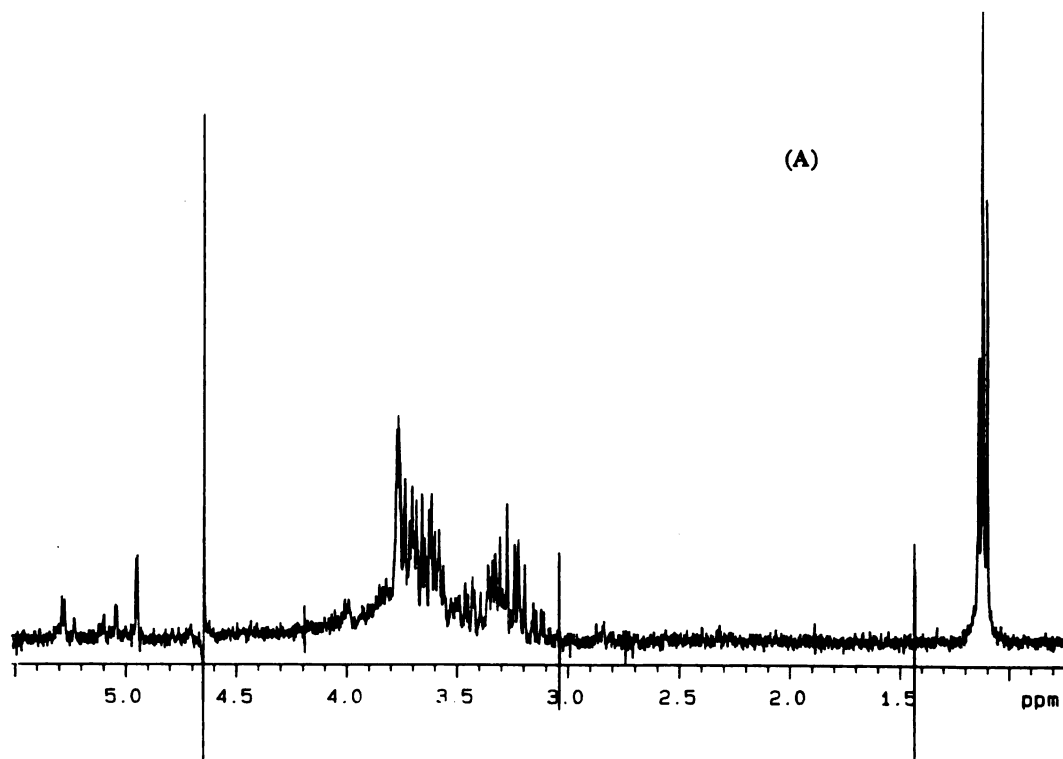
Figure 2

Figure 3. GC profiles of the trimethylsilylated (-)-2-butyl glycoside derivatives of (A) the polysaccharide, and (B) the standard mixture of L-rhamnose, D-glucosamine, and D-mannosamine in 3 : 1 : 1 molar ratio. Note that the signal patterns in both chromatograms are the same, and this confirm the absolute configurations of the sugar components.
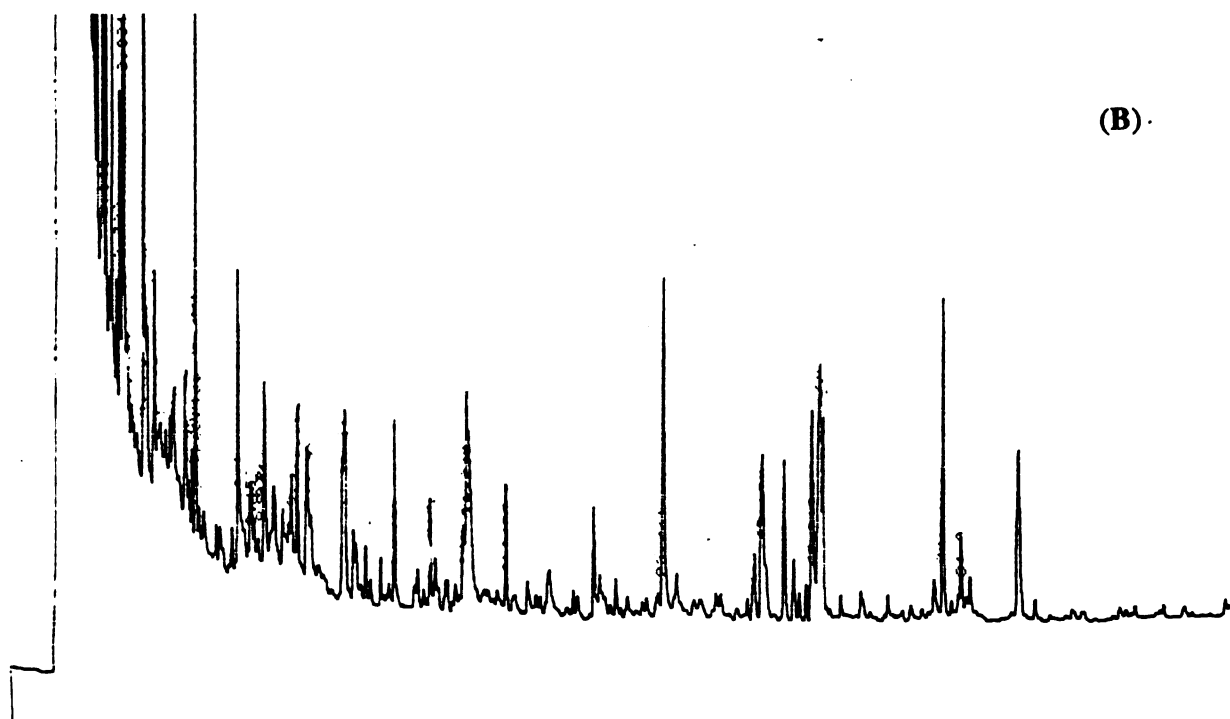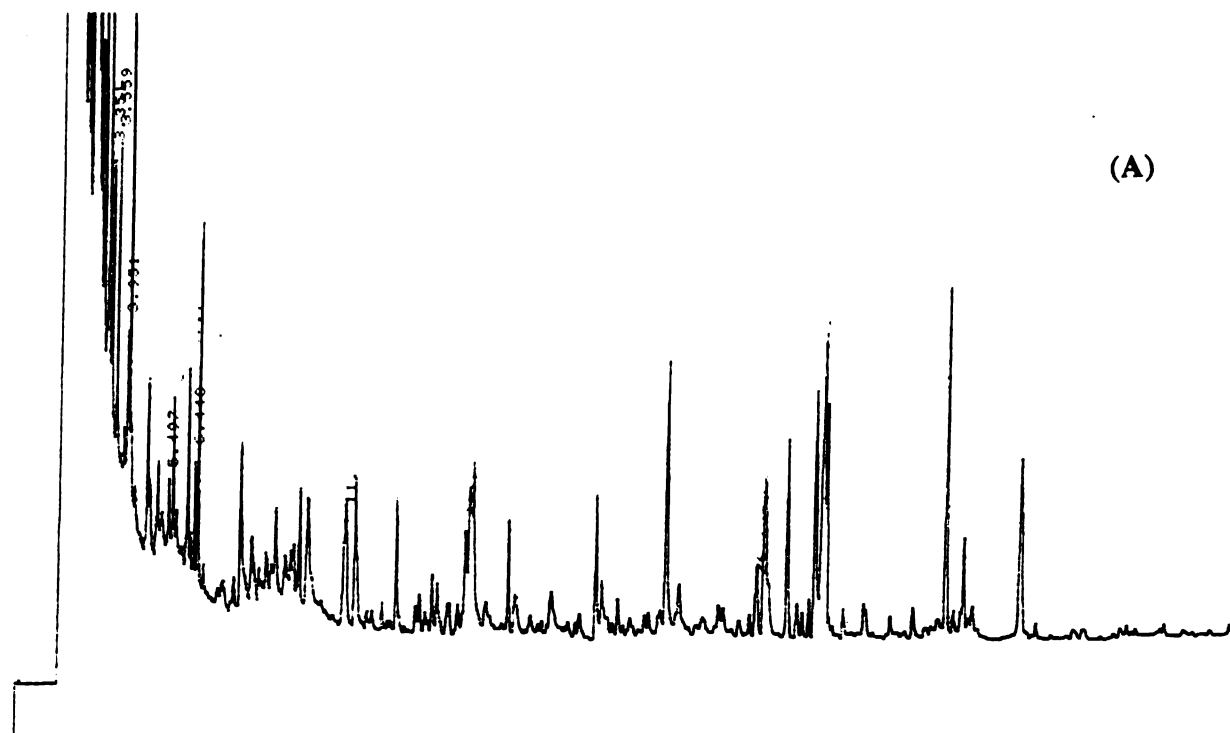
Figure 3

Figure 4. ^{1}H NMR spectrum of the polysaccharide. There are five anomeric proton resonances between 4.7 ppm and 5.2 ppm. Two singlets at 2.0 ppm and 2.04 ppm are due to the N-acetyl groups of the N-acetyl glucosamine and N-acetyl mannosamine residues. Three doublets between 1.15 ppm and 1.25 ppm are assigned to the 6-deoxy groups of the three rhamnosyl residues. Signals between 3.2 ppm and 4.6 ppm are due to the rest of sugar protons. R_1-1 refers to the anomeric proton of the rhamnosyl residue 1. R_2-1 and R_3-1 refer to that of rhamnosyl residue 2 and 3, respectively. M-1 ang G-1 refer to the anomeric proton signals of N-acetyl mannosamine and N-acetyl glucosamine, respectively.
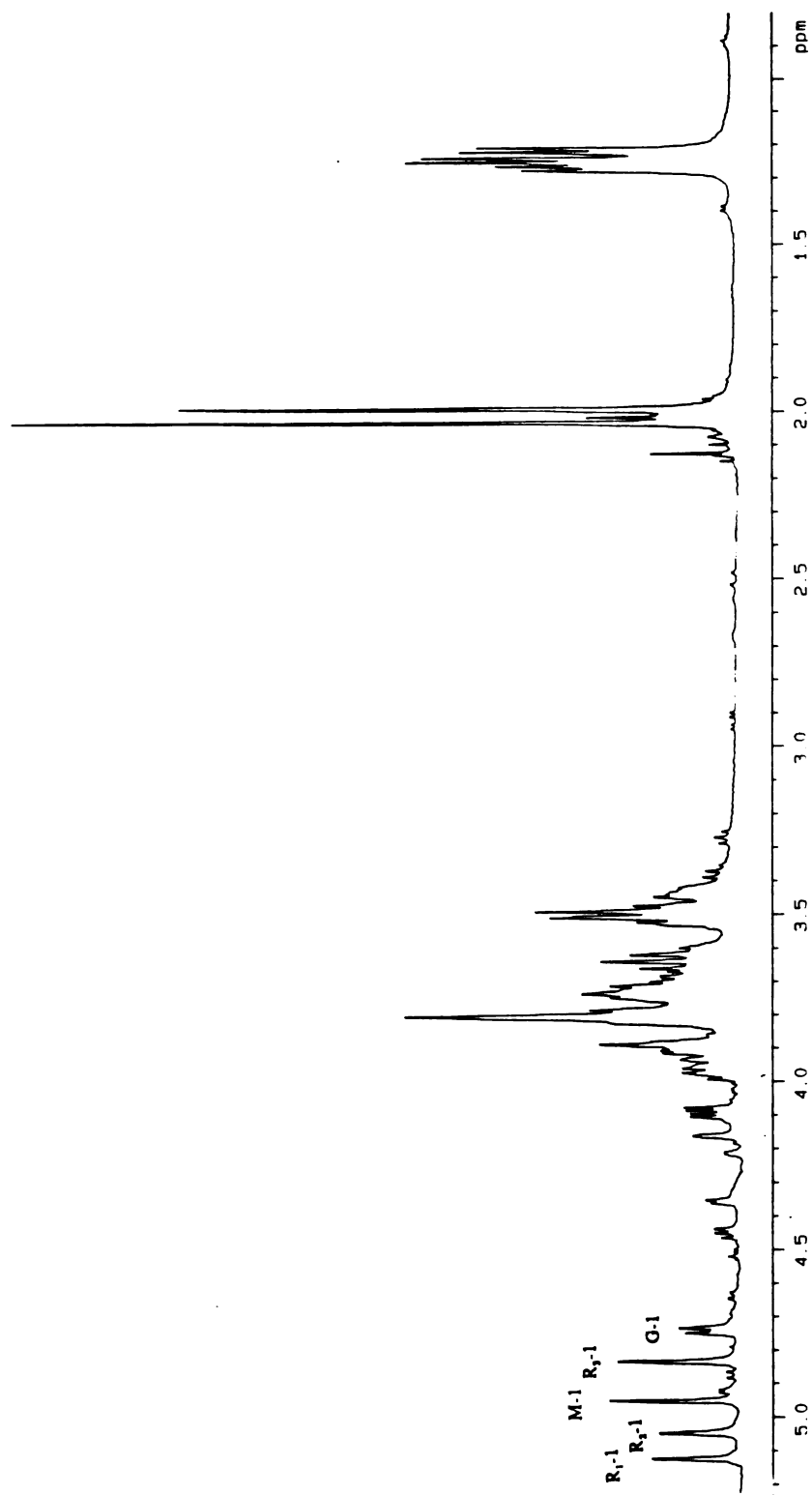
Figure 4

The ^{13}C-NMR spectrum of the polysaccharide (Figure 5) showed 34 signals. Two of them (at 173.8 ppm and 174.3 ppm) corresponded to carbonyl carbons. Five anomeric carbon signals appeared in the region of 94-102 ppm. The ^{13}C-DEPT spectrum (Figure 6) showed the presence of 5 methyl carbons, 2 methylene carbons and 25 methine carbons. Three methyl carbon signals at 16.0 ppm, 16.4 ppm and 16.4 ppm were assigned to the three 6-deoxy carbons of rhamnosyl residues. Two methyl carbon signals at 21.6 and 21.8 ppm were assigned to the acetyl moieties of N-acetyl glucosamine and N-acetyl mannosamine. Two methine carbon signals at 52.4 ppm and 55.2 ppm were assigned to the 2-positions of the amino sugars. The two methylene carbon signals at 60.0 and 60.6 ppm were assigned to the C-6 CH_2OH groups in N-acetyl glucosamine and N-acetyl mannosamine. Furthermore, the chemical shifts of these resonances indicated that these hydroxyl groups were unsubstituted. These 1D NMR results, combined with GC data, suggested that the polysaccharide consisting of a pentasaccharide repeating unit containing three rhamnosyl (Rha), one N-acetyl glucosamine (GlcNAc) and, one N-acetyl mannosamine (ManNAc) residues.

The ^{1}H NMR spectrum agreed well with the ^{13}C NMR spectrum, as demonstrated by the ^{1}H-^{13}C HMQC NMR spectrum (Figure 7). The five anomeric proton signals correlated with the five anomeric carbon signals. The proton signal at 4.75 ppm correlated with the ^{13}C signal at 101.5 ppm and was assigned to the anomeric proton of N-acetyl glucosamine. The ^{13}C signal at 95.0 ppm was assigned to the anomeric carbon of N-acetyl mannosamine. The ^{1}H-^{13}C correlated pair of signals at 4.95 ppm and 95.0 ppm were thus assigned to the anomeric position of N-acetyl mannosamine. The ^{1}H-^{13}C correlated pair of signals at 5.13 ppm and 101.1 ppm were assigned to the anomeric position of rhamnose (Rha1). The ^{1}H-^{13}C correlated pair of signals at 5.05ppm and 100.5ppm were assigned to the anomeric position of another rhamnosyl residue (Rha2). The ^{1}H-^{13}C correlated pair of signals at 4.84 ppm and 100.9 ppm were assigned to the other rhamnose moiety (Rha3) of the polysaccharide. The ^{13}C signals around 80 ppm were assigned to the carbons involved
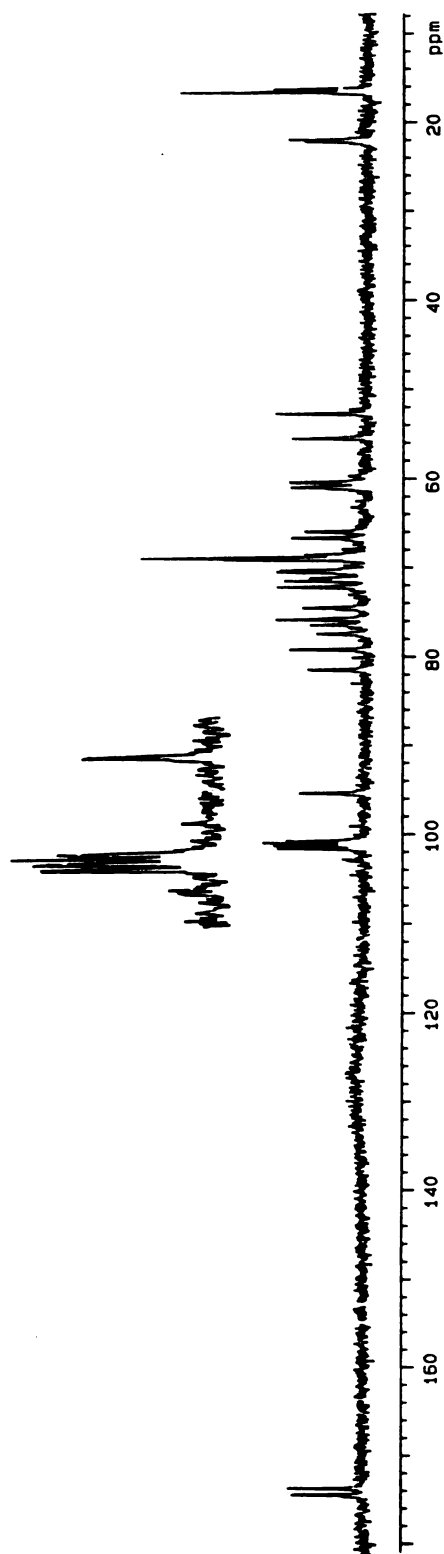
Figure 5. ^{13}C NMR spectrum of the polysacchride. The five anomeric carbon signals appear in the region of 94-102 ppm. Two N-acetyl carbon signals are at 21.6 ppm and 21.8 ppm. Three 6-deoxy carbon signals are at 16 ppm. The five anomeric carbon resonances are expanded in the insert.

Figure 6. ^{13}**C DEPT spectrum of the polysaccharide.** Note the two 6-hydroxymethy carbons appear at 60.0 ppm and 60.6 ppm. The C-2 resonances of the amino sugar residues are at 52.4 ppm and 55.2 ppm.
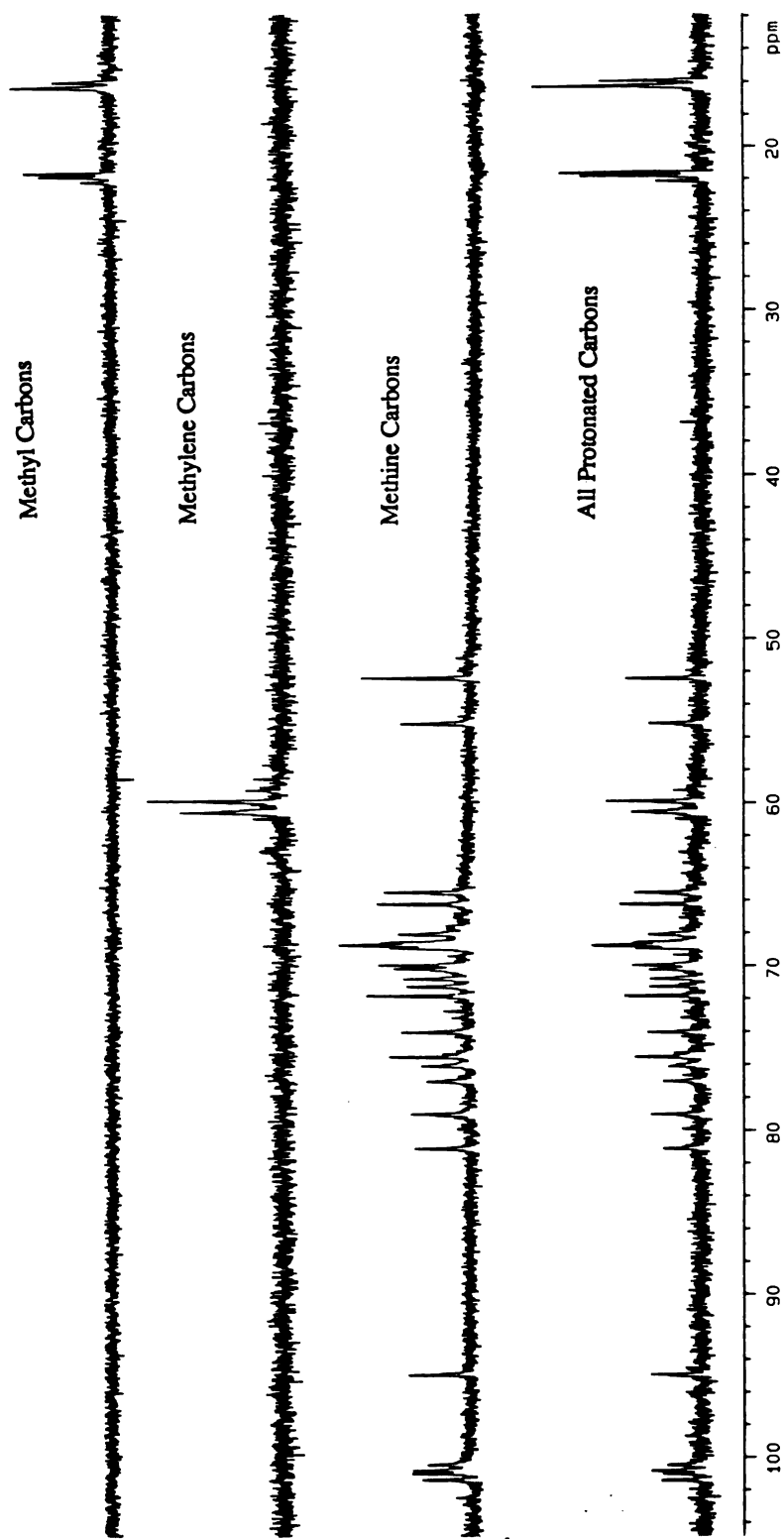
Figure 6

Figure 7. ^{1}H-^{13}C HMQC spectrum of the polysaccharide. The proton signals agree well with the carbon signals. Note the assignment for the five cross peaks corresponding to the substitution sites. R_2-2 refers to the C-2 of rhamnosyl residue 2, R_2-3 to C-3 of rhamnosyl residue 2, R_1-3 to C-3 of rhamnosyl residue 1, R_3-3 to C-3 of rhamnosyl residue 3 and G-4 to C-4 of N-acetyl glucosamine. Note that there is no such cross peak for the N-acetyl mannosamine residue.

in substitutions. These are typically 10 ppm downfield of unsubstituted carbon signals. The five such substituted ^{1}H-^{13}C correlated signal pairs are as follows: 3.58 ppm and 81.1 ppm, 3.95 ppm and 79.1 ppm, 3.76 ppm and 77.0 ppm, 4.22 ppm and 76.1 ppm, and 3.42 ppm and 75.5 ppm.

The proton signals belonging to one continuous spin system were traced from the proton HOHAHA spectrum (Figure 8). The Rha1 anomeric proton signal at 5.13 ppm was found to be connected to another signal at 4.16 ppm, past which there was no transfer because of the small H-1-H-2 coupling. This signal was thus assigned to H-2 of Rha1. Similarly, the ^{1}H signal at 4.22 ppm was assigned to H-2 of Rha2. This latter resonance was one that was found to be correlated with a resonance assigned to a carbon involved in a linkage by the HMQC experiment. The O-2 of Rha2 was thus found to be linked to another residue. The signal at 3.86 ppm was assigned to H-2 of Rha3 since no spin transfer was observed beyond this resonance. The anomeric proton signal of N-acetyl glucosamine showed connectivities with four other signals, at 3.79 ppm, 3.62 ppm, 3.47 ppm and 3.42 ppm. Further proton signals assignments were accomplished through the spin connectivities in the DQF-COSY spectrum (Figure 9). Of the resonances for the N-acetyl glucosamine residue, the signal at 3.79 ppm was correlated with the H-1 signal in the DQF-COSY spectrum. This signal was, therefore, assigned to H-2. The H-2 signal was also correlated with the signal at 3.62 ppm. The latter signal at 3.62 ppm was, therefore, assigned to H-3. The hydroxymethyl signals at 3.89 ppm and 3.74 ppm (traced from HMQC cross peaks corresponding to C-6 signal at 60.0 ppm), were correlated with the the signal at 3.47 ppm. This latter resonance was thus assigned to H-5 of the N-acetyl glucosamine residue. Finally, the signal at 3.42 ppm was assigned to H-4. This position was substituted, as indicated by its large ^{13}C chemical shift from the HMQC results. Thus all of the proton signals of N-acetyl glucosamine residue were assigned. The proton signal assignments from HOHAHA and DQF-COSY studies, combined with HMQC results, revealed that the N-acetyl glucosamine residue was substituted at the 4-position. Rha1 and

Figure 8. **HOHAHA spectrum of the polysaccharide.** Note the connectivities for the amino sugar residues and rhamnosyl residues. The labels denote the glycosyl components as described in Figure 4.

42
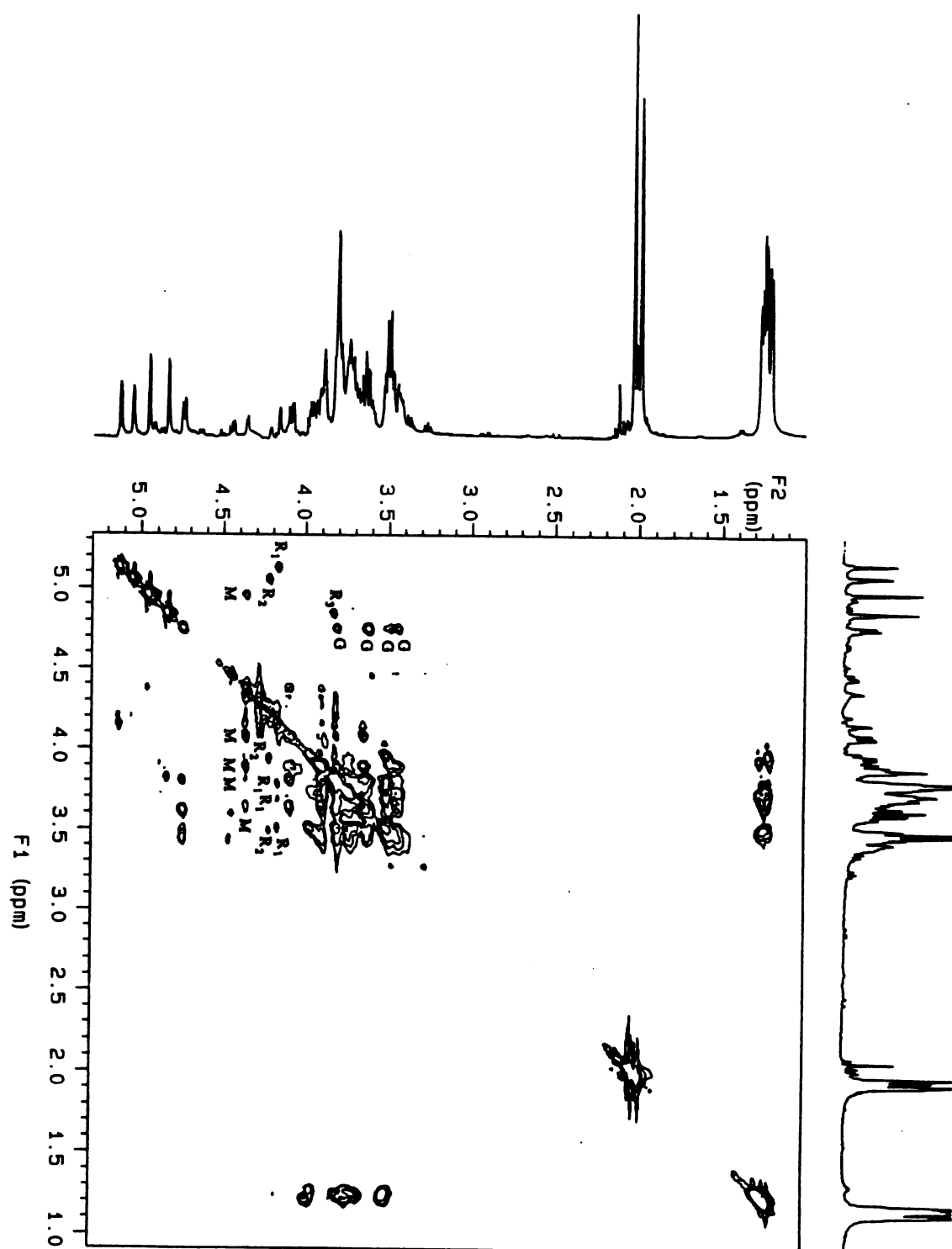
Figure 8

Figure 9. **DQF-COSY spectrum of the polysaccharide.** Note the assignments for the various connectivities. For instance, G-1, 2 denotes the cross peak correlating the signals for H-1 and H-2 of N-acetyl glucosamine.
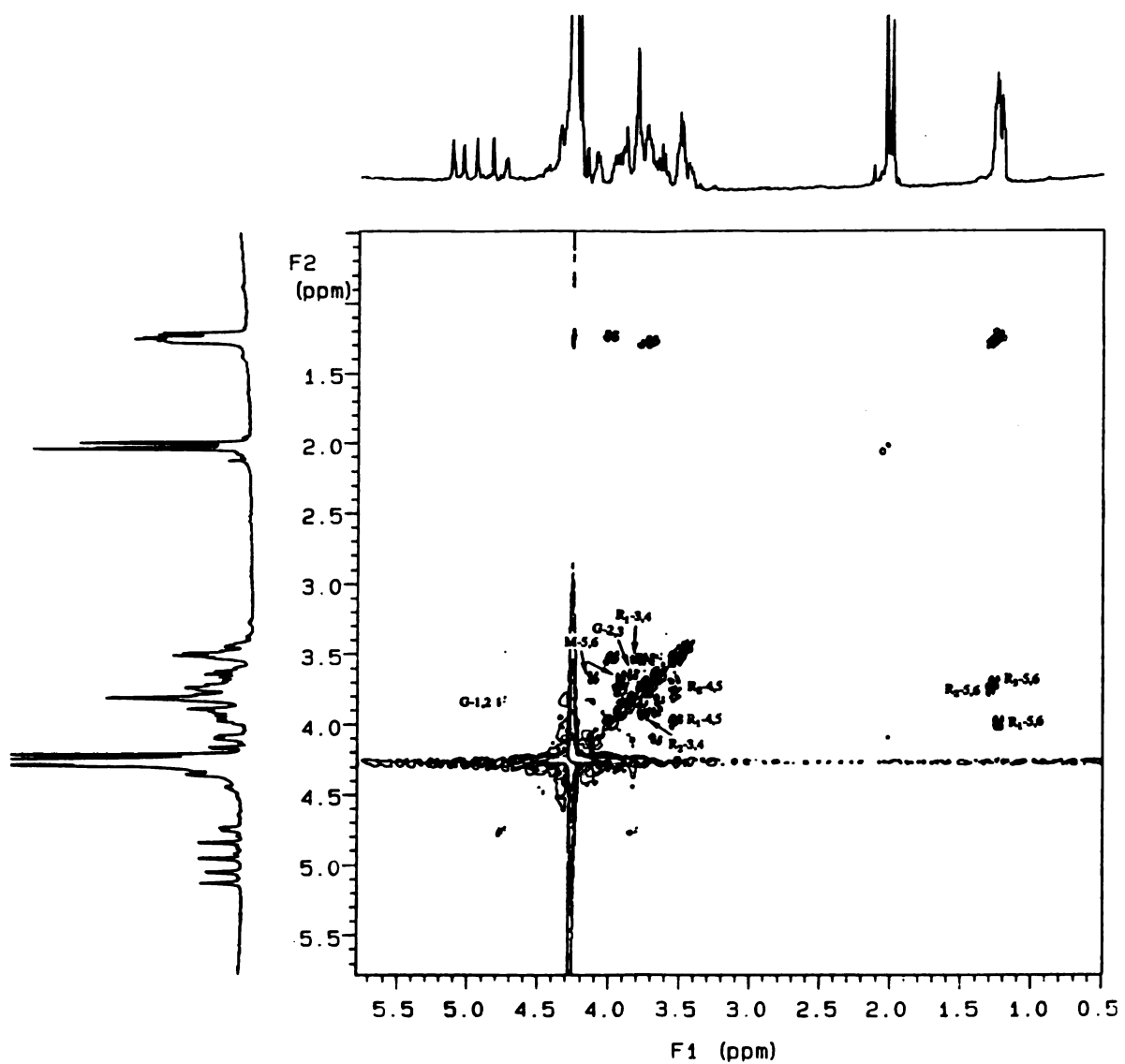
Figure 9

Rha3 were linked at the 3-positions, and Rha2 was substituted at both the 2- and 3-positions. It was concluded that the N-acetyl mannosamine was a side chain substituent since it bore no substituents itself.

The foregoing substitution results were confirmed by two successive periodate oxidations. Gas chromatography analysis (Figure 10A) showed that the first periodate oxidation caused the degradation of the N-acetyl mannosamine residue, indicating that both the O-3 and O-4 positions were not substituted. Since the O-6 was known to be unsubstituted (DEPT), the conclusion that N-acetyl mannosamine was a side chain substituent was thus confirmed. Further periodate oxidation of the first oxidation product did not cause any further degradation of the N-acetyl glucosamine and rhamnosyl residues (Figure 10B). This confirmed that N-acetyl glucosamine was linked in the 4-position, and that the three rhamnosyl residues were 3- linked. This also confirmed that the N-acetyl mannosamine was linked to the 2-position of the rhamnose2 residue. This conclusion could be reached because it was known from the HOHAHA and HMQC experiments that this position was substituted. The main-chain substituent must, therefore, be in the 3-position, rendering the branching residue inert to further periodate oxidation after the side chain N-acetyl mannosamine residue was removed.

The sequence of the polysaccharide was deduced by partial hydrolysis and 1D and 2D n.O.e. experiments. The polysacharide was partialy hydrolyzed under condition strong enough to cleave the anomeric linkages of the rhamnosyl residues but not the anomeric linkages of acetamino sugars. The products were analyzed by FAB-MS (Figure 11). Disaccharides of GlcNAc-(1-3)-Rha or ManNAc-(1-2)-Rha were identified in the mass spectrum, but no trisaccharide species were detected. These results excluded the possible sequence of β-D-GlcNAc-(1-3)-α-L-Rha2-(2-1)-α-D-ManNAc (two acetamino residues linked to the same rhamnose residue). Two possible sequnces of the repeating unit thus remained. These were: -3)-α-L-Rha-(1-3)-(α-D-(1-2)-)-α-L-Rha2-(1-4)-β-D-GlcNAC-

Figure 10. GC profiles of the alditol acetate derivatives of the first (A) and the second (B) periodate oxidation products. Note that the mannosamine was oxidized in the first reaction. Glucosamine and rhamnose were not affected by the periodated treatment.

Figure 10

48
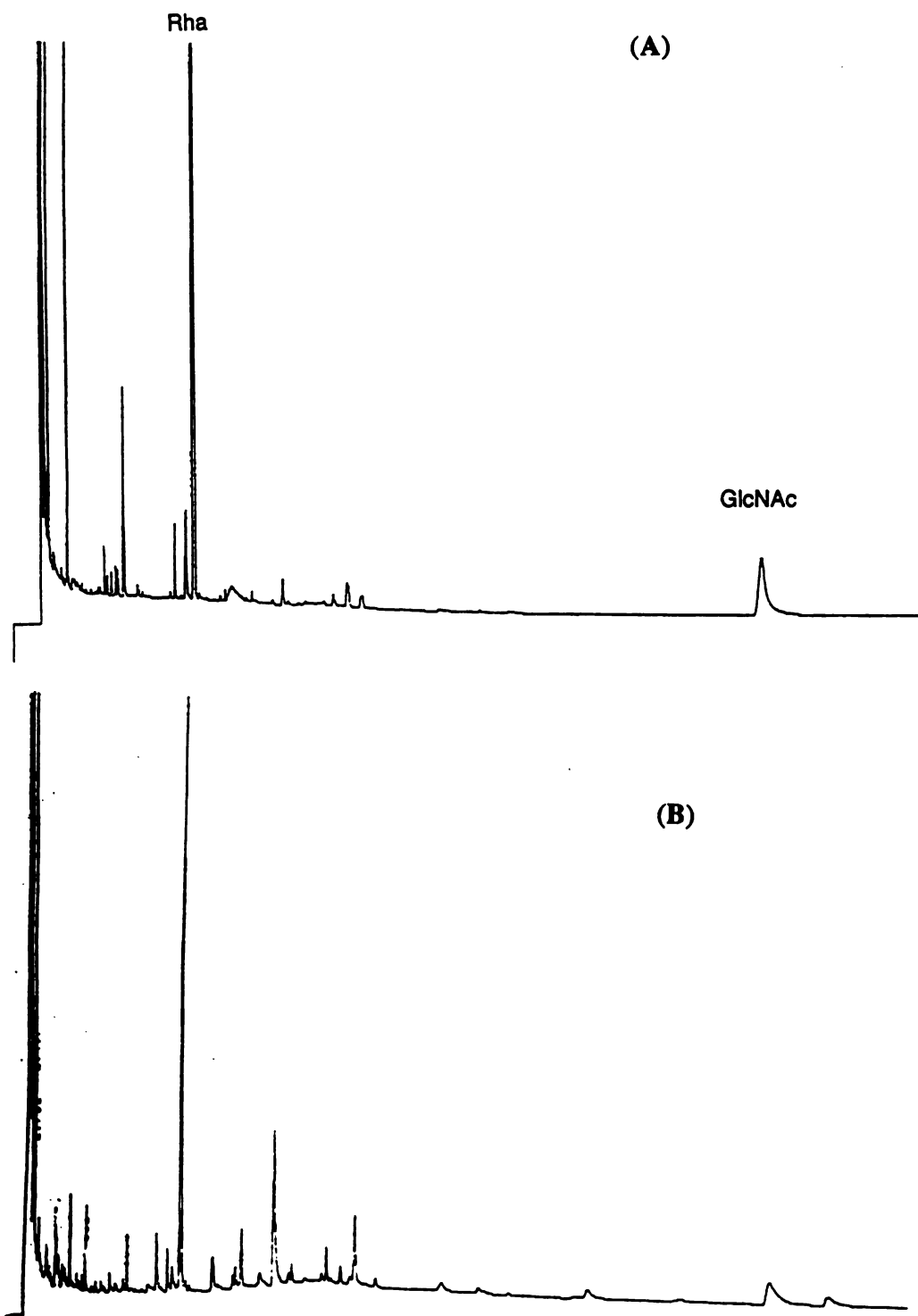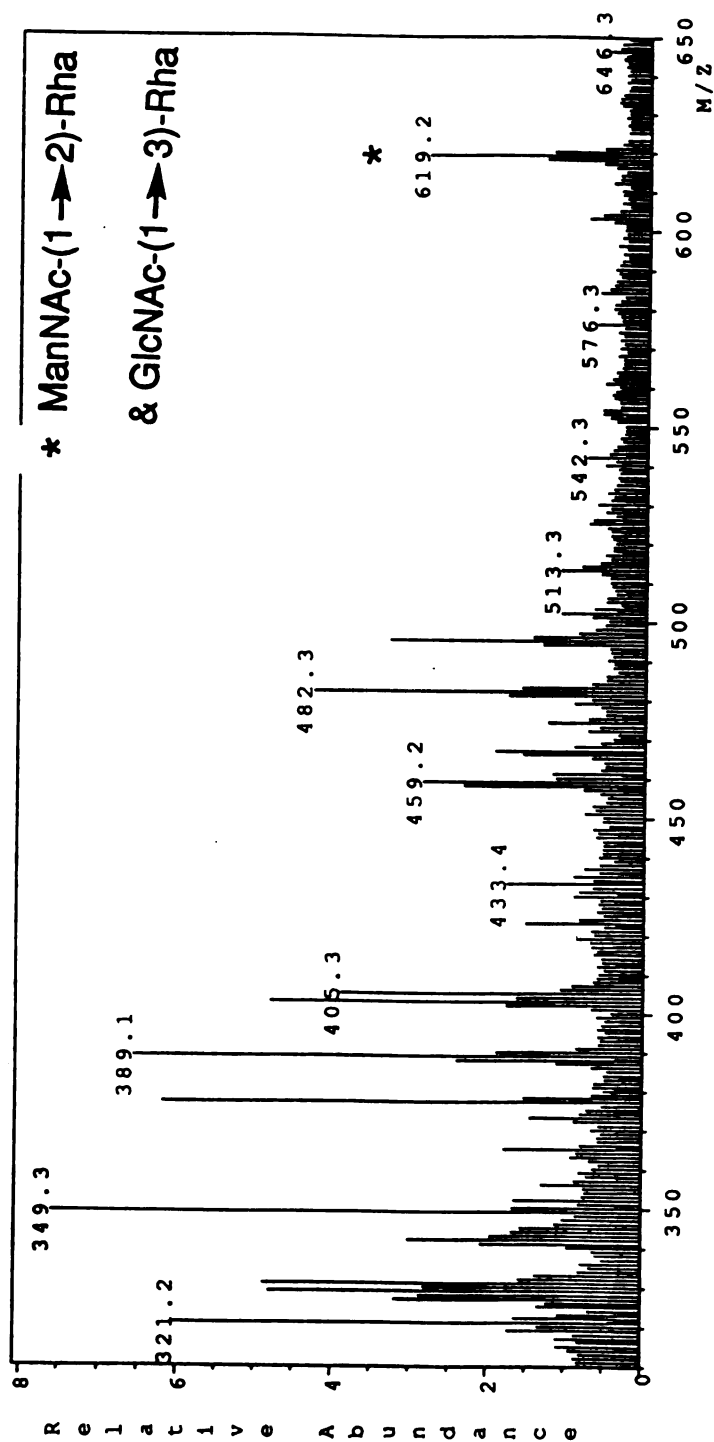
Figure 11. FAB mass spectrum of the partially hydrolyzed polysaccharide. Note that m/z at 389.1 was from the per-acetylated GlcNAc or ManNAc and 321.2 was from the per-acetylated Rha. Dissaccharide fragments are identified, but no trisaccharide fragment was detected.

(1-3)-α-L-Rha-(1-, and -4)-β-D-GlcNAc-(1-3)-α-L-Rha-(1-3)-(α-D-ManNAc-(1-2)-)-α-L-Rha2-(1-3)-α-L-Rha-(1-. In the NOESY spectrum (Figure 12), H-1 of the rhamnose 1 showed correlation with H-2 of the rhamnose 2 and the H-2 of itself. H-1 of the N-acetyl mannosamine residue was observed to be spatially close to H-2 of the rhamnose 1. This suggested the sequence: -3)-α-L-Rha1-(1-3)-(α-D-ManNAc-(1-2)-)-α-L-Rha2-(1-. The anomeric proton of the N-acetyl glucosamine showed an n.O.e. effect with H-2 of Rha3, indicating the partial sequence: -4)-β-D-GlcNAc-(1-3)-α-L-Rha3-(1-. In the 1D n.O.e. (Figure 13), the anomeric proton of N-acetyl glucosamine showed an n.O.e. with the N-acetyl protons of both N-acetyl glucosamine and N-acetyl mannosamine, thus indicating that these two amino sugar residues were close to each other. The structure of the repeating unit was therefore readily shown to be that in Figure 14. The complete ^{1}H NMR assignments for the carbohydrate ring protons are given in Figure 15. The ^{1}H and ^{13}C chemical shifts are tabulated in Table 1.

This study describes the complete structural characterization of the O-antigen component of the lipopolysaccharide of *Rhizobium trifolii* 4S. The structure of the capsular polysaccharide of this organism has already been determined (10). The number of residues and their configurations were readily determined by GC, GC-MS and NMR spectroscopy. The combination of DEPT, HOHAHA, and DQF-COSY spectroscopy allowed the determination of the sites of linkages of the residues. The linkage positions were fully supported by periodate oxidation, which clearly demonstrated that the rhamnosyl residues were 3-linked and, therefore, had no vicinal diol functions which could be oxidized. The same experiment also confirmed that the N-acetyl mannosamine residue was terminally linked as its primary hydroxyl group was unacetylated (DEPT) and the 3- and 4-positions were both free, since it was completely oxidized by periodate oxidation. Failure to remove any other residue in a second oxidation prove that the N-acetyl mannosaamine component was, in fact, the sole side chain residue. Only one glycosyl
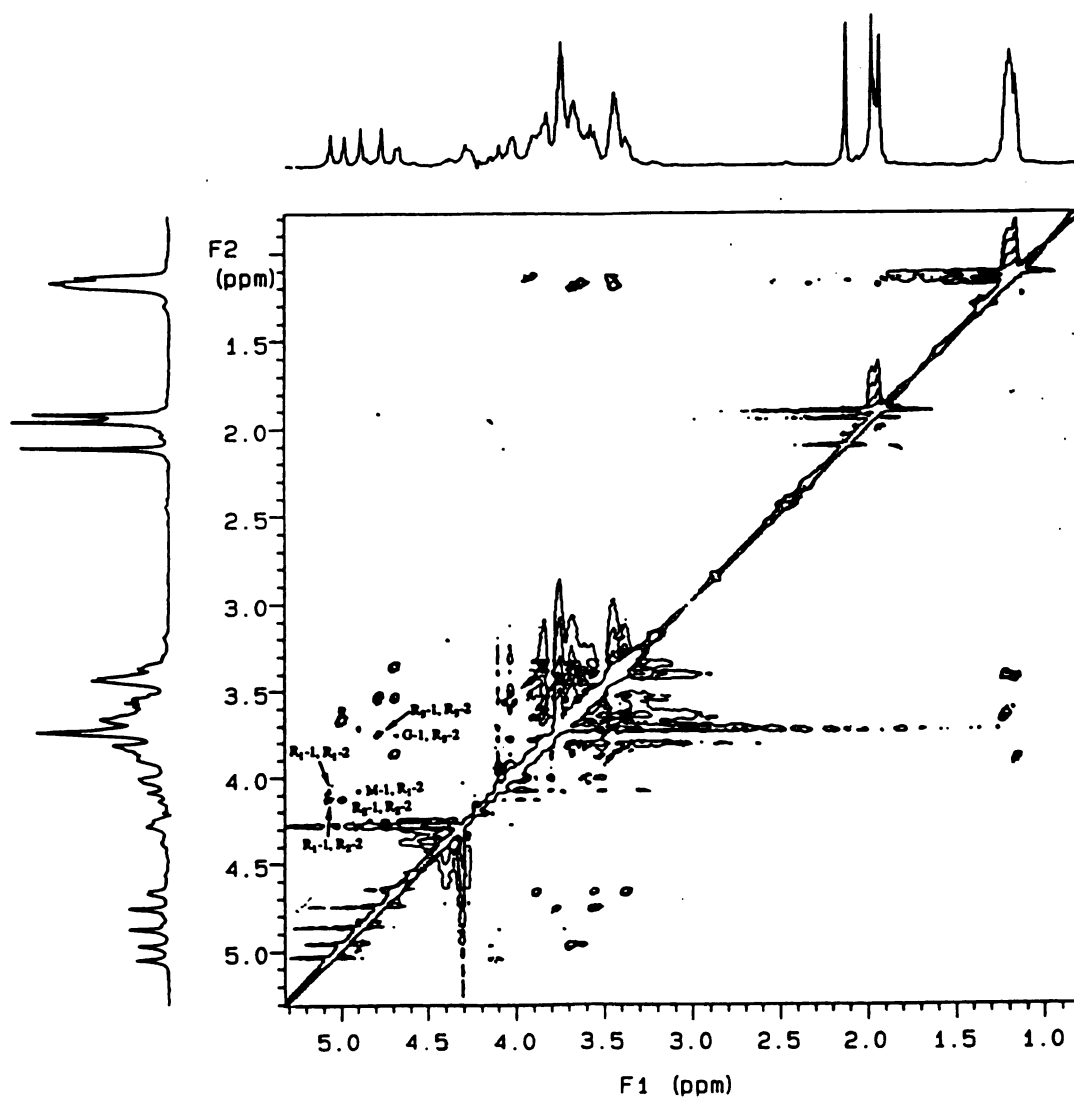
Figure 12. NOESY spectrum of the polysaccharide. Note the correlation between each anomeric proton signal and other sugar proton signals.

51
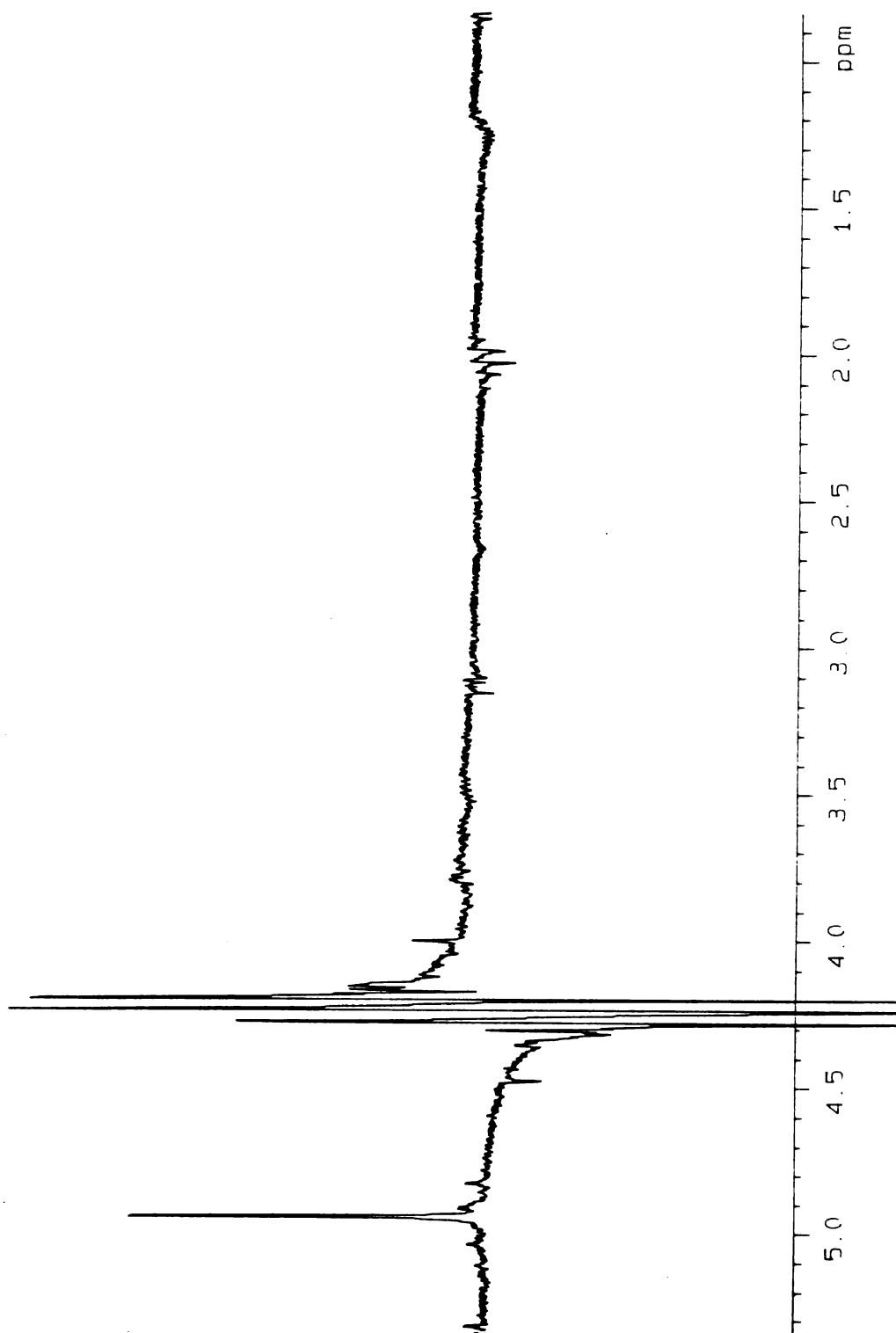
Figure 13. NOE spectrum of the polysaccharide. The anomeric signal of N-acetylmannosamine residue at 4.95 ppm was irradiated and showed n.O.e. with N-acetyl protons of both GlcNAc and ManNAc residues.
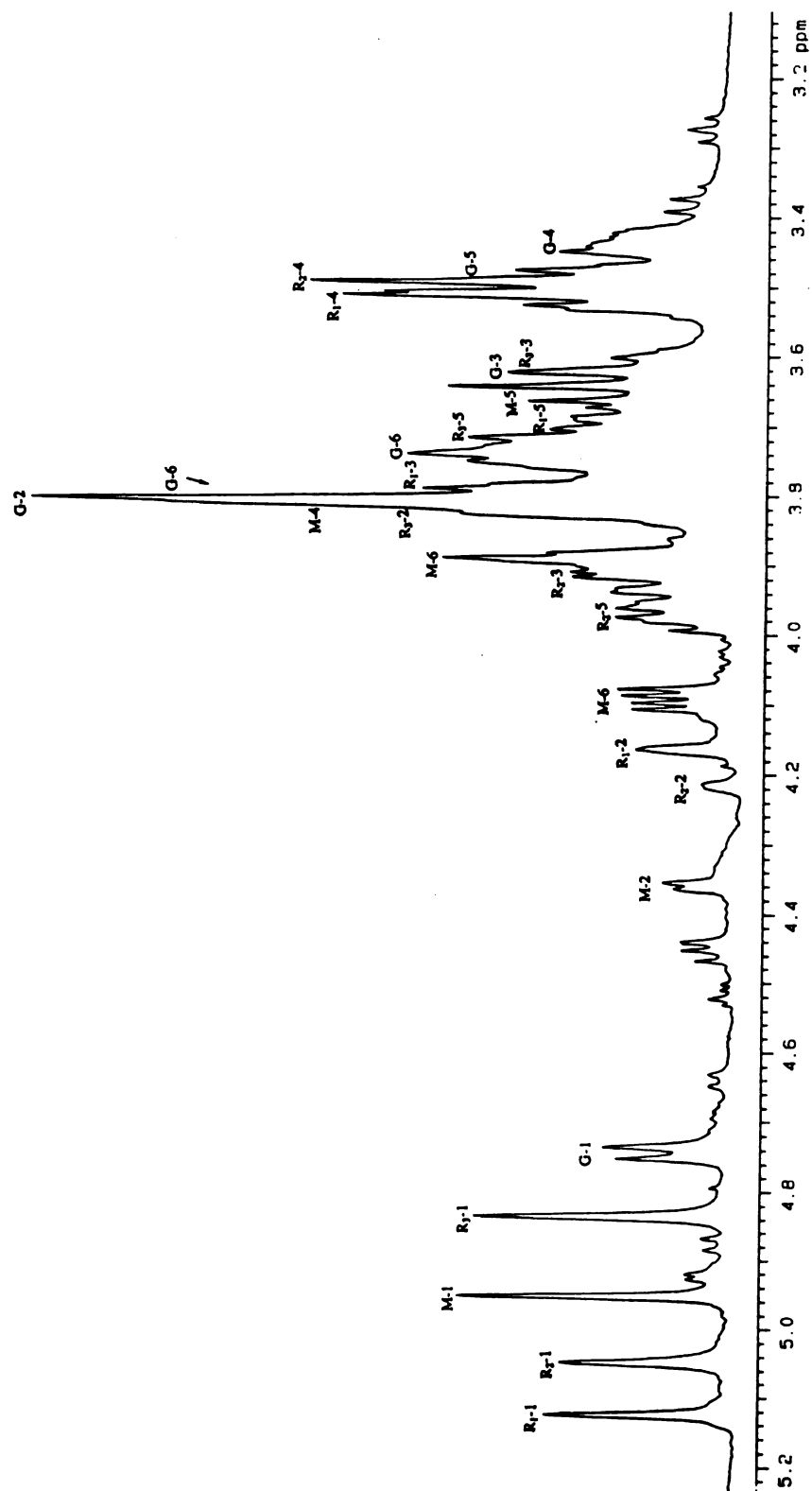
Figure 14. Part of the ¹H NMR spectrum with detailed assignments.
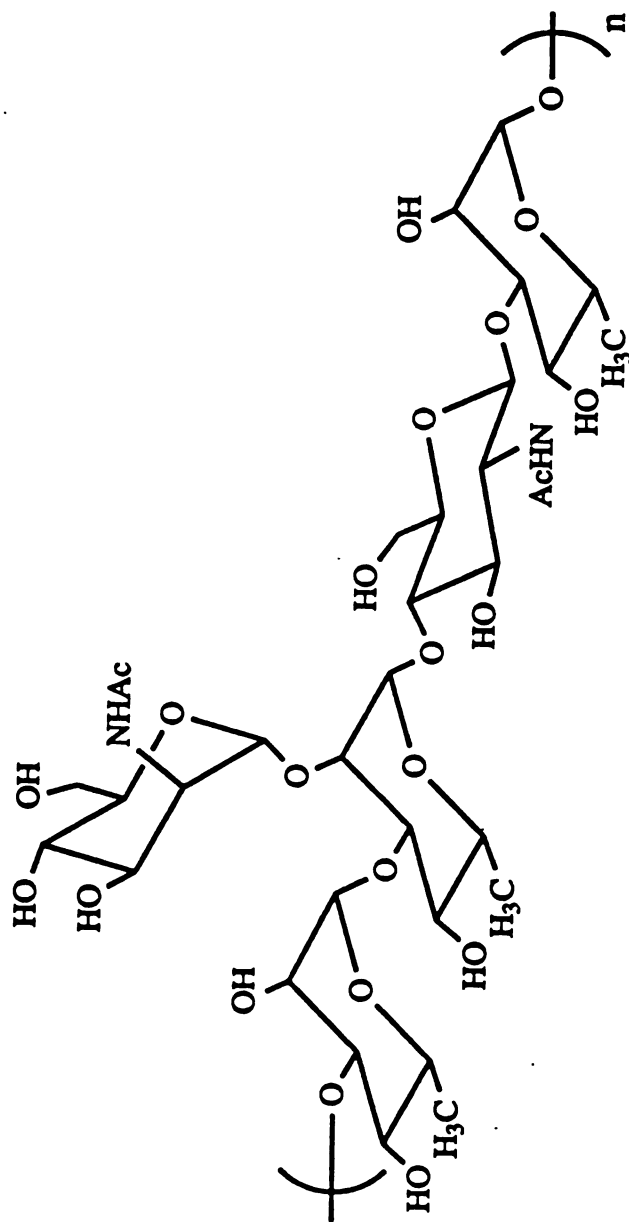
Figure 15. Structure of the polysaccharide.

Table 1. The ¹H and ¹³C chemical shifts of the polisaccharide. Note that the remaining un-assigned ¹³C signals are between 68.0-74.0 ppm.

¹H (ppm)	H-1	H-2	H-3	H-4	H-5	H-6
Rha1	5.13	4.16	3.76	3.52	3.69	1.25
Rha2	5.05	4.22	3.95	3.49	3.97	1.25
Rha3	4.84	3.86	3.58	3.70		1.25
ManNAc	4.95	4.36		3.83	3.67	4.09, 3.90
GlcNAc	4.75	3.79	3.62	3.42	3.47	3.89, 3.74

¹³C (ppm)	C-1	C-2	C-3	C-4	C-5	C-6
Rha1	101.1	65.9	77.0		68.7	16.0
Rha2	100.5	76.1	79.1		68.9	16.0
Rha3	100.9		81.1		68.8	16.0
ManNAc	95.0	52.4			68.1	60.0
GlcNAc	101.5	55.2		75.5		60.6

component (a rhamnosyl residue) bore two substituent glycosyl residues (not counting the anomeric sites). The final proposed sequence was readily confirmed by partial hydrolysis and n.O.e. measurements. The anomeric configurations were unequivocally established by the NMR chemical shifts and coupling constants of the anomeric protons.

REFERENCES

1. Maier, R. J.; Brill, W. J. *J. Bacteriol.* **1978**, 133, 1295-1299.

2. Maier, R. J.; Brill, W. J. *J. Bacteriol.* **1976**, 127, 763-769.

3. Stacey, G.; Paau, A. S.; Noel, K. D.; Maier, R. J.; Silver, L. E.; Brill, W. J. *Arch. Microbiol.* **1982**, 132, 219-224.

4. Carlson, R. W.; Garcia, F.; Noel, D.; Hollingsworth, R. I. *Carbohydr. Res.* **1989**, 195, 101-110.

5. Zhang, Y.; Hollingsworth, R. I.; Priefer, U. B. *Carbohydr. Res.* **1992**, 231, 261-271.

6. Priefer, U. B. *J. Bacteriol.* **1989**, 171, 6161-6168.

7. Carlson, R. W.; Hollingsworth, R. I.; Dazzo, F. B. *Carbohydr. Res.* **1988**, 176, 127-135.

8. Hollingsworth, R. I.; Carlson, R. W.; Garcia, F.; Gage, D. *J. Biol. Chem.* **1989**, 264, 9294-9299.

9. Hollingsworth, R. I.; Carlson, R. W.; Garcia, F.; Gage, D. *J. Biol. Chem.* **1989**, 265, 12752.

10. Amemura, A.; Harada, T.; Abe, M.; Higashi, S. *Carbohydr. Res.* **1983**, 115, 165-174.

11. Higashi S.; M. Abe, M. *J. Gen. Appl. Microbiol.* **1978**, 24, 143-153.

12. Hollingsworth R. I.; Lill-Elghanian, D. A. *J. Biol. Chem.* **1989**, 264, 14039-14042.

13. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. *Carbohydr. Res.* **1978**, 62, 349-357.

14. Dubois, M.; Gills, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F. *Analy. Chem.* **1956**, 28, 350-356.

15. Bax, A.; Subramanian, S. *J. Magn. Reson.* **1986**, 67, 565-569.

16. Ernst, R. R.; Bodenhansen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Oxford University Press: Oxford, 1987; pp. 431-440.

17. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, 65, 355-360.

18. Sangers, J. K.; Hunter, B. K *Modern NMR Spectroscopy*; Oxford University Press: Oxford, 1993; pp 193-197.

CHAPTER 3

A SOLVENT SYSTEM FOR THE HIGH RESOLUTION PROTON NMR

SPECTROSCOPY OF MEMBRANE LIPIDS

ABSTRACT

NMR spectroscopy has tremendous potential as an analytical tool for studying the compositions and structural details of complex mixtures of lipids. Unfortunately, because of aggregate formation, lipids generally give poorly resolved NMR spectra and there is no general solvent system which gives good spectra of all classes of lipids. A simple solvent system for the high resolution NMR spectroscopy of membrane lipids has been developed. This solvent system is composed of a mixture of pyridine-d_5, deuterium chloride in deuterium oxide, methanol-d_4 and chloroform-d in a volume ratio of 1:1:2:10, respectively. The use of this solvent system gives uniformly well-resolved ^{1}H-NMR spectra over a wide selection of lipid classes including the lipid A region of bacterial lipopolysaccharides, phospholipids and large gangliosides. The use of this solvent system makes the undesirable practice of methylation of the phosphate groups in lipid A and other molecules, in order to acquire high resolution spectra unnecessary. It also increases the general utility and potential for NMR studies on bacterial adaptation, serum lipid composition and membrane lipid alterations in cancer cells.

INTRODUCTION

The chemistry of membrane lipids is a critical aspect of several areas of biology and biochemistry. Areas in which a detailed knowledge of membrane structure and composition are important include cell structure reorganization during bacterial adaptation, the organization and composition of specialized membrane structures as are found in plants and photosynthetic bacteria and studies of mammalian cell membranes during oncogensis e.g. in tumor tissue (1). Nuclear magnetic resonance spectroscopy has the capability of making detailed structural and compositional studies of membrane lipid mixtures possible (2). This is especially true since the advent of the newer NMR methodologies which make the analysis of complex mixtures potentially routine. Pulse sequence such as that used in the HOHAHA experiment (3) makes the tracing and assignment of spin connectivities (and even the extraction of a "pure" one-dimensional spectrum) of a single component of a mixture possible. The HMQC experiment (4) allows the correlation between phosphorous and protons possible (as well as those between carbon or nitrogen and protons) and the newer "triple" resonance probe designs make the simultaneous determination of the different nuclei in one molecular species possible (5). In fact NMR spectroscopy has recently been suggested as a tool for following the compositions of the membranes of cancer cells (6, 7).

Despite its potential for use in membrane structural and compositional analysis, one major problem which limits the use of NMR is the fact that there is no universal NMR solvent or solvent system which gives uniformly good, well resolved peak for all of the lipid classes, many of which occur simultaneously on the same cell surface. Hence, whilst one can obtain moderately good spectra of some gangliosides in dimethylsulfoxide provided all traces of metals are first removed (8), these conditions do not give good spectra for phospholipids, in fact, some phospholipids are insoluble in dimethylsulfoxide.

More non-polar lipids may tend to be soluble in solvents containing high proportions of chloroform but the more polar gangliosides are completely insoluble in such solvents and, if they are, do not yield interpretable spectra because of extreme line broadening. Molecules such as the lipid A component of lipopolysaccharides usually yield very poor spectra in any solvent. In order to obtain high resolution spectra of such molecules, the charged phosphate groups are often esterified by methylation with diazomethane (9). This practice is undesirable since diazomethane is capable of methylating alcohols and amines and methylates carboxylic acid functions instantaneously. Diazomethane is also extremely prone to explosion. There is, in fact, no universal solvent in which uniformly high quality NMR spectra can be obtained over the large cross section of lipid classes.

We are interested in developing NMR methods for studying the component structures and compositions of complex lipid mixtures. Such systems include entire membrane preparations from both bacterial and mammalian sources. These studies require the use of NMR solvents which give universally good spectra for all classes of lipid molecules. The quality of the spectra should allow the detection of even long range splitting. In this study, we describe the use of a simple solvent system which consistently gives uniformly well-resolved ^{1}H-NMR spectra over a wide selection of lipid classes including the lipid A region of bacterial lipopolysaccharides, phospholipids and large (usually intractable) gangliosides. This combination of solvents was designed to reduce the magnitude of the electrostatic forces between lipid head groups while being non-polar enough to disperse their interacting hydrocarbon chains. We think that the use of this and similar solvents will greatly facilitate the use of NMR spectroscopy in the study of complex lipid mixtures. This should facilitate NMR studies on serum lipids, bacterial membrane changes during adaptation and membrane lipid alterations in cancer cells, three major areas in which NMR spectroscopy has tremendous promise.

MATERIALS AND METHODS

Materials

Monophosphoryl lipid A from *Salmonella minnesota* R595 and diphosphoryl lipid A from *Escherichia coli* K12 (D31m4) were purchased from List Biological Laboratories, Inc. Disialoganglioside GD1a, L-α-phosphatidylcholine, L-α-phosphatidyl-L-serine (PS) and L-α-phosphatidylethanolamine (PE) were obtained from Sigma chemical company. Chloroform-d (D, 99.8%) and methanol-d4 (D, 99.8%) were purchased from Cambridge isotope laboratories. Pyridine-d5 (99%) and deuterium chloride (37 wt%) solution in deuterium oxide were purchased from Aldrich chemical company.

Methods

The NMR solvent system was prepared as a mixture of pyridine-d5, deuterium chloride in deuterium oxide, methanol-d4 and chloroform-d. These solvents were mixed in a volume ratio of 1:1:2:10, respectively. The concentration of the NMR samples were 3 mg/ml for lipid A samples, 6 mg/ml for phospholipids and 0.8 mg/ml for the disialoganglioside. All NMR spectra were recorded on a Varian VXR500 spectrometer operating at 500 MHz for protons. Experiments were conducted at 25 °C or 35 °C. The NMR experiments were conducted using chloroform as the lock signal. The ^{1}H-NMR chemical shifts were referenced relative to the chloroform signal at 7.24 ppm. The ^{31}P-NMR spectra were acquired at 202.33 MHz with a spectral width of 1583 Hz. H_3PO_4 (85%) was used as an external reference. The ^{1}H/^{31}P-HMQC NMR experiments were acquired with 256 data set at 2048 data points each and 16 transients per data set.

RESULTS

Phospholipids. Phosphatidylethanolamine (PE) and phosphatidylserine (PS) were examined using the present solvent system and both of them yielded well-resolved ^{1}H-NMR spectra. In the ^{1}H-NMR spectrum of PE (FIG. 1), the methylene protons of the 3-position of the glycerol backbone were nonequivalent and gave one doublet of doublets at 4.03 ppm and another at 4.25 ppm. This was due to coupling with each other and with the methine proton at the 2-position. The signal at 4.25 ppm was more strongly coupled to the methine proton signal than the one at 4.03 ppm. This methine proton signal appeared at 5.12 ppm as a complex multiplet, due to coupling with both sets of methylene protons in the glycerol backbone. The methylene protons of the -C$\underline{H}_2$-OP substructure gave a doublet of doublets centered at 3.94 ppm, resulting from the coupling to the methine proton of the glycerol backbone and the ^{31}P. Of the ethanolamine head group, the -C$\underline{H}_2$-N protons appeared as a triplet at 3.15 ppm. The splitting was due to coupling with the other methylene protons of the ethanolamine group. These latter methylene protons appeared at 4.08 ppm as a multiplet, due to coupling with the above -C$\underline{H}_2$-N protons as well as with the ^{31}P of the phosphate head group. For the fatty acyl protons, the triplet at 0.74 ppm was assigned to the terminal methyl protons. The signals at 1.14 ppm were assigned to the methylene protons of the alkyl chains. There were two sets of triplets, one at 2.16 ppm and the other at 2.18 ppm. These were assigned to the α-methylene protons of fatty acyl chains. The unresolved signal at 1.45 ppm was assigned to the β-methylene protons of the fatty alkyl chains.

The ^{1}H NMR spectrum of PS (FIG. 2) was very different to that of PE. Of the protons of the glycerol backbone, the doublet of doublets at 4.02 ppm and 4.23 ppm were each assigned to one of the methylene protons of the 3-position of the glycerol substructure. The multiplet at 5.12 ppm was assigned to the methine proton of the glycerol

Figure 1. **The ^{1}H NMR spectrum of PE showing assignments for the head group protons.** The protons are numbered. The unlabeled multiplet at 3.22 ppm is from deuterated methanol and the broard singlet at ~4.65 ppm is from hydroxyl protons in the solvent.

Figure 1

Figure 2. **The ¹H NMR spectrum of PS.** The signals to the left of the one marked "2" are due to vinyl protons in the fatty acid chain. The ones between 2.6 and 2.8 ppm are due to allylic protons.
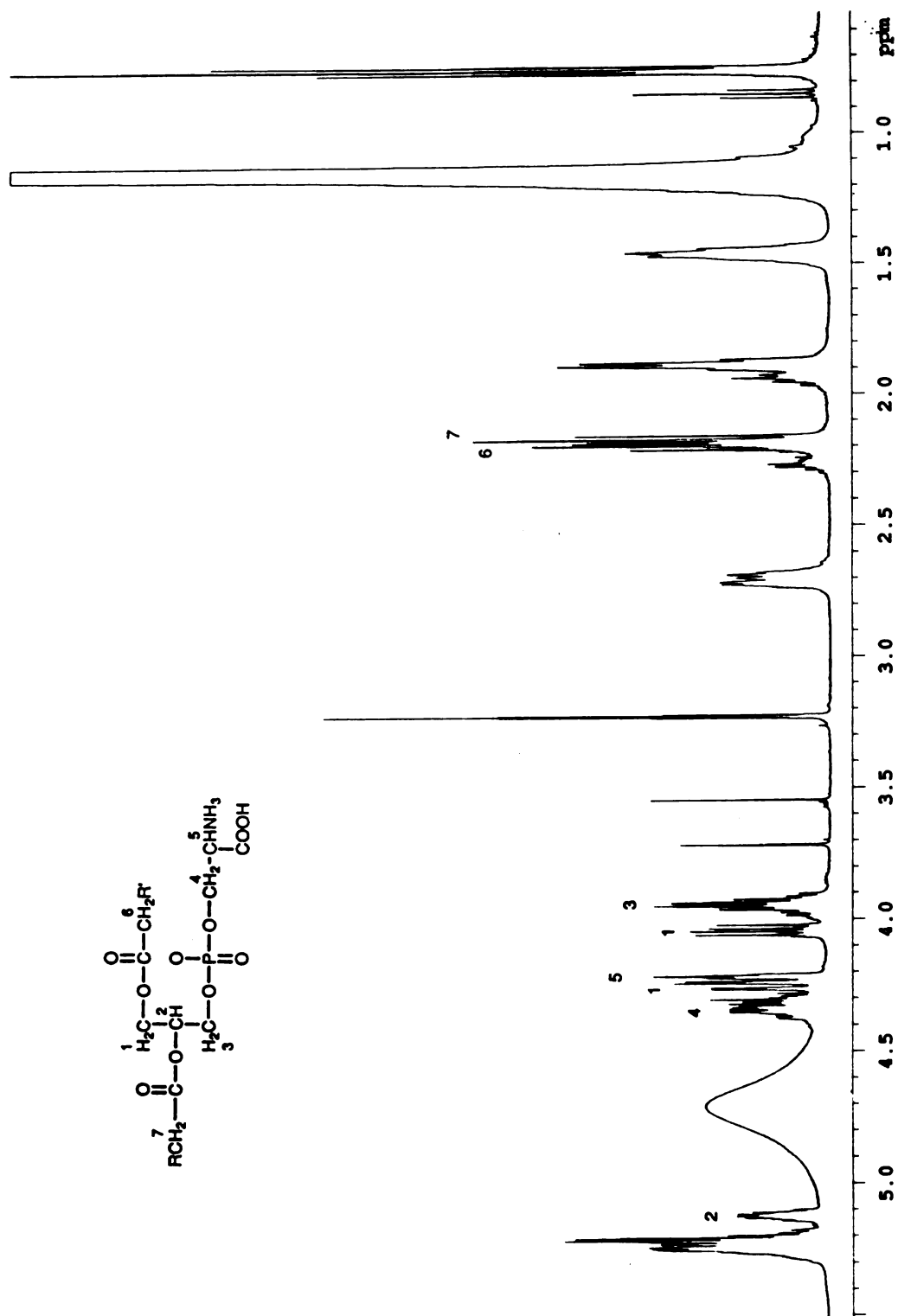
Figure 2

position of the glycerol moiety. The multiplet at 4.32 ppm was assigned to the methylene protons of the serine head group. The methine proton of the serine head group gave rise to a multiplet at 4.20 ppm. Weak coupling with the phosphorous of the phosphate group was observed. Signals for C-2 protons of the fatty acyl chains appeared as triplets at 2.16 ppm and 2.18 ppm. The β-methylene proton signals appeared at 1.44 ppm. Unsaturation in the fatty acyl chain was evident from signals at 5.24 and 5.19 ppm which was assigned to the protons of the -CH=CH- group. The signals at 1.87 ppm were assigned to the allylic methylene protons. Since this PS sample was extracted from Bovine brain, heterogeneity of the fatty acyl chains might account for the presence of the low intensity resonances at 2.25 ppm, 1.92 ppm and 0.83 ppm. The proton resonance assignments for PE and PS are summarized in TABLE 1.

Mixture of phospholipids. A mixture of phosphatidylethanolamine, phosphatidylserine and phosphatidylcholine was tested in the solvent system. The ^{1}H NMR spectrum (FIG. 3) of the mixture was well-resolved. The presence of PE and PS were verified from their ^{1}H resonance by comparing to their standard spectrum. The presence of PC was recognized from its distinct $-N^+(CH_3)_3$ proton signal, which appeared as a huge sharp singlet at 3.2 ppm. The presence of PC was also confirmed by its other ^{1}H signals as labelled in the spectrum. The relative lipid contents of the mixture can be obtained by their ^{1}H signal peak ratio.

Lipid A. Typical lipid A molecules such as those from *S. minnesota* and from *E. coli* are both composed of a β-D-glucosamine-(1'-6)-α-D-glucosamine disaccharide head group. Monophosphoryl lipid A has one phosphoryl group at the 4'-position of the distal glucosaminyl residue. Diphosphoryl lipid A has two phosphoryl groups at positions 4' and 1 of the glucosaminyl residues. The disaccharide head group is acylated by four (R)-3-hydroxyltetradecanoic acid residues at positions 2, 3, 2', and 3'. Some of these fatty acyl residues are further acylated at their 3-hydroxyl positions by other fatty acids. The

Table 1. The ^{1}H-chemical shift assignments of phospholipids.

		δ (ppm), J (Hz)
PE:		
glycerol backbone	-^{1}CH$_2$-OCO-	4.03 (doublet of doublet, J 12.0, 3.5)
		4.25 (doublet of doublet, J 12.0, 3.5)
	-^{2}CH-	5.12 (multiplet)
	-^{3}CH$_2$-OP-	3.94 (doublet of doublet, J=6.5, 5.5)
ethanolamine group	-PO$_4$CH$_2$	4.08 (multiplet)
	-CH$_2$NH$_3$	3.15 (triplet, J=5.0)
fatty acyl chain	-COCH$_2$-R	2.18 (triplet, J=8.0)
	-COCH$_2$-R'	2.16 (triplet, J=8.0)
	-COCH$_2$CH$_2$-	1.45 (multiplet)
	-CH$_2$-	1.14
	-CH$_3$	0.74 (triplet, J=7.0)
PS:		
glycerol backbone	-^{1}CH$_2$-OCO-	4.02 (doublet of doublet, J 12.0, 7.0)
		4.23 (doublet of doublet, J 12.0, 7.0)
	-^{2}CH-	5.12 (multiplet)
	-^{3}CH$_2$-OP-	3.93 (multiplet)
serine head group	-PO$_4$CH$_2$	4.32 (multiplet)
	-CH(COOH)NH$_3$	4.20 (multiplet)
fatty acyl chains	-CH$_3$	0.74 (triplet, J=6.8)
	-CH$_2$-	1.13
	-COCH$_2$CH$_2$-	1.44
	-COCH$_2$-R	2.18 (triplet, J=8.0)
	-COCH$_2$-R'	2.16 (triplet, J=8.0)
	-CH=CH-CH$_2$-	1.87
	=CH-CH$_2$-CH=	2.68
	-CH=CH-	5.19
		5.24
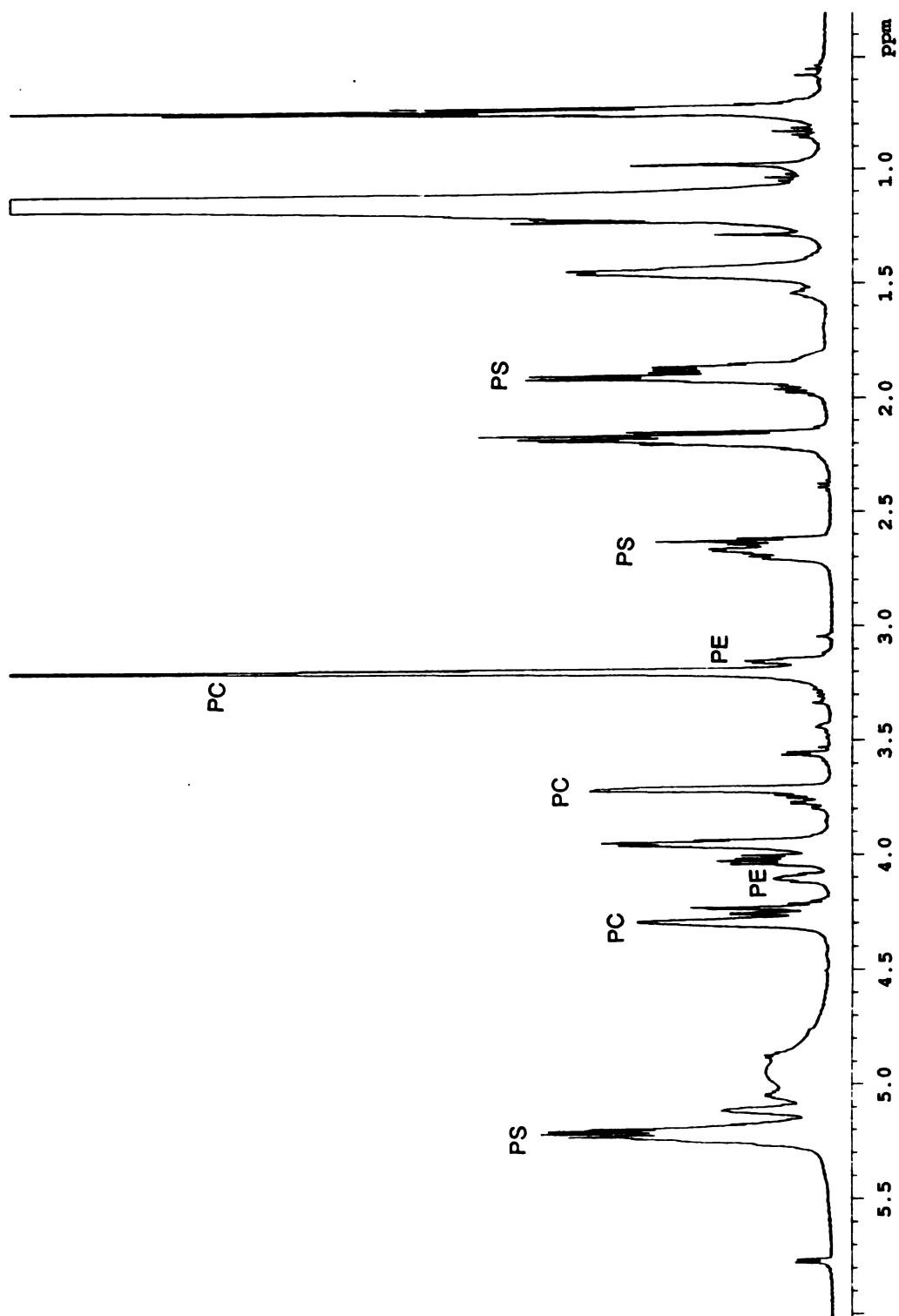
Figure 3. The ^{1}H NMR spectrum of a mixture of PE, PS and PC.

structural features of these lipid A molecules were revealed in their ^{1}H NMR spectra. n both the monophosphoryl lipid A (FIG. 4) and diphosphoryl lipid A (FIG. 5) spectra, the dominant signals at 1.15 ppm and 0.78 ppm were from the methylene and methyl groups of fatty acyl chains. The signals between 2.1-2.7 ppm were assigned to the α-methylene protons (adjacent to the carbonyl group) of the fatty acyl chains. The 3-position protons of the 3-hydroxyl fatty acyl chains gave signals in the region of 4.7-5.2 ppm. The other resonances between 3.0-5.5 ppm were due to the protons on the carbohydrate head group. Detailed assignments for all protons are possible using two-dimensional NMR experiments such as HOHAHA (which shows cross peaks between all the spins in a coupled system), and DQF-COSY (which shows cross peaks between spins that are directly coupled to each other) (10). Spectra of the mono- and diphosphorylated lipid A molecules in pure chloroform are shown in Fig. 6 for comparison.

The ^{31}P NMR spectra provided easy tools of distinguishing different types of lipid A head groups. As shown in the FIG. 7, monophosphoryl lipid A and diphosphoryl lipid A were easily distinguishable by their ^{31}P-NMR spectra. While the diphosphoryl lipid A showed two groups of resonance at 0.45 ppm and 2.5 ppm, the monophosphoryl lipid A showed one group of resonance between 1.0-1.4 ppm. Instead of showing a single resonance for each phosphorous site, a group of resonances appeared. This indicated that the lipid A samples were heterogeneous. The heterogeneity might represent different conformation or different aggregation forms of lipid A. The sharp signal at -0.50 ppm was assigned to the free phosphorylethanolamine which was present as an impurity.

The correlation between proton and phosphorous resonances in the lipid A molecules were obtained from the ^{1}H-^{31}P HMQC NMR spectra. For monophosphoryl lipid A, as shown by the cross peaks in the HMQC spectrum (FIG. 8), the phosphorus resonance at 1.0 and 1.4 ppm were correlated with the proton signal at 4.1 ppm. Since the phosphoryl group is at 4'-position, this proton signal was, therefore, from H-4' of the

Figure 4. The ^{1}H NMR spectrum of the monophosphoryl lipid A from *S. minnesota* in in 1:1:2:10 pyridine-D_5 / DCl / CD_3OD / $CDCl_3$. The multiplet at 3.2 ppm is from deuterated methanol and the HOD line at 4.6 ppm is suppressed by presaturation.

Figure 4

73
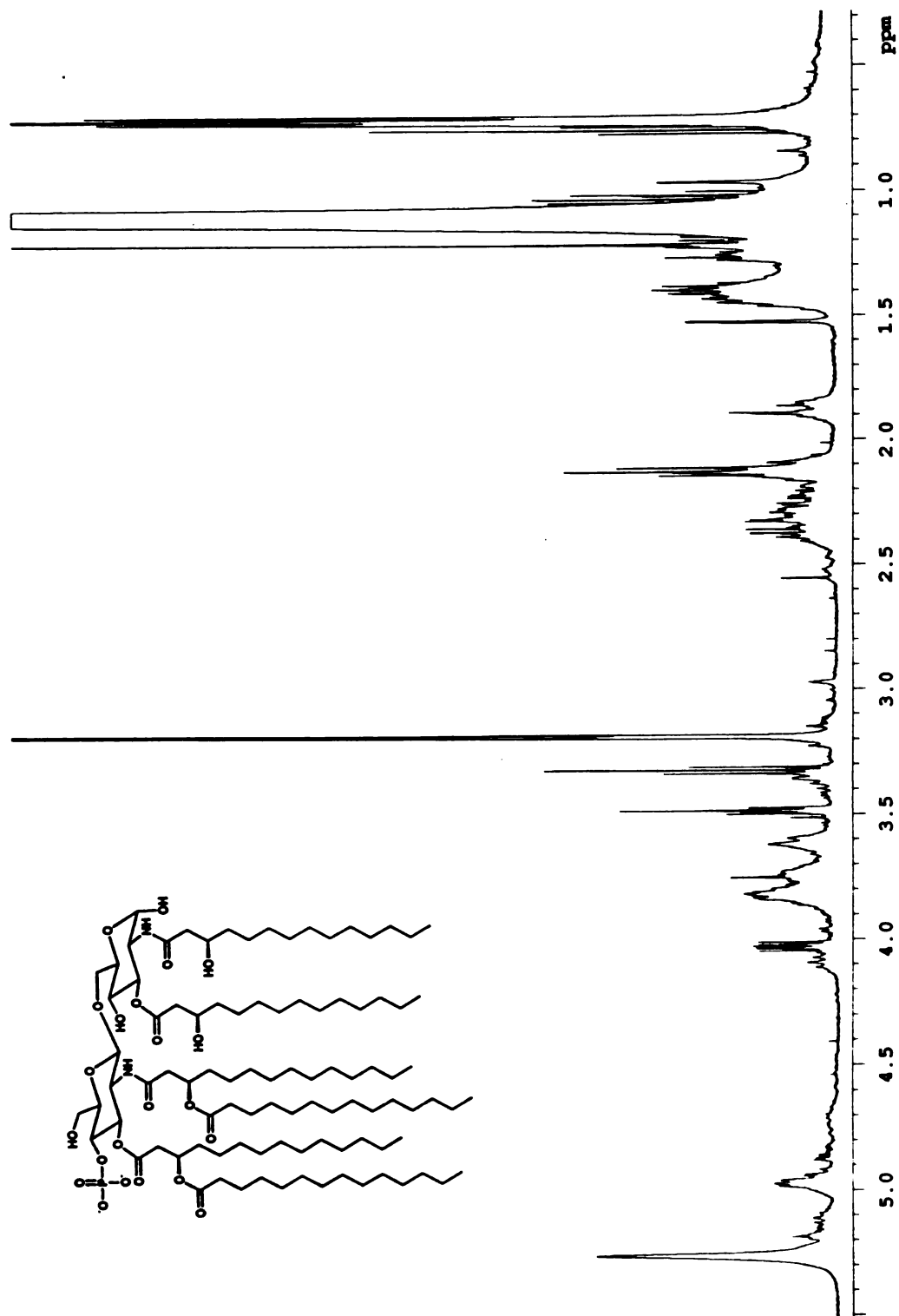
Figure 5.   The ^{1}H NMR spectrum of the diphosphoryl lipid A from E. *coli* in 1:1:2:10 pyridine-D$_5$ / DCl / CD$_3$OD / CDCl$_3$.
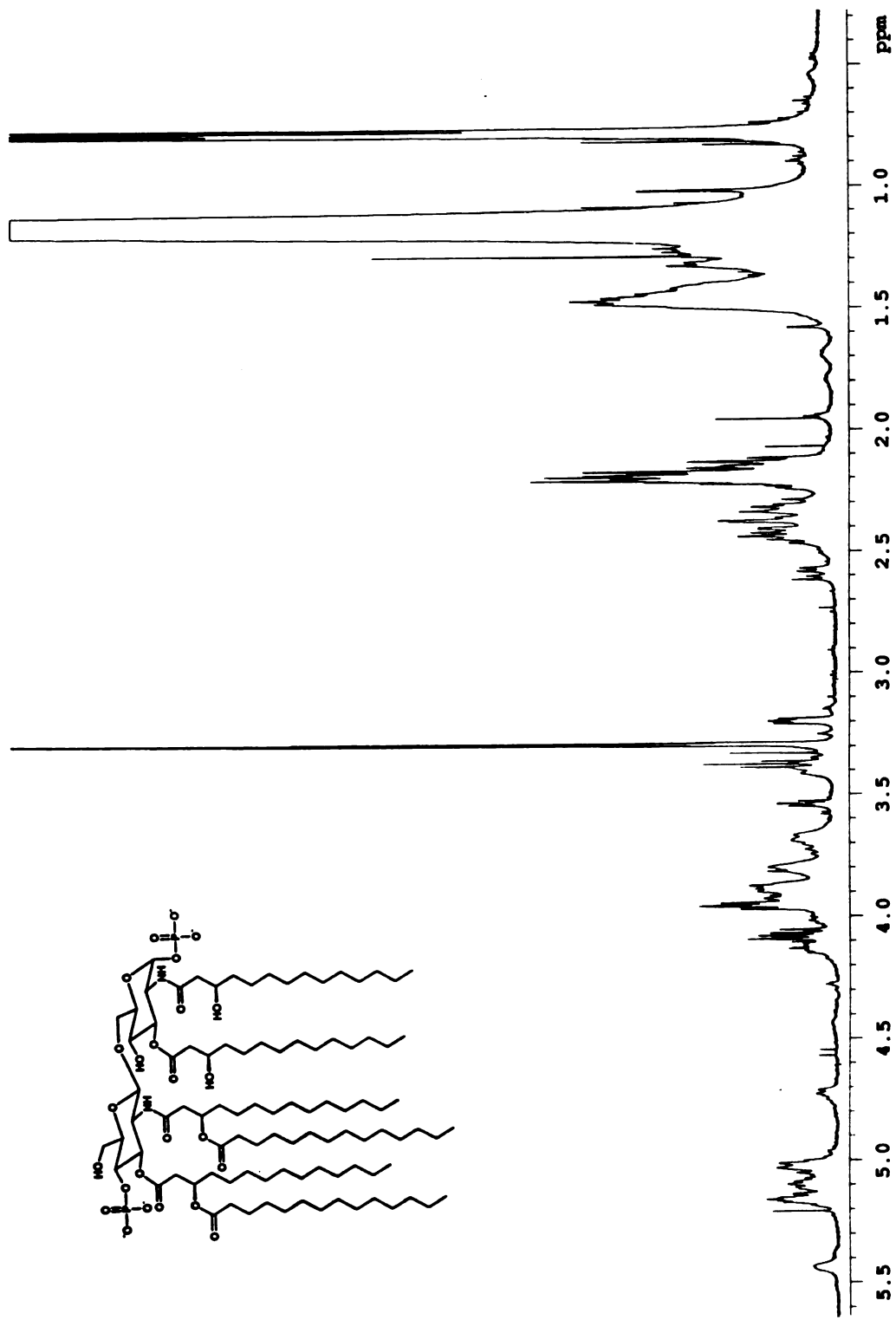
Figure 5

Figure 6.   The ^{1}H NMR spectra of the monophosphoryl lipid A (A) and the diphosphoryl lipid A (B) in CDCl$_3$.

(A)

(B)

Figure 6

(A)

(B)
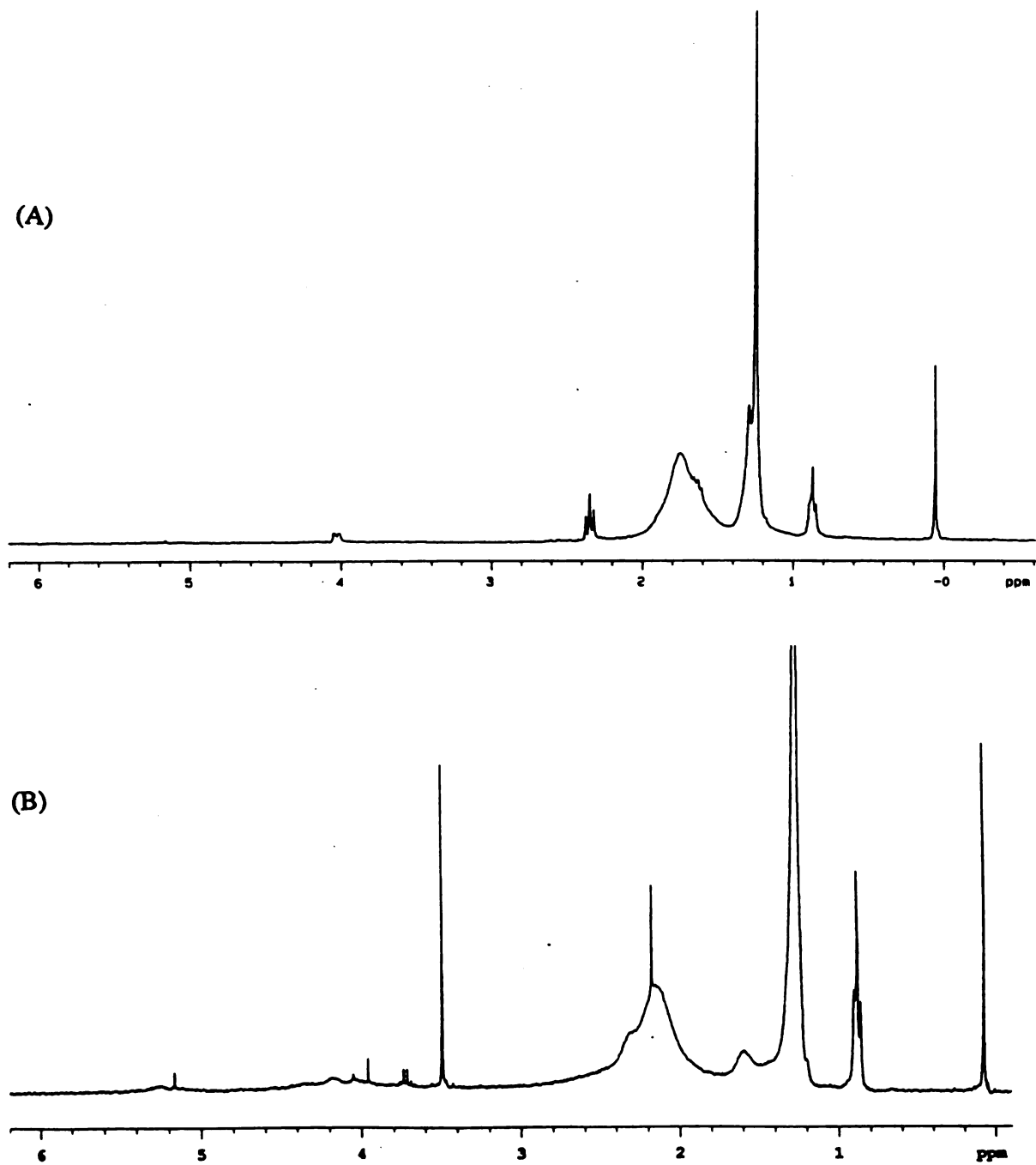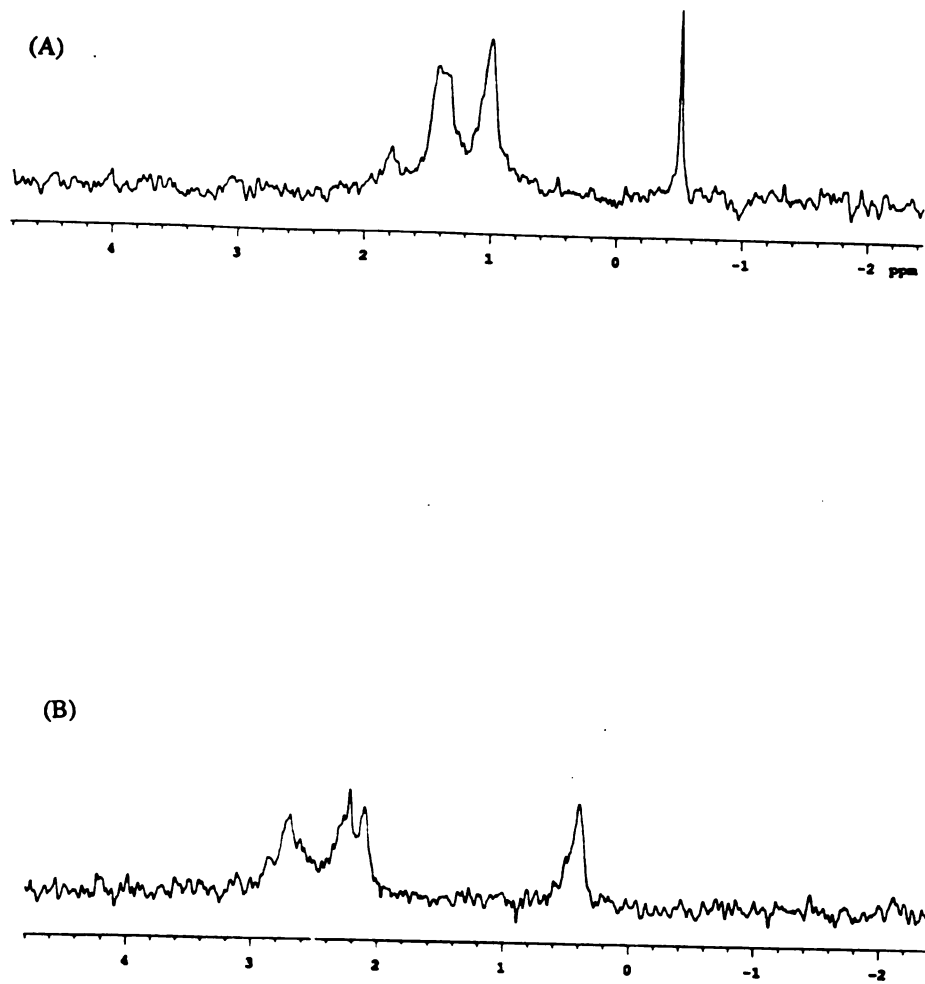
Figure 7.   The ³¹P NMR spectra of the monophosphoryl lipid A (A) and the diphosphoryl lipid A (B).

disaccharide backbone. In addition, the phosphorous resonance was weakly correlated with the proton resonance at 3.9 ppm and 3.8 ppm. The weak correlation arose from long range couplings. The sharp signal at -0.50, apparently from trace quantities of a phosphorylethanolamine group, was correlated with two multiplet at 4.15 ppm and 3.49 ppm. These two multiplet were thus present in the impurity and not part of the lipid A head group. In the case of diphosphoryl lipid A (FIG. 9), the ^{31}P signal at 0.45 ppm was correlated with the ^{1}H-signals at 5.40 ppm. The signal at 5.40 ppm was assigned to the H-1 of the disaccharide backbone. The ^{31}P signal at 2.5 ppm was correlated with the signal at 3.98 ppm, which was thus assigned to the H-4'. Long range couplings were also observed through the low intensity cross peaks.

Gangliosides. Gangliosides are a class of fairly complex lipids. Disialoganglioside GD1a, as an example, is composed of fatty acids, sphingosine, galactose, glucose, galactosamine, and sialic acid. It was, however, soluble in the present solvent system and gave well-resolved ^{1}H-NMR spectrum (FIG. 10). The dominant signals between 0.6-1.5 ppm were assigned to the methylene and methyl protons of fatty acid chains. The multiplet at 2.15 ppm was assigned to the allylic methylene protons of the sphingosine base, while the resonance at 5.3 ppm and 5.6 ppm were assigned to the -CH=CH- protons. The triplet at 2.5 ppm was assigned to the methylene protons adjacent to a carbonyl group. The resonances between 2.6-3.0 ppm were assigned to the methylene protons at the 3-position of the sialic acid residues. The carbohydrate proton resonances appeared in the region of 3.0-4.7 ppm. The four sets of doublets between 4.4-5.0 ppm were assigned to the anomeric protons of the galactose, glucose, and galactosamine residues, while the doublets were due to the axial-axial coupling with the 2-position protons. In addition to the disialoganglioside, the sample appeared to contain other contaminating lipids (probably free fatty acids) as well. This was suggested by the ratio of fatty acid signals to carbohydrate signals.
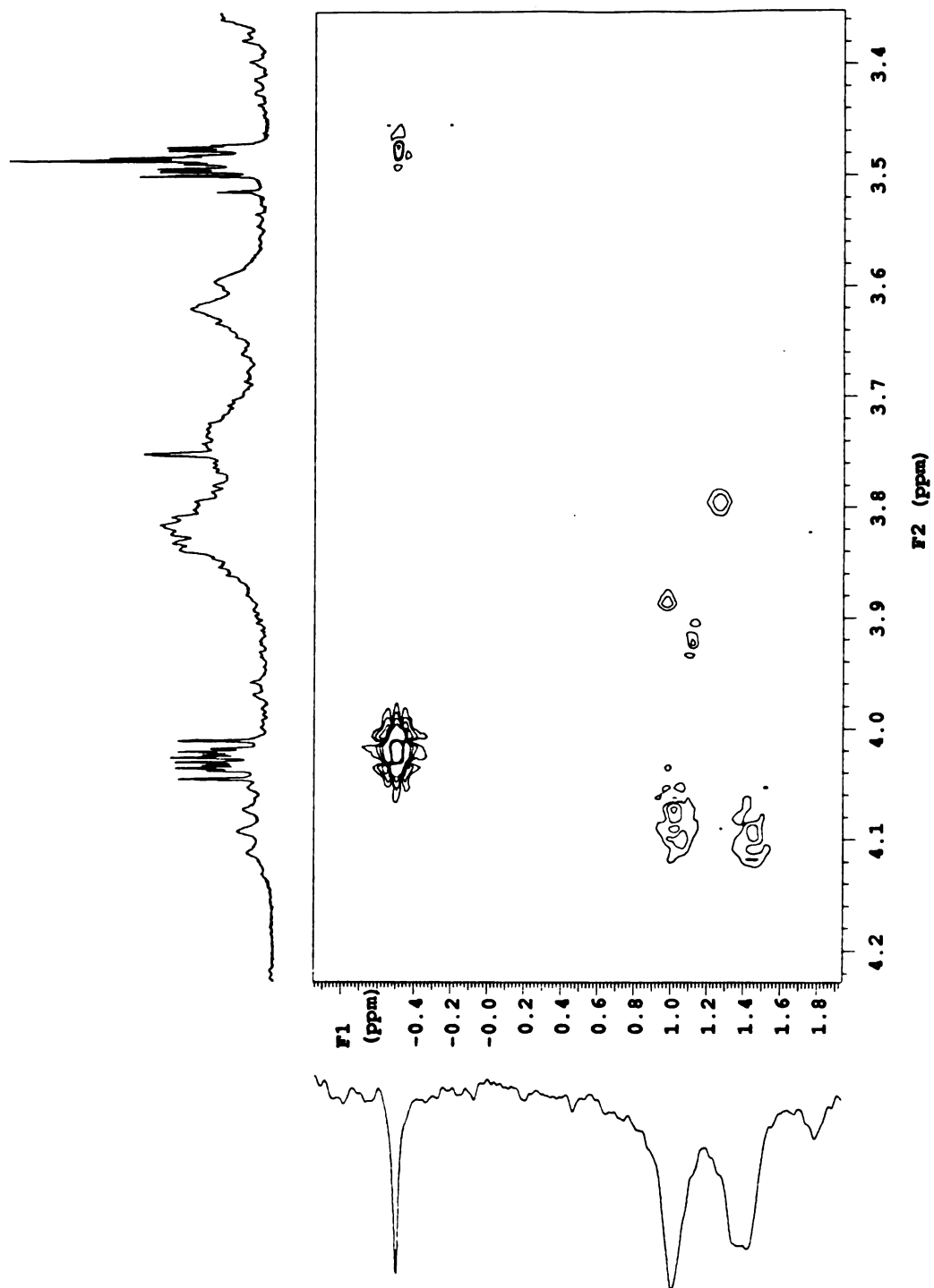
Figure 8.   The ^{1}H-^{31}P HMQC spectrum of the monophosphoryl lipid A.
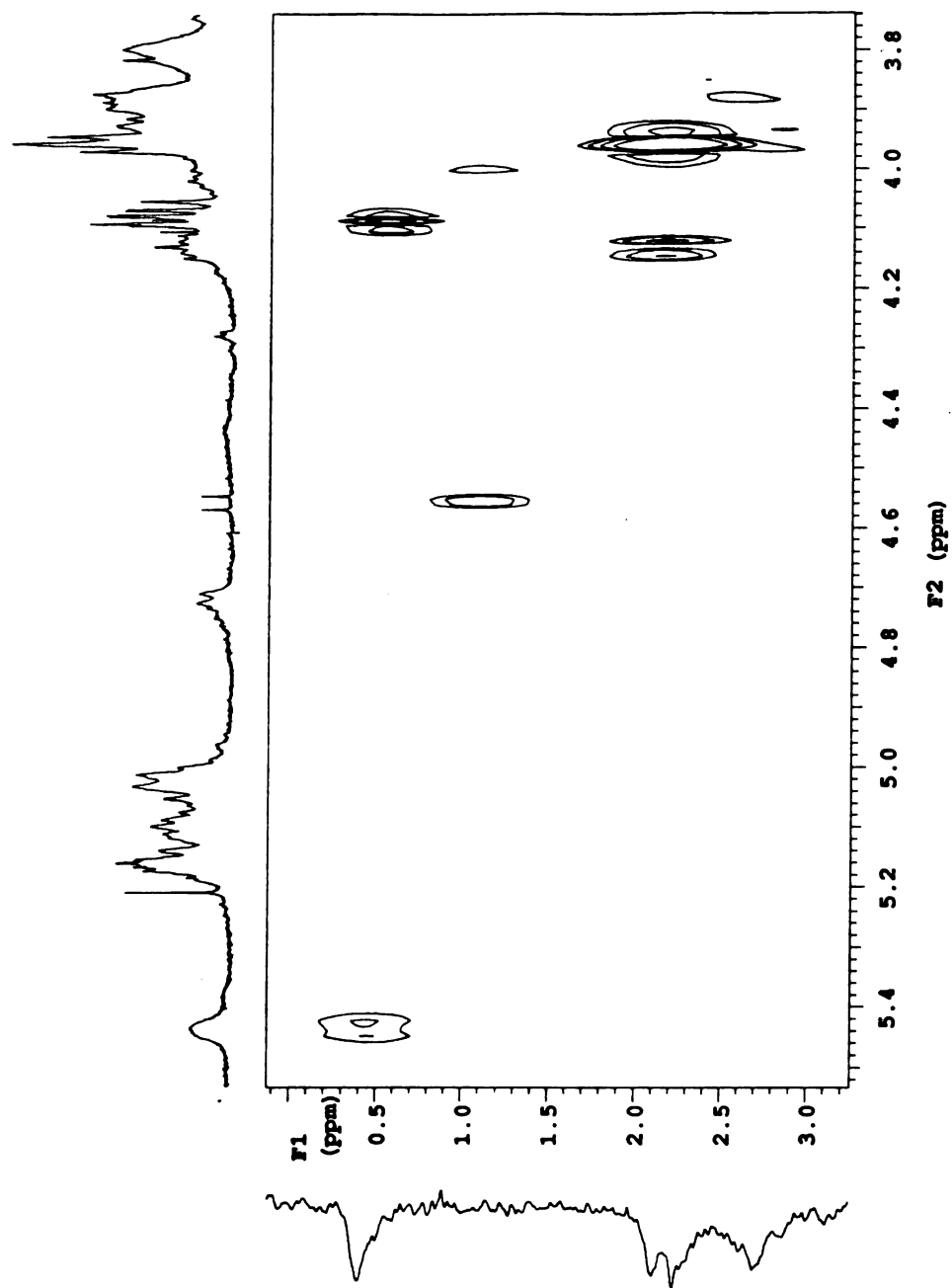
Figure 9. The ^{1}H-^{31}P HMQC spectrum of the diphosphoryl lipid A.

Figure 10. The ¹H NMR spectrum of the disialoganglioside.

Figure 10

DISCUSSION

Phospholipids, lipopolysaccharides and gangliosides are examples of complex lipids which are found in cell membranes. Phospholipids occur almost universally. Lipopolysaccharides are found on bacterial cell surfaces and gangliosides are found on cells of the nervous system. These molecules represent a fairly broad spectrum of lipid structures. The primary reason why structural analyses by NMR spectroscopy is difficult with these lipids is because of their asymmetry in polarity. They all possess strongly polar or hydrophilic regions which tend to interact with the similar regions of other molecules or with each other. There are also hydrophobic regions consisting of long alkyl chains which also tend to self-associate. This leads to aggregate formation. Because of this molecular association, the motion of lipid molecules in solution tend to be much slower than one would predict based on molecular weights of each molecule. The rotational correlation times (τ_c) are much larger than expected and the NMR tranverse or spin-spin relaxation times (T_2) tend to be very small. This leads to line broadening and poor signal resolution. In severe cases, the presence of molecular anisotropies, because of the relatively slowly tumbling structures, leads to unusual line shapes or powder spectra. The approach taken here was to remove the bulk of these intermolecular forces by using large polarizable groups as counter ions for the charged regions of the lipid molecules. The pyridinium ion is ideal for such a function. The choice of chloroform in the solvent mixture was aimed at combating the hydrophobic interactions of the alkyl chains. Methanol was used to maintain a one-phase system. It should by noted that of these solvent components, only methanol absorbs in the area of usual interest in the NMR spectra of lipids.

Although there is not much structural similarity between the classes of lipids used in this study, the results were uniformly good. The phospholipids used in the study were two with different head group charges. This did not affect the results to any great degree since both classes of lipids gave excellent spectra. The choice of ganglioside was made

deliberately because of the carbohydrate chain complexity. A very well resolved spectrum was obtained showing well resolved signals for the carbohydrate protons and the sphingosine base.

Both the proton and ^{31}P NMR spectra of the lipid A molecules were very informative. No attempt was made to make a detailed analysis of the spectra but the NMR signal for the protons at the sites of phosphorylation, the anomeric resonances and the signal for protons at the sites of acylation were readily observable. The spectrum was of high quality. The sample of monophosphorylated lipid A contained some phosphoryl ethanolamine as a contaminant.

The results obtained here have important implications for the use of NMR spectroscopy as a general method for studying the structures of complex lipids and lipid mixtures. The formulation is simple and the mixture can be premade and stored. The spectra are good over a wide spectrum of lipid classes.

REFERENCES

1. Ravindranath, M. H.; Tsuchida, T.; Morton, D. L.; Irie, R. F. *Cancer* **1991**, 67, 3029-3035.

2. Sparling, M. L.; Zidovetzki, R.; Muller, L.; Chan, S. I. *Anal. Biochem.* **1989**, 178, 67-76.

3. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, 65, 355-360.

4. Bax, A.; Subrramanian, S. *J. Magn. Reson.* **1986**, 67, 565-569.

5. Clore, G. M.; Gronenborn, A. M. *Science* **1991**, 252, 1390-1399.

6. Kriat, M.; Vion-Dury, J.; Confort-Gouny, S.; Favre, R.; Viout, P.; Scciaky, M.; Sari, H.; Cozzone, P. J. *J. Lipid Res.* **1993**, 34, 1009-1019.

7. Williams, P. G.; Helmer, M. A.; Wright, L. C.; Dyne, M.; Fox, R. M.; Holmes, K. T.; May, G. L.; Mountford, C. E. *FEBS Lett.* **1985**, 192, 159-164.

8. Brocca, P.; Acquotti, D.; Sonnino, S. *Glycoconjugate J.* **1993**, 10, 441-446.

9. Imoto, M.; Kusumoto, S.; Shiba, T.; Naoki, H.; Iwashita, T.; Rietschel, E. T.; Wollenweber, H. W.; Galanos, C.; Luderitz, O. *Tetrahedron Lett.* **1983**, 24, 4017-4020.

10. Ernst, R. R.; Bodenhansen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Oxford University Press: Oxford, 1987; pp 431-440.

CHAPTER 4

MAJOR REVISION OF THE PROPOSED STRUCTURE OF THE LIPID A REGION OF THE LIPOPOLYSACCHARIDES OF *RHIZOBIUM LEGUMINOSARUM* BIOVARS

ABSTRACT

Major revisions are made to the proposed structure of the lipopolysaccharide lipid A of *Rhizobium leguminosarum* biovars (Bhat, U., Forsberg, S. & Carlson, R. W., (1994) *J. Biol. Chem.*, 269, 14402-14410). Features of the previously proposed structure include a trisaccharide headgroup unit in which one residue was a glucosamine aldonic acid, an α-1,4-linkage between a galacturonic acid residue and a glucosamine residue the absence of acyloxyacyl fatty acids and a complete lack of acylation of the galacturonic acid residue. We show here that, among other revisions: The lipid A of *Rhizobium leguminosarum* biovar *trifolii* is a disaccharide and not a trisaccharide and does not contain a glucosamine aldonic acid residue; The linkage between the galacturonic acid residue and glucosamine is 1,6- and not 1,4-; The galacturonic acid residue bears a total of four acyl groups, three of which are fatty acids including 27-hydroxyoctacosanoic acid and that acyloxyacyl groups are present. Other findings described here include the fact that 3-hydroxybutyrate is an acyl substituent of the lipid A (previously reported by Bhat et al.) and the new finding that the glucosamine residue bears a lactyl ether at the 1-position. These studies involved extensive use of homonuclear and heteronuclear correlation 1 and 2-dimensional NMR spectroscopy and mass spectrometry using isotope labelling with electron impact and electrospray ionization. These methods were complemented by linkage-specific and site-specific chemical cleavage methods. The revised structure has the same structural and charge elements of the typical lipid A molecule as is observed in enterobacterial lipopolysaccharides: A disaccharide with two negatively charged groups, one on each of the opposite ends of the structure. In this case,

the charged groups are carboxylate and not phosphate as in the *enterobacteriaceae*.

INTRODUCTION

The structure of the lipopolysaccharides of *Rhizobium* has always been an area of interest ever since it became clear that they were, somehow, involved in determining the outcome of the symbiotic process between these bacteria and their legume plant hosts (1-3). Much of the earlier work only reported glycosyl composition but eventually the structure of two core components were elucidated for *R. trifolii* ANU843 (4-6). The structures of the O-antigen components and the lipid A regions, however, defied complete characterization. Eventually, a structure of the lipid A region of *R. trifolii* was proposed (7). Unfortunately, this came before the discovery that 27-hydroxyoctacosanoic acid was a predominant fatty acid in the lipid A of these bacteria (8) and that galacturonic acid was also a component of this molecule (9). These structural features were not accommodated in the first proposed structure. It was also demonstrated that lipid A preparations of *R. trifolii* ANU843 also contained small amounts of glucosamine uronic acid residue to which 27-hydroxyoctacosanoic acid was attached (10). A structure was later proposed for the lipid A of *R. meliloti* (11).

Despite attempts at determining their structures, the lipid A regions of the lipopolysaccharides of the various rhizobia still remain somewhat of an enigma. They represent one of the more radical departures from the well characterized diphosphorylated

glucosamine disaccharide template that so characterize the lipid A regions of the *Enterobacteriaceae* lipopolysaccharides. In a recent publication (12) a structure was proposed for the lipid A region of *Rhizobium leguminosarum* bv *phaseoli* and indeed all of the *Rhizobium leguminosarum* biovars of which *Rhizobium leguminosarum* bv *trifolii* ANU 843 is a member. This was an unusual structure that accommodated the more recent findings of galacturonic acid and 27-hydroxyoctacosanoic acid (8, 9) and included the presence of a third carbohydrate residue (an aldonic acid derivative of glucosamine) with a tentative fatty acid residue at its 5-position (Figure 1). This proposed structure was based almost exclusively on mass spectral and methylation analysis data. In these analyses, there were stated problems such as under methylation and, apparently, deacylation and ester migrations. There were also no supporting NMR spectroscopic analyses to substantiate the proposed structure, several elements of which conflicted with the findings in our laboratory.

The major strategy we used in this study was to undertake a thorough NMR spectroscopic analysis of the lipid A molecule with and without the fatty acyl groups. Such analyses are usually not possible with lipid A molecules because of extensive line broadening due to small transverse relaxation times. This is the result of extensive aggregate formation. Our NMR spectroscopic analyses were facilitated by the recent development of a multicomponent solvent system in our laboratory which gives well resolved spectra for all classes of lipids (13). The ability to acquire well resolved, high resolution spectra of the lipid A molecule made it possible to perform several NMR spectroscopy experiments. These included DEPT to determine the number of primary hydroxyl groups present and hence the

number of hexoses with primary alcohols at their C-6 positions; $^1H/^{13}C$ HMQC experiments to confirm the number of nitrogen-bearing carbons in the structure as well as the anomeric linkages of the sugar residues and linkage position; and TOCSY and DQF-COSY to trace the spin connectivities and assign the positions of substitution. Our second strategy was to effect selective de-N-acylation and elimination of acyl groups at the 3-position of the galacturonic acid residue as well as dehydration of unsubstituted 3-hydroxy fatty acyl groups. This was accomplished by the use of a reaction involving 4-dimethylamino pyridine and trifluoroacetic anhydride (14). In the presence of these reagents, hydroxyl groups so placed as to give an α,β-unsaturated system on elimination do so readily on trifluoroacetylation in the presence of dimethylamino pyridine which acts as a base. The NH proton of amides are readily acylated and the product readily hydrolyzed in the presence of water to form the free amine and carboxylic acid. In the case of a galacturonic acid residue, once the 4-hydroxyl group has eliminated to form a hex-4-enopyranosyl uronic acid residue (an α,β-unsaturated system) during prolonged reaction, protonation of the 3-substituent results in ready elimination to extend the conjugation. Yet another aspect of the approach was to develop conditions that allowed selective O-deacylation or total deacylation. This allowed unambiguous assignment of the N-linked fatty acids. These analyses are supported by a wealth of data from several mass spectrometry experiments including unambiguous deuterium labelling studies. Our findings are now presented here. Among several other aspects to be addressed by way of revision, we show that, contrary to the findings of Bhat et. al., the lipid A of *Rhizobium leguminosarum* bv *trifolii* does not contain a glucosamine aldonic acid residue at its terminus, that the linkage between the galacturonic acid residue and

glucosamine is 1,6 and not 1,4 as was proposed, and that the galacturonic acid residue is substituted with three fatty acid residues and is not unsubstituted as was proposed. We further report that a 3-hydroxybutyryl residue is esterified to the 27-hydroxyoctacosanoic acid residue, and that the reducing end of the glucosamine residue is substituted by a lactyl group in a glycosidic linkage. Our proposed structure is shown in Figure 2.

MATERIALS AND METHODS

Bacterial Cultures and LPS and Lipid A Isolation. *Rhizobium trifolii* ANU 843 was grown to early stationary stage in modified Bergensen's (BIII) medium at 30 °C. The LPS was isolated and purified as described previously (10). The purified LPS (about 15 mg) was hydrolyzed with 10 ml of 1% acetic acid at 100 °C for 2 hrs. After cooling to room temperature, the solution was extracted with 4 volumes of 5:1 chloroform / methanol. The organic layer containing lipid A was collected after centrifugation (8000 rpm for 10 min) and the aqueous layer was again extracted with chloroform / methanol. The combined organic extracts were dried under a stream of nitrogen.

Figure 1. The proposed structure of *R. leguminosarum* lipid A as proposed by Bhat et al. (1994). R_1 is variable and can be 3-OH 14:0, 3-OH 16:0, 3-OH 18:0, 3-OH 15:0, or 27-OH 28:0. It is predominantly 3-OH 14:0 and 27-OH 28:0. R_2 can be 3-OH 14:0, 3-OH 16:0, or 3-OH 18:0. The GlcN N-acyl residue is predominantly 3-OH 14:0, while the major GlcN-onate N-acyl is 3-OH 16:0 or 3-OH 18:0. No acyl substituents are present on the GalA residue. There are a total of 5-6 fatty acyl residues/GlcN indicating that all available positions of the GlcN (C-1, C-2, and C-3) and of the GlcN-onate (C-2, C-3, C-4, and C-5) are acylated. The linkage of the 27-OH 28:0 residue to C-5 of GlcN-onate is hypothetical and the linkage of the GlcN-onate residue at C-6 is assumed.
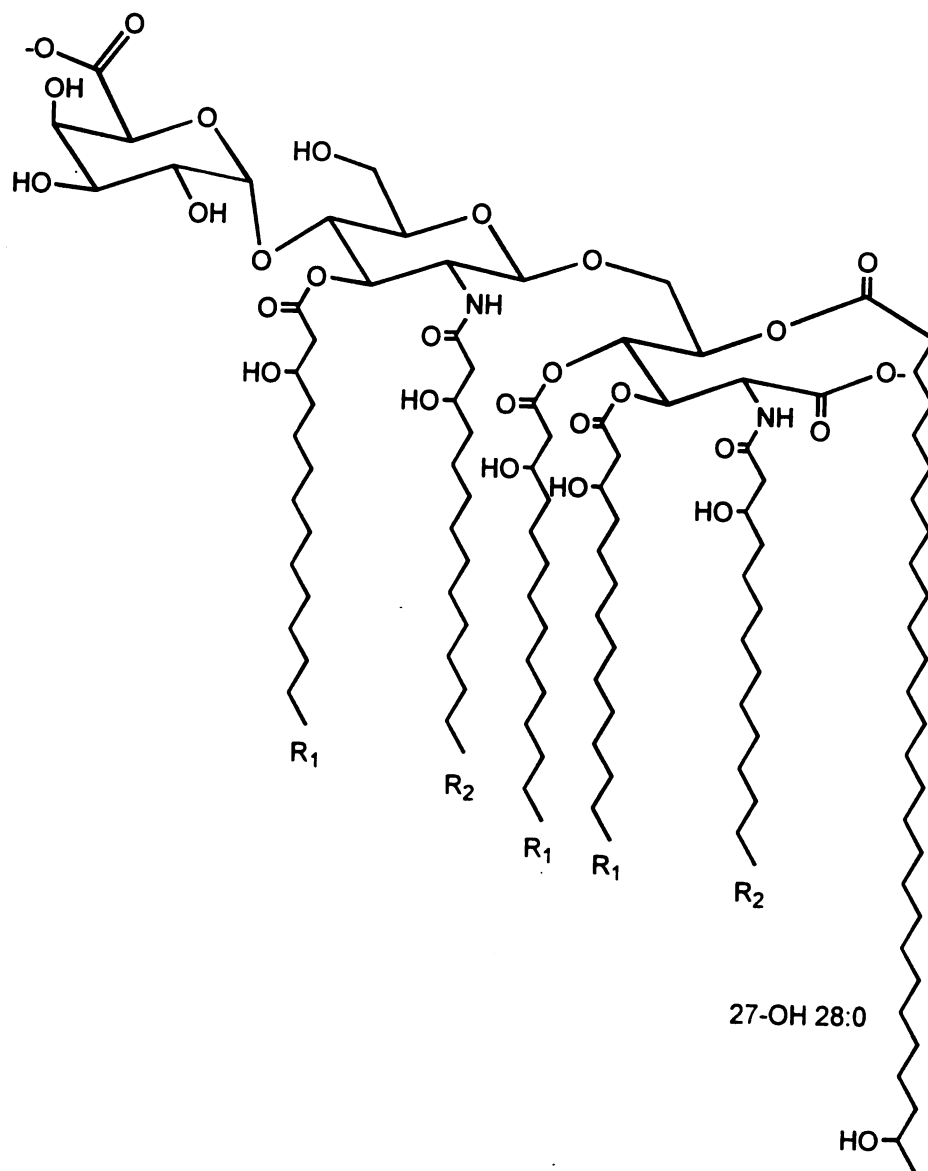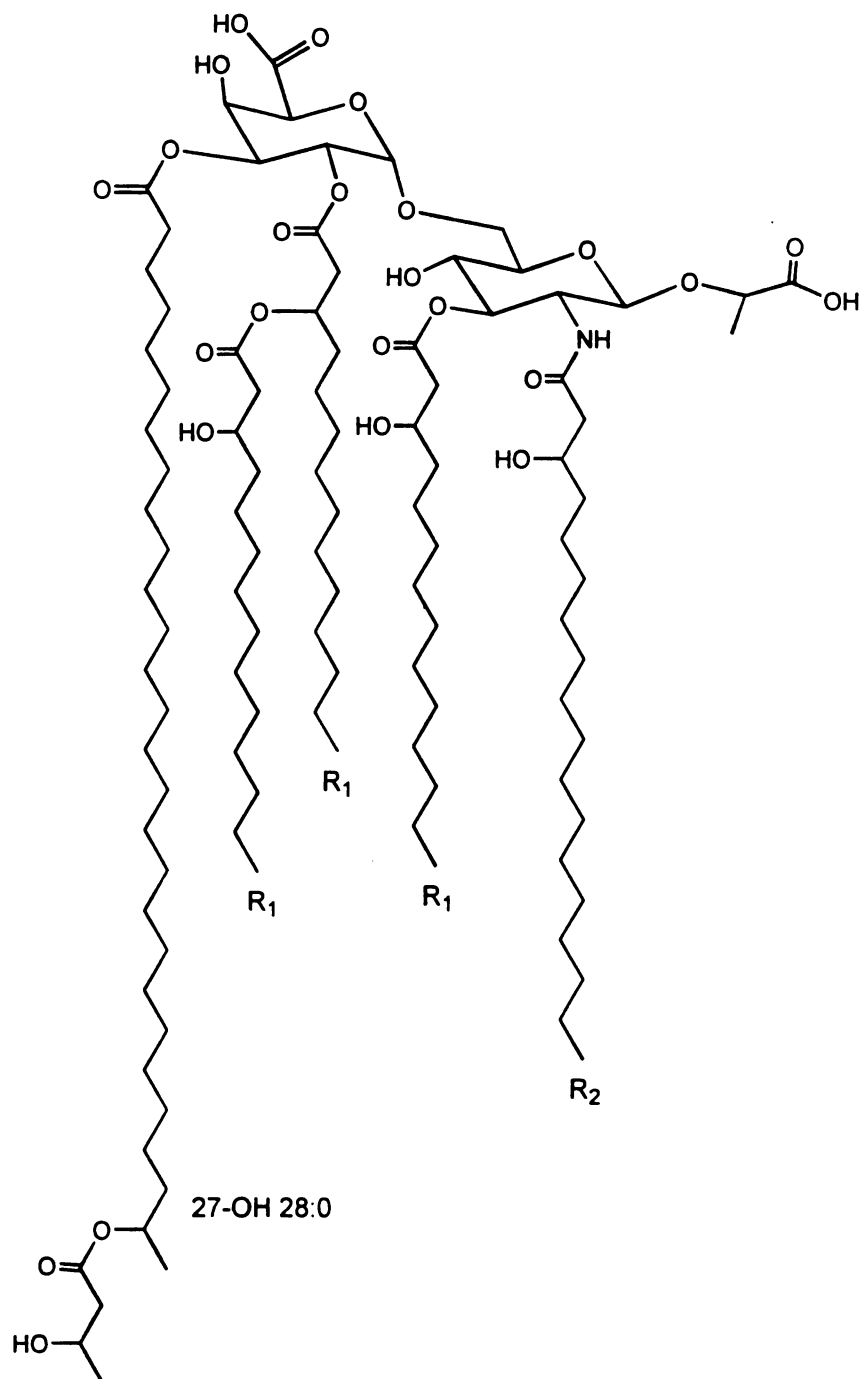
27-OH 28:0

Figure 1

Figure 2. The revised structure of *R. leguminasorum* lipid A. R₁ can be 3-OH 14:0, 3-OH 15:0, 3-OH 16:0 and 3-OH 18:0, but usually 3-OH 14:0. R₂ can be 3-OH 16:0, 3-OH 18:0 and 3-OH 14:0, but usually 3-OH 16:0 or 3-OH 18:0. There are 4 acyl residues on GalA and 2 fatty acyl residues on GlcN.

Figure 2

Fatty Acid Analysis. A sample (1 mg) of lipid A was placed in 1 ml of 5%

hydrochloric acid in methanol and methanolyzed in a Teflon-lined screw-cap vial at 70 $^{\circ}$C

for 30 hrs. The methanolysis product was cooled to room temperature, concentrated to

dryness under a stream of nitrogen and then partitioned between 4 ml of hexane and 2 ml

of water. The hexane fraction was collected and dried under nitrogen and subjected to GC

and GC-MS analysis. GC analysis was performed on a Hewlett Packard model 5890 gas

chromatograph using a DB1 capillary column with helium as the carrier gas. A

temperature program was used as follows: an initial column temperature 150 $^{\circ}$C was held

for 2 min, the temperature was then increased to 300 $^{\circ}$C at 3 $^{\circ}$C/min and held at this value

for 20 mins. Mass spectra were acquired on a JEOL JMS-AX505H mass spectrometer.

De-O-acylation of Lipid A. A sample ($\sim$ 2 mg) of dried lipid A was added 1 ml

of 1.5 M potassium hydroxide in 1-propanol containing 2 mg of sodium borohydride and

0.5 mg of thiophenol. The solution was heated at 120 $^{\circ}$C for 2 hrs and then heated at 74

$^{\circ}$C for a further 20 hrs. After cooling to room temperature, the reaction mixture was

acidified with formic acid and then extracted with 3 ml of chloroform. The aqueous

fraction was dried on a rotary evaporator and then separated on a Bio-gel G10 column

eluting with water. Fractions of $\sim$ 2 ml were collected and those containing carbohydrate

were identified by phenol / sulfuric acid assay (15). They were pooled and dried on a

rotary evaporator. The chloroform fraction was partitioned between 2 ml of 4:1 hexane

/ toluene and 1 ml of 10:1 methanol / water containing 1 % formic acid. The lower

hexane / toluene layer was collected and subjected to ^{1}H-NMR spectroscopic and fatty acid

analysis.

Total Deacylation of Lipid A. This was carried out as described for the de-O-acylation except that n-butanol (1.5 ml), 6M sodium hydroxide (1.5ml) and thiophenol (20 mg) were used. After heating for 2 hours at 120 °C, the reaction mixture was cooled and 2 ml of ethanol were added. The mixture was then heated again at 75 °C for 16 hours and then at 120 °C for a further 2 hours. The mixture was cooled, acidified with formic acid and partitioned between 1:1 chloroform / hexane (20 ml) and water (20 ml). The aqueous fraction was analyzed by ^{1}H-NMR spectroscopy and then subjected to gel filtration chromatography as described above for the de-O-acylation reaction.

De-N-acylation of Lipid A and Dehydration of Free 3-hydroxyl Groups of Fatty Acids. Lipid A sample (2 mg) was treated with 10 mg of 4-dimethylaminopyridine (DMAP) and 1 ml of trifluoroacetic anhydride (TFAA). The solution was sonicated for 2 min and heated at 70 °C for 16 hrs. After cooling to room temperature, the reaction was quenched with 1 ml of water. The solution was then extracted with hexane (2x2 ml). The combined hexane extracts were dried under a stream of nitrogen and partitioned between the polar and nonpolar phases of a mixture of 2 ml of hexane and 2 ml of methanol containing about 50 μl of water. The hexane layer was dried and partitioned again between the phases of a biphasic system composed of chloroform, methanol, water and hexane in a ratio of 1 : 4 : 0.01 : 4 respectively. The layers were analyzed by ^{1}H-NMR spectroscopy and the top layer was adsorbed onto a C-18 reverse phase cartridge in 9 : 1 methanol / water and eluted with 6 ml fractions each of methanol / water (9 : 1), ethanol

and chloroform in that order. These fractions were analyzed by ^{1}H-NMR spectroscopy and by GC and GC-MS after converting the fatty acids to their methyl esters.

Glycosyl Composition Analysis. The glycosyl fraction obtained after de-acylation of the lipid A was dissolved in hydrochloric acid in methanol (1%) and heated at 40 °C for 2 hrs. Methanol (0.1 ml) was added and the sample was dried under nitrogen. Water (0.5 ml) containing sodium borodeuteride (1 mg) was added and reduction allowed to proceed at room temperature for 1 hr. To the reduced product was then added 2.4 N hydrochloric acid dropwise until effervescence ceased and the solution was dried under nitrogen. The reduced product was then converted to alditol acetate derivatives according to a procedure described previously (16).

NMR Spectroscopy. The ^{1}H-NMR spectra of fatty acid samples were recorded in chloroform-d at 300 MHz on a VARIAN VXR300 spectrometer. All other spectra were acquired at 35 °C at 500 MHz for ^{1}H and 125 MHz for ^{13}C on a VARIAN VXR500 spectrometer. The lipid A sample was dissolved in a solvent system (13), composed of a mixture of pyridine-d_5 / 37 wt% deuterium chloride in deuterium oxide / methanol-d_4 / chloroform-d in a volume ratio of 1:1:2:10 respectively. The proton chemical shifts are quoted relative to the $CDCl_3$ line at 7.24 ppm and the ^{13}C chemical shifts are relative to $CDCl_3$ line at 77.0 ppm. DQF-COSY (17) was acquired using a total of 256 data sets with 32 transients at 2048 data points each. Data for the TOCSY experiment (18) was obtained using similar acquisition and processing conditions. A mixing time of 80 ms was used.

Mass Spectrometry. FAB mass spectra were recorded at a JEOL JMS-HX110

mass spectrometer with dinitrobenzyl alcohol as the matrix in negative ion mode. ES-MS

spectra were recorded on a Fisons Platform mass spectrometer. Samples were dissolved

in 1% ammonium hydroxide in chloroform / methanol (5 : 1).

RESULTS AND DISCUSSION

The glycosyl and lipid composition of the lipid A was in agreement with published

results (19, 20). There were 5 fatty acids (Figure 3), namely, 3-hydroxytetradecanoic acid

(3-OH 14:0), 3-hydroxypentadecanoic acid (3-OH 15:0), 3-hydroxyhexadecanoic acid (3-

OH 16:0), 3-hydroxyoctadecanoic acid (3-OH 18:0), and 27-hydroxyoctacosanoic acid

(27-OH 28:0), in a molar ratio of 2:0.5:1:1:1. The carbohydrate compositional analyses

indicated that galacturonic acid and glucosamine were present in equal amounts. The

presence of galacturonic acid was clearly defined by the 100% incorporation of two

deuterium atoms at one of the primary positions of the alditol acetate of galactitol during

borodeuteride reduction. This led to doublets of ions of equal intensities at m/z 217/219,

289/291 and 361/363 in the mass spectrum of galactitol hexa-acetate (Figure 4A). These

corresponded to primary fragmentation ions. In contrast, mass spectrometric analyses

indicated that no label was incorporated into the 2-amino-2-deoxyalditol acetate derived

from the amino sugars (Figure 4B). This indicated that no glucosamine aldonic acid was

present as proposed earlier (12). This could be concluded since the two predominent ions
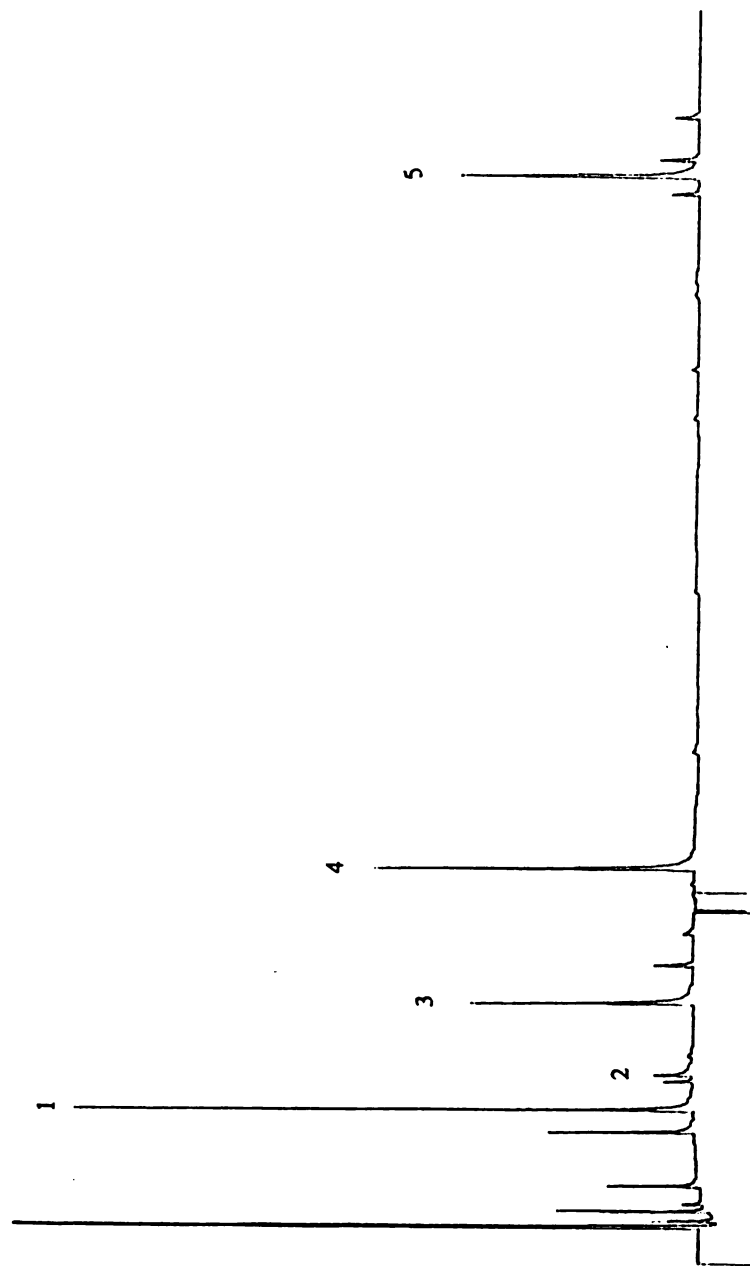
Figure 3. GC profile of the methyl ester derivatives of the total fatty acids released from *R. trifolii* ANU843 lipid A. Signals 1-5 correspond to 3-OH 14:0, 3-OH 15:0, 3-OH 16:0, 3-OH 18:0 and 27-OH 28:0 respectively.

Figure 4. GC-MS fragmentation pattern for prereduced and deuterium labelled alditol acetate derivatives of (A) the galacturonic acid and (B) the glucosamine residue of *Rhizobium trifolii* ANU843 lipid A. All major ions of the GalA derivatives are doublets separated by 2 mass units due to deuterium labelling at the former carboxylic acid end of the alditol. Note that there is no isotope labelling of the GlcN residue.
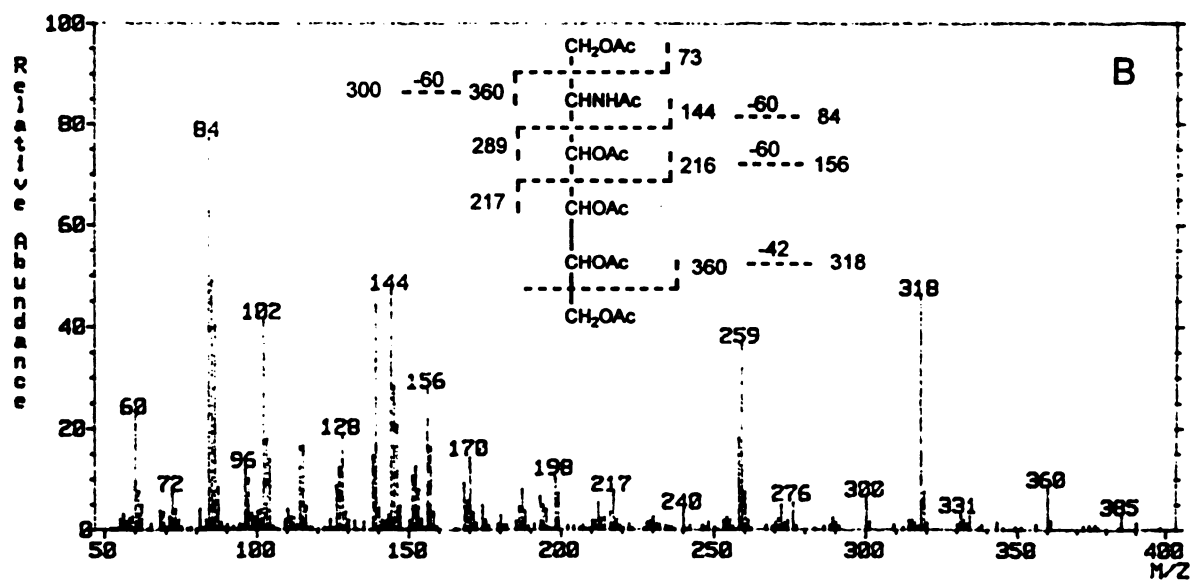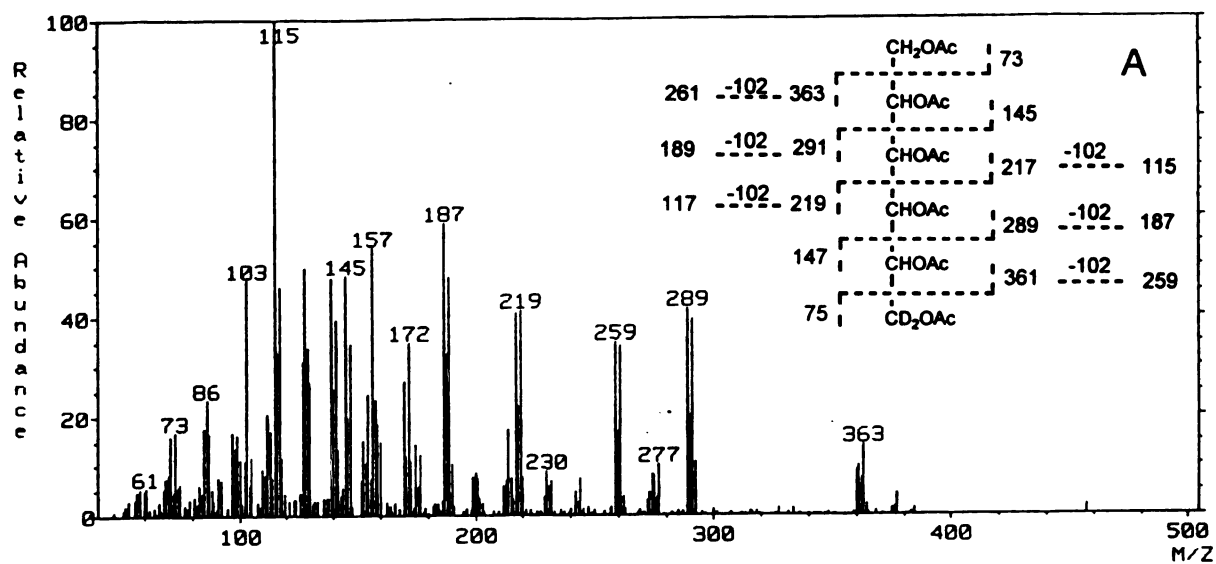
Figure 4

at m/z 144 and 84 would have been accompanied by two other ions at m/z 146 and 86 due

to incorporation of two deuterium atoms. The ions at m/z 144 and 84 arise from scission

between C2 and C3 to yield the first one followed by loss of acetic acid to yield m/z 84.

Since the glucosamine peak contained predominant ions at m/z 144 and 84 and not 145 and

85, it could also be concluded that the reducing end was not free but was involved in

a glycosidic linkage and, therefore, not susceptible to reduction.

The ester and amide-linked fatty acids could be distinguished by selective de-O-

acylation of the lipid A with thiophenol under basic conditions. This yielded a mixture

of free fatty acids and the N-acylated lipid A headgroup. The free fatty acids partitioned

into the hexane fraction of a hexane / aqueous methanol biphasic system. Analysis of the

free fatty acid fraction by ^{1}H-NMR spectroscopy confirmed the presence of 3-hydroxy

fatty acids (the AB component of an ABX spin system between 2.5 and 2.8 ppm and a 1-H

multiplet at 5.48 ppm) as well as that of 27-hydroxyoctacosanoic acid (a sharp doublet at

1.18 ppm and a sextet at 4.91 ppm). The identities of the 3-hydroxy fatty acids were

determined by GC and GC-MS to be 3-OH C14, 3-OH C15, 3-OH C16, and 3-OH C18

in an approximate 3:2:1:1 molar ratio (Figure 5A). The presence of 27-

hydroxyoctacosanoic acid was also confirmed by this technique. The amide-linked fatty

acids were then released from the headgroup by methanolysis and their identities

determined by GC and GC-MS to be 3-OH C16, 3-OH C18 and small amount of 3-OH

C14 (Figure 5B). These results clearly showed that the only fatty acid that was

Figure 5. (A) GC profile of the methyl ester derivatives of the O-linked fatty acids. (B) GC profile of the methyl ester derivatives of the N-linked fatty acids. Peaks 1-5 are from 3-OH 14:0, 3-OH 15:0, 3-OH 16:0, 3-OH 18:0 and 27-OH 28:0 respectively. The molar ratio is approximately 3:2:1:1:1. N-linked fatty acids are 3-OH 14:0, 3-OH 16:0 and 3-OH 18:0. Note that 27-OH 28:0 is not present in (B) and, therefore, is exclusively O-linked.
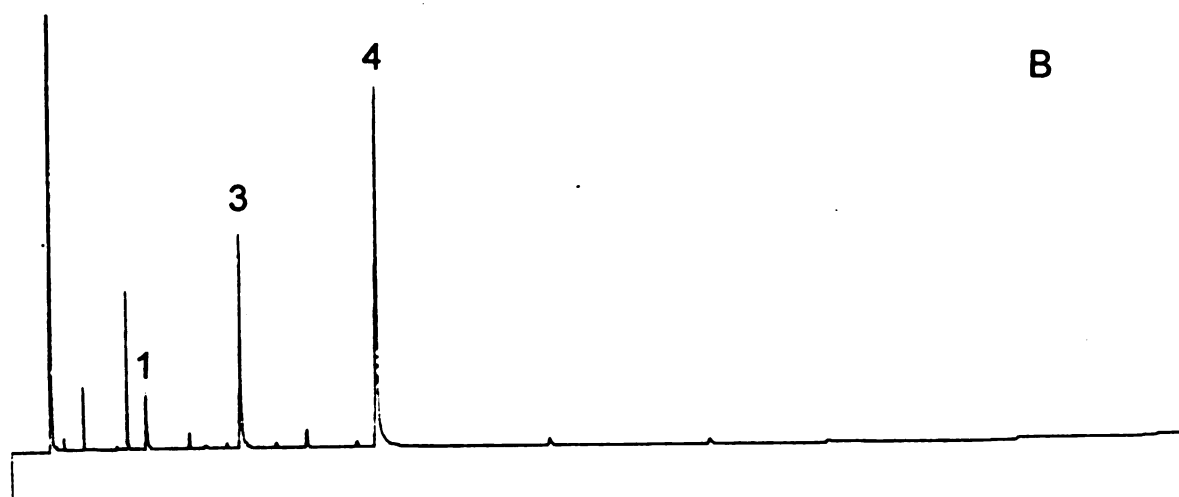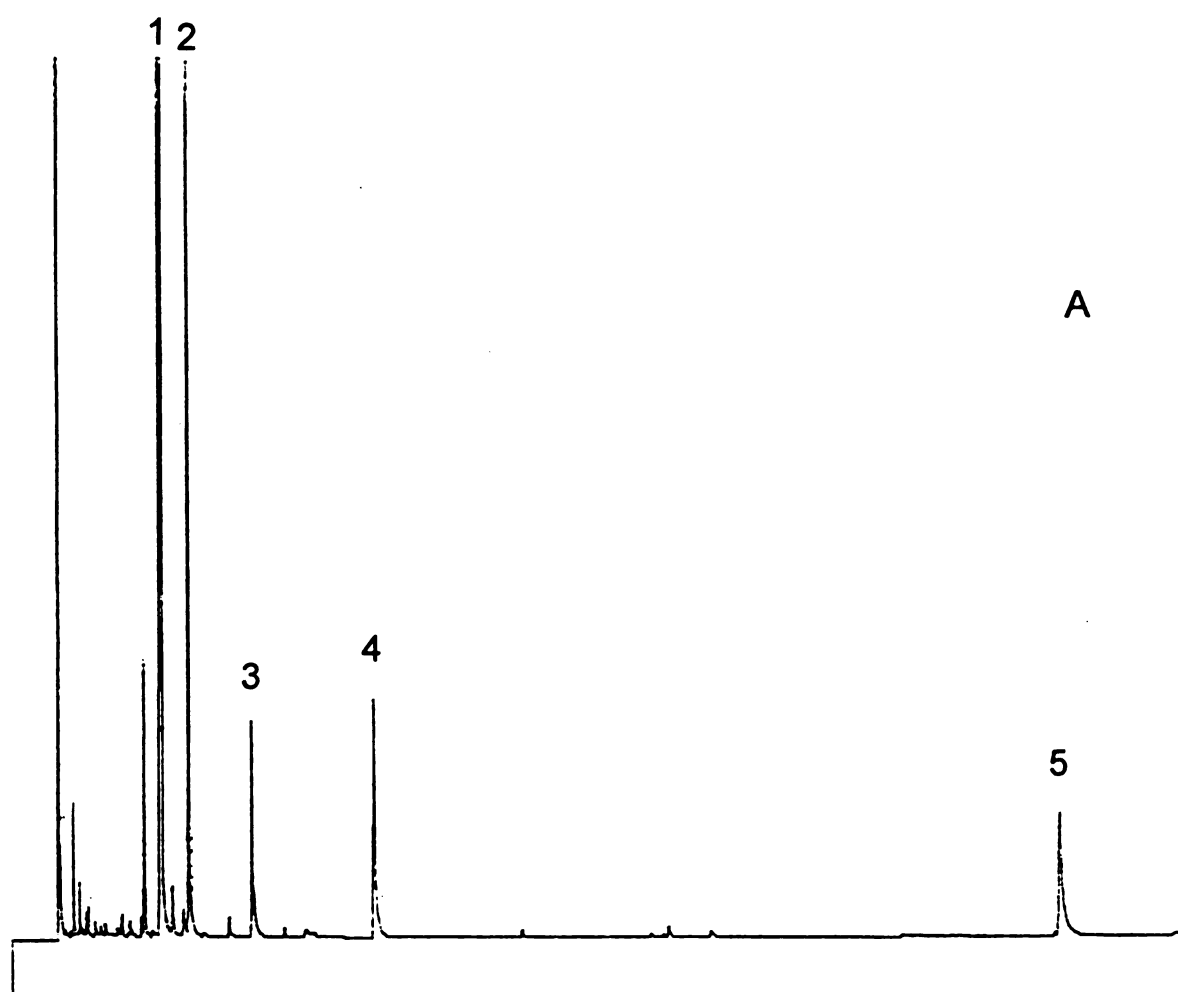
Figure 5

ester-linked was 27-hydroxyoctacosanoic acid.

The O-deacylated lipid A headgroup was further analyzed by negative ion electrospray mass spectrometry. This analysis yielded a spectrum containing two major peaks at m/z 680 and 718 (Figure 6). Since it was already established that galacturonic acid and glucosamine were present in the molecule, in addition to 3-hydroxyhexadecanoic acid or a 3-hydroxyoctadecanoic acid residue (due to heterogeneity), these ions could be assigned to a disaccharide containing galacturonic acid and glucosamine in addition to one or other of these fatty acids plus a substituent with a mass suggesting a lactyl or some other 3-carbon acid residue. It is important to note here that this definitively shows that another 6-carbon glycosyl residue (i.e. glucosamine aldonic acid as proposed by Bhat et al.) is not attached to the lipid A head group. Total de-acylation of the lipid A under more forcing conditions indicated the presence of a lactyl residue since the proton NMR spectrum of the aqueous fraction, after removal of the fatty acids by extraction, clearly showed the presence of a doublet at 1.22 ppm (J = 6.5 Hz) that could be assigned to a lactyl group (Figure 7). This spectrum also confirmed that all of the fatty acids were removed and it indicated that a 3-hydroxybutyryl residue was also associated with the lipid A molecule. This was suggested by the presence of another doublet at 1.08 ppm and a pair of mutually coupled doublet of doublets between 2.1 and 2.4 ppm with chemical shifts and splitting characteristic of 3-hydroxybutyric acid (21).

Information on the sites of substitution of the hydroxyl groups of the fatty acids and the carbohydrate rings was obtained by a selective dehydration and de-N-acylation reaction

Figure 6. ES-MS spectrum of the O-deacylated lipid A. The [M-H]⁻ and [M+K-H]⁻

ions at m/z 680 and 718 suggest the presence of a lactyl group in addition to a GlcA-GlcN

disaccharide and a 3-OH 16:0 residue. The potassium ion comes from the KOH used in

the hydrolysis.
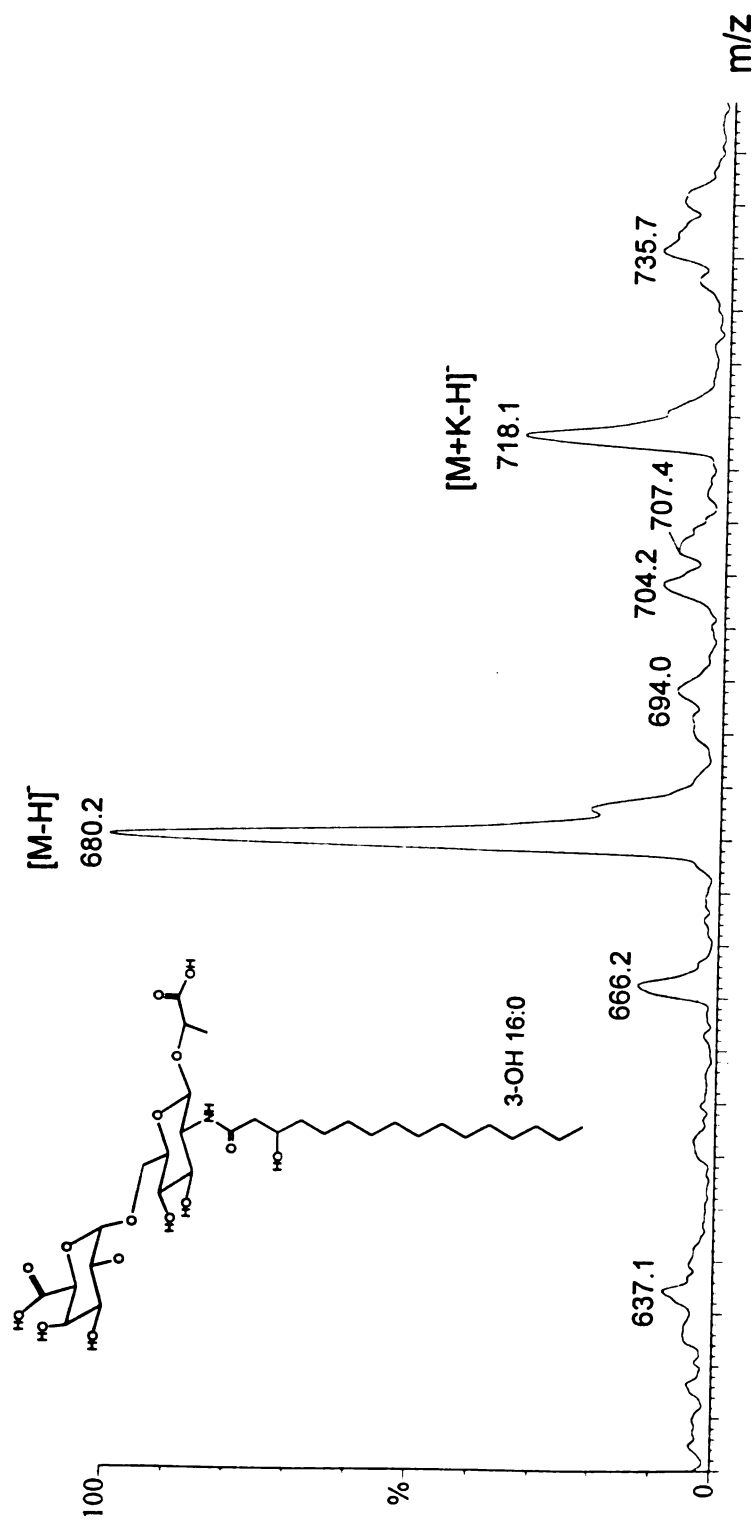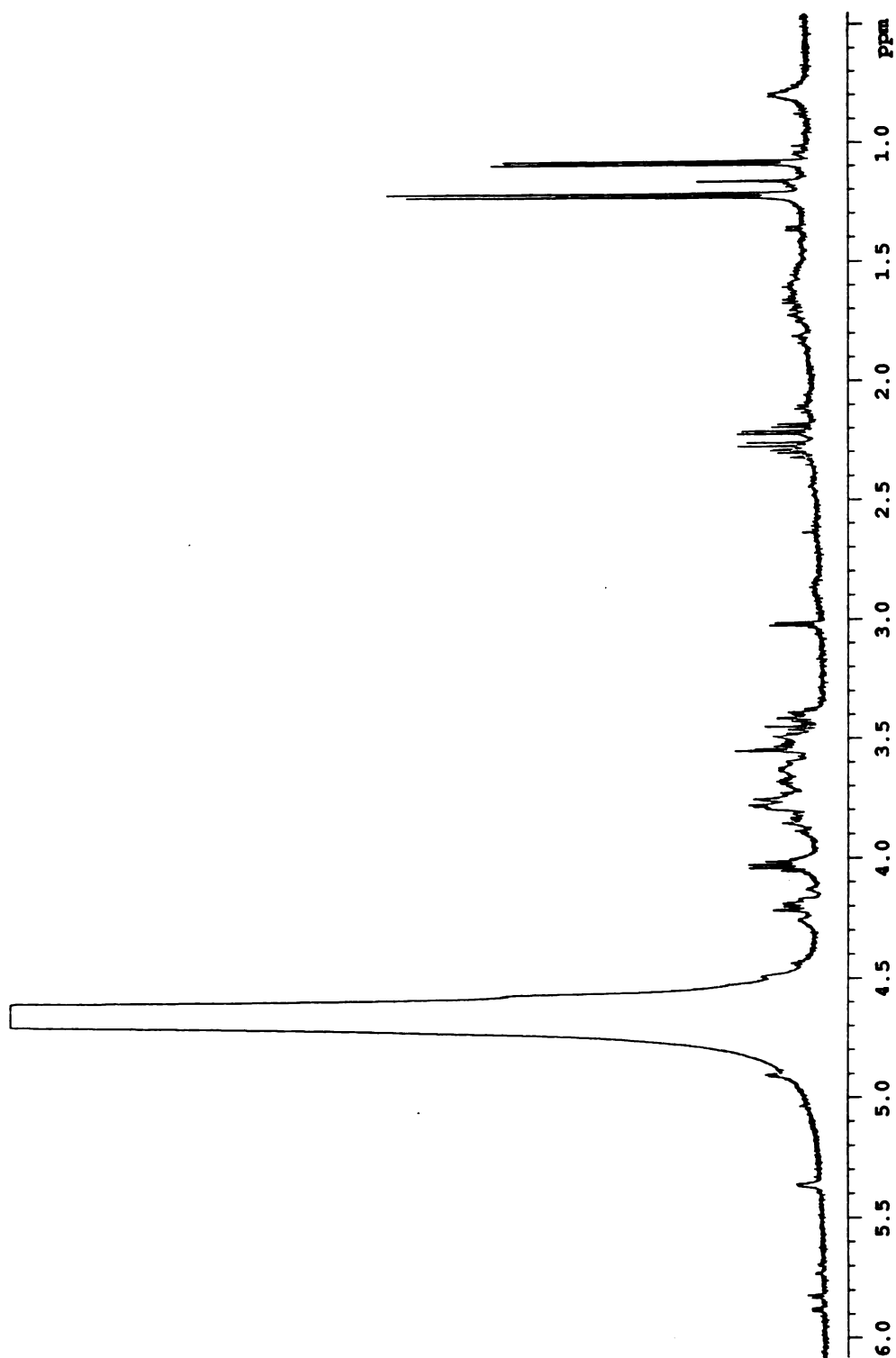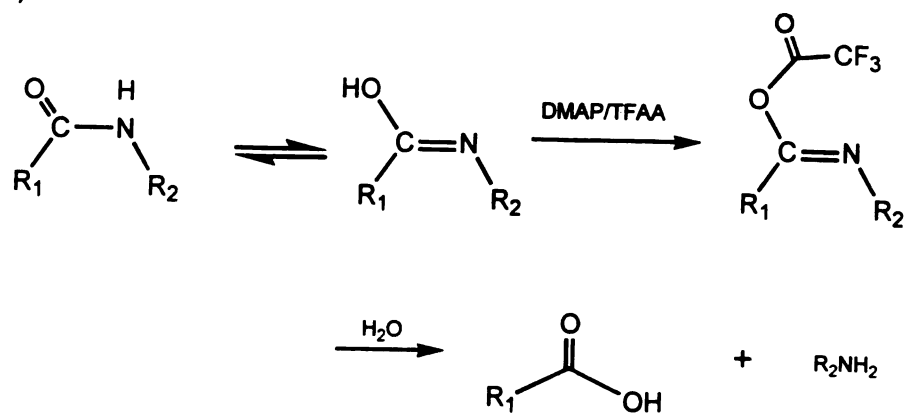
Figure 6

Figure 7. **^{1}H-NMR spectrum of the aqueous fraction of the totally delipidated lipid A after the removal of fatty acids by extraction.** The doublet at 1.22 ppm is due to a lactyl group. The doublet at 1.08 ppm and the pair of doublet of doublet between 2.16 and 2.35 ppm are due to the methyl and methylene protons of the 3-hydroxylbutyryl residue, respectively.

Figure 7

using 4-dimethylamino pyridine and trifluoroacetic anhydride (14). The de-N-acylation

can be rationalized mechanistically as shown in Figure 8A. The N-linked fatty acids were

released by this process and the hydroxyl groups that were β-to carbonyl functions were

eliminated yielding α,β-unsaturated esters. Using this procedure, 27-hydroxyoctacosanoic

acid was released from the lipid A molecule. This indicated that this fatty acid (which is

known not to be N-linked) was attached to some other site that was prone to

cleavage. This is true in the case of the 3-position of galacturonic acid. An acetoxy group

here is susceptible to acid catalyzed elimination to form an extended conjugated system

after the 4,5-unsaturation (which requires that the 4-hydroxyl group be unsubstituted) is

effected (Figure 8B). The free 27-hydroxyoctacosanoic acid was recovered by reverse

phase chromatography and its purity established by GC (Figure 9). These results,

therefore, established that the 27-hydroxyoctacosanoyl residue was attached to the 3-

position of the galacturonic acid residue. A ^1H NMR spectroscopy analysis (Figure 10)

of the liberated 27-hydroxyoctacosanoic acid fraction indicated that it was, in fact,

acylated with a 3-hydroxybutyryl residue thus explaining the 3-hydroxybutyryl signals in

the NMR spectrum of the aqueous fraction from the total deacylation. A pair of mutually

coupled doublet of doublets at 2.5-2.8 ppm were assigned to the A and B components of

an ABX spin system. This splitting pattern is characteristic of the signals for the

methylene protons of a 3-hydroxybutyryl substituent (21). The A component appeared at

2.61 ppm (J = 5.1, 8.1 Hz), and the B signal at 2.70 ppm (J = 8.1, 5.1 Hz). The

chemical shifts were consistent with a CH_2 being adjacent to a carbonyl group. The X-

component (the methine proton) was assigned to the downfield sextet at 5.48 ppm. The

Figure 8. (A) N-deacylation reaction with DMAP / TFAA. (B) Mechanism of elimination of the 3-substituent of GalA with DMAP / TFAA to form an extended conjugated system.

(A)

(B)
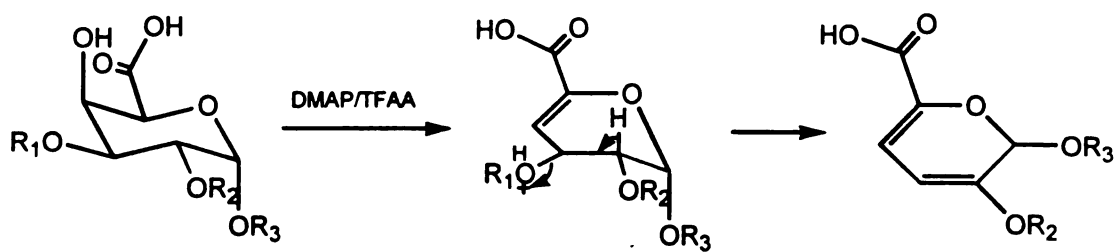
Figure 8
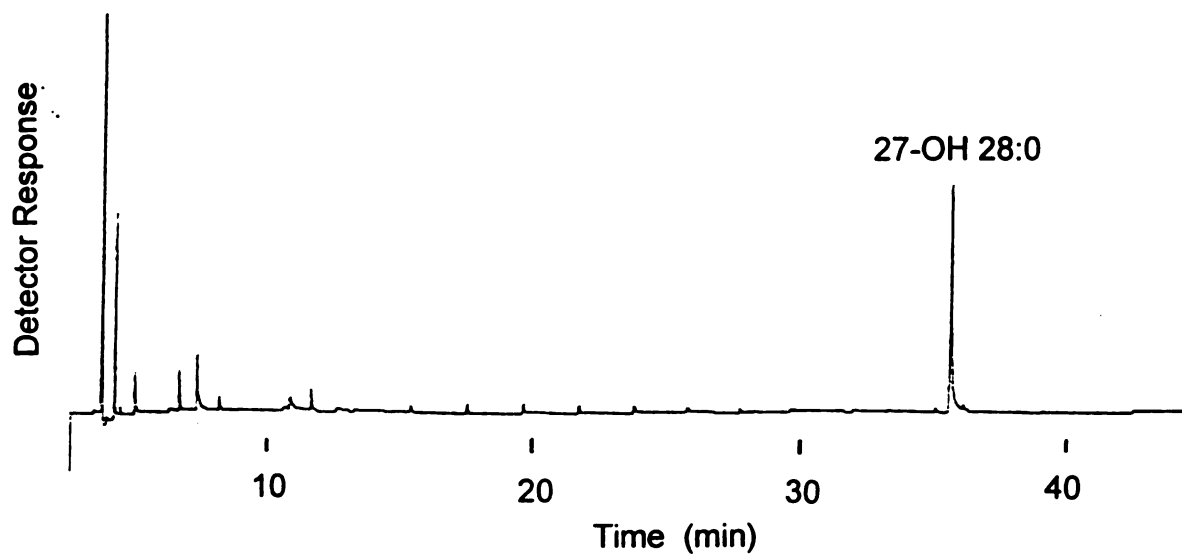
Figure 9. Gas chromatogram of the methanolysed fraction containing 27-OH 28:0 liberated by trifluoroacetic anhydride / dimethylamino pyridine. The late eluting peak is from the methyl ester of 27-OH 28:0. The earlier one is probably from 3-hydroxylbutyrate.

Figure 10. **¹H NMR spectrum of the fraction containing 27-OH 28:0 liberated by trifluoroacetic anhydride / dimethylamino pyridine.** As shown in the difference decoupling spectra in the inset, irradiation of the sextet at 4.91 ppm causes the collapse of a 3H-doublet at 1.18 ppm due to the methyl group of the 27-OH 28:0 residue. The downfield chemical shift of the methine signal (4.91 ppm) indicates that the hydroxyl group is acylated. Irradiation of the sextet at 5.48 ppm causes the collapse of the pair of mutually coupled doublet of doublet between 2.5 and 2.8 ppm and the doublet at 1.40 ppm assignable to a 3-hydroxybutyryl residue.
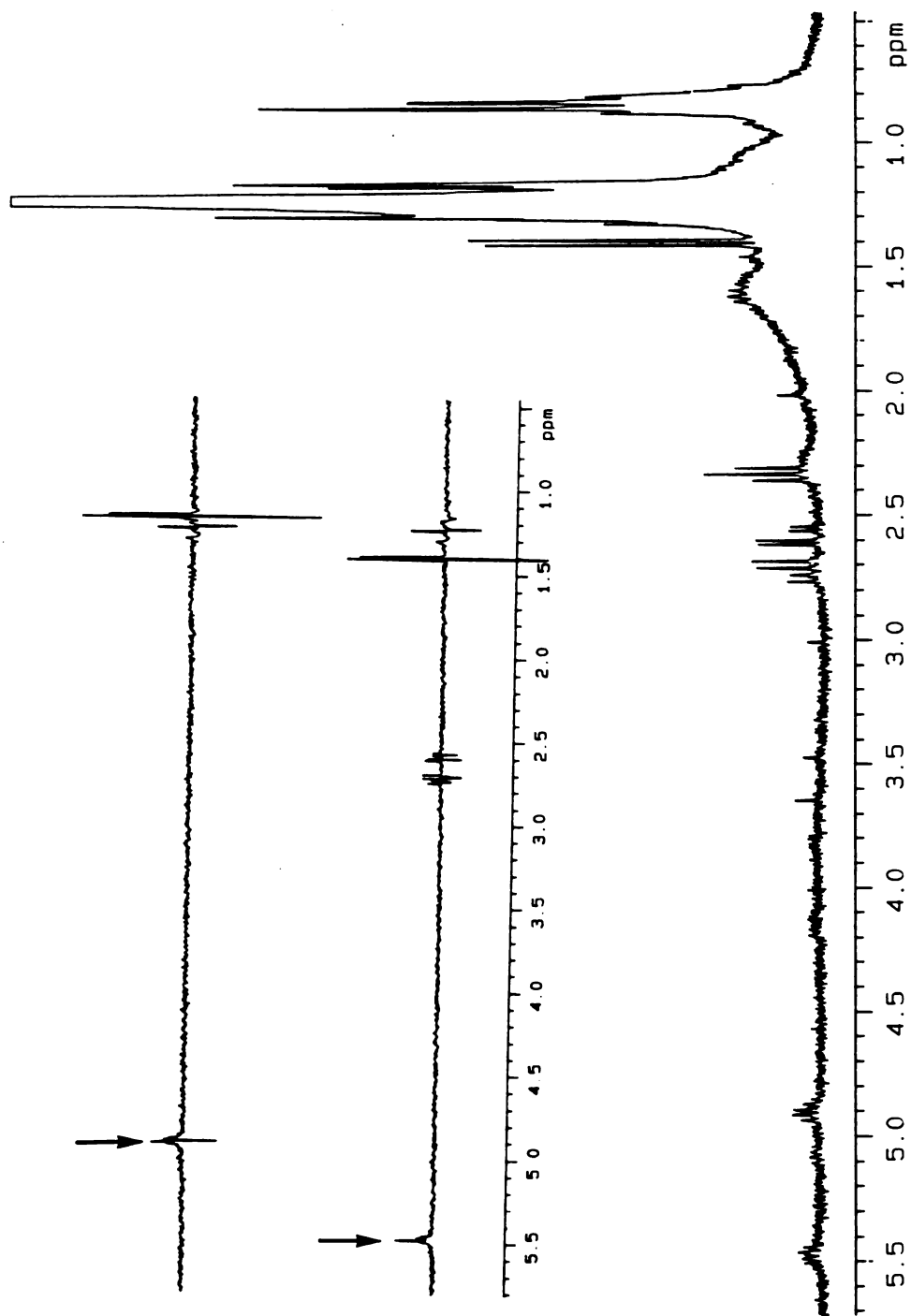
Figure 10

downfield shift of the 3-hydroxylbutyryl methine proton indicated that the hydroxyl group was trifluoroactylated from the reaction. Irradiation of this sextet led to collapse of the 3H-doublet at 1.40 ppm, and the methylene doublet of doublets at 2.5-2.8 ppm (see inset to Figure 10). Another sextet was also observed in the spectrum at 4.91 ppm. Decoupling of this signal led to the collapse of a doublet at 1.18 ppm. This latter resonance was assigned to a methyl group split by a single proton. The sextet at 4.91 ppm was assigned to the methine proton at the 27-position of the 27-hydroxyoctacosanoyl residue and the doublet to the methyl group. The downfield position of the methine proton was consistent with the hydroxyl group being acylated (8). A triplet at 2.34 ppm was assigned to the methylene protons adjacent to the carbonyl group of the C-28 fatty acid. FAB mass spectrum (Figure 11) confirmed these results by the presence of a pseudo-molecular ion [M-H]⁻ at m/z 526 expected for a 3-hydroxybutyryl group esterified to a 27-hydroxyoctacosanoic acid residue. The trifluoroacetyl group is labile in hydroxylic solvents. Both NMR and GC-MS analyses of the products formed by the dehydration / de-N-acylation reaction clearly indicated that the location of the different 3-hydroxy fatty acids are not very site specific but that the distribution in chain lengths was the result of heterogeneity. Hence, some degree of dehydration of all of the various 3-hydroxy fatty acids was observed. One more important result from this experiment was that the hydroxyl group of 3-hydroxytetradecanoic acid was substituted (in an acyloxyacyl group) by another fatty acid and was, therefore, protected from acylation. Unsaturated fatty acids (Figure 12) were clearly defined in the ¹H-NMR spectra (Figure 13) by their characteristic downfield signals. A doublet of triplets at 5.80 ppm (J = 15.6, 1.8 Hz) and
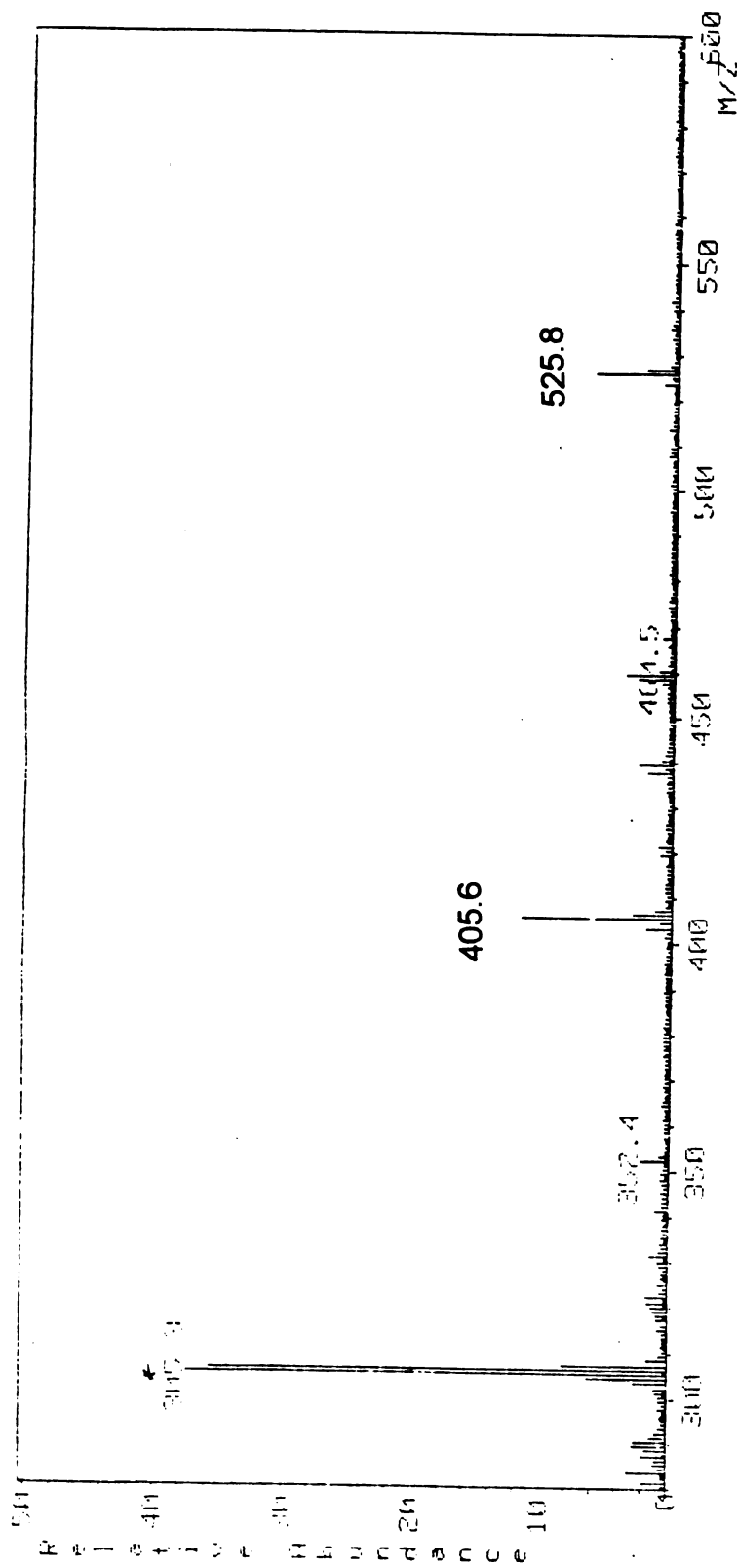
Figure 11. FAB mass spectrum of the 27-OH C28 containing fraction. Note that the [M-H]⁻ ion at 526 indicates that the 27-hydroxyoctacosanoic acid is esterified by a 3-hydroxybutyl group. The (*) signal is from the matrix.
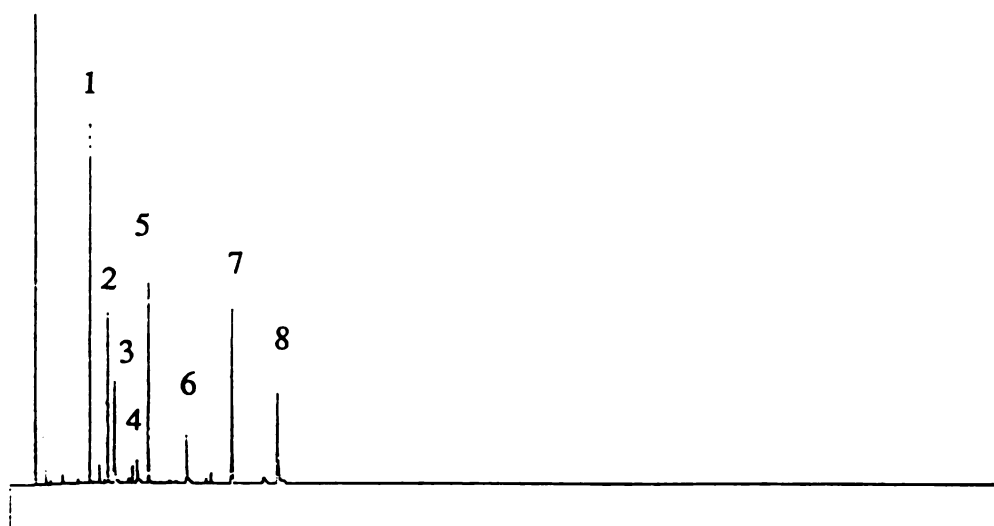
Figure 12. GC profile of the methyl esters of the unsaturated fatty acids containing fraction. Peaks 1-8 correspond to 14:1, 15:1, 3-OH 14:0, 3-OH 15:0, 16:1, 3-OH 16:0, 18:1, 3-OH 18:0, respectively.

HOOC-CH^a=CH^b-CH$_2$^{c}-(CH$_2$)$_n$^{d}CH$_3$^{e}

n=9, 10, 11, 13
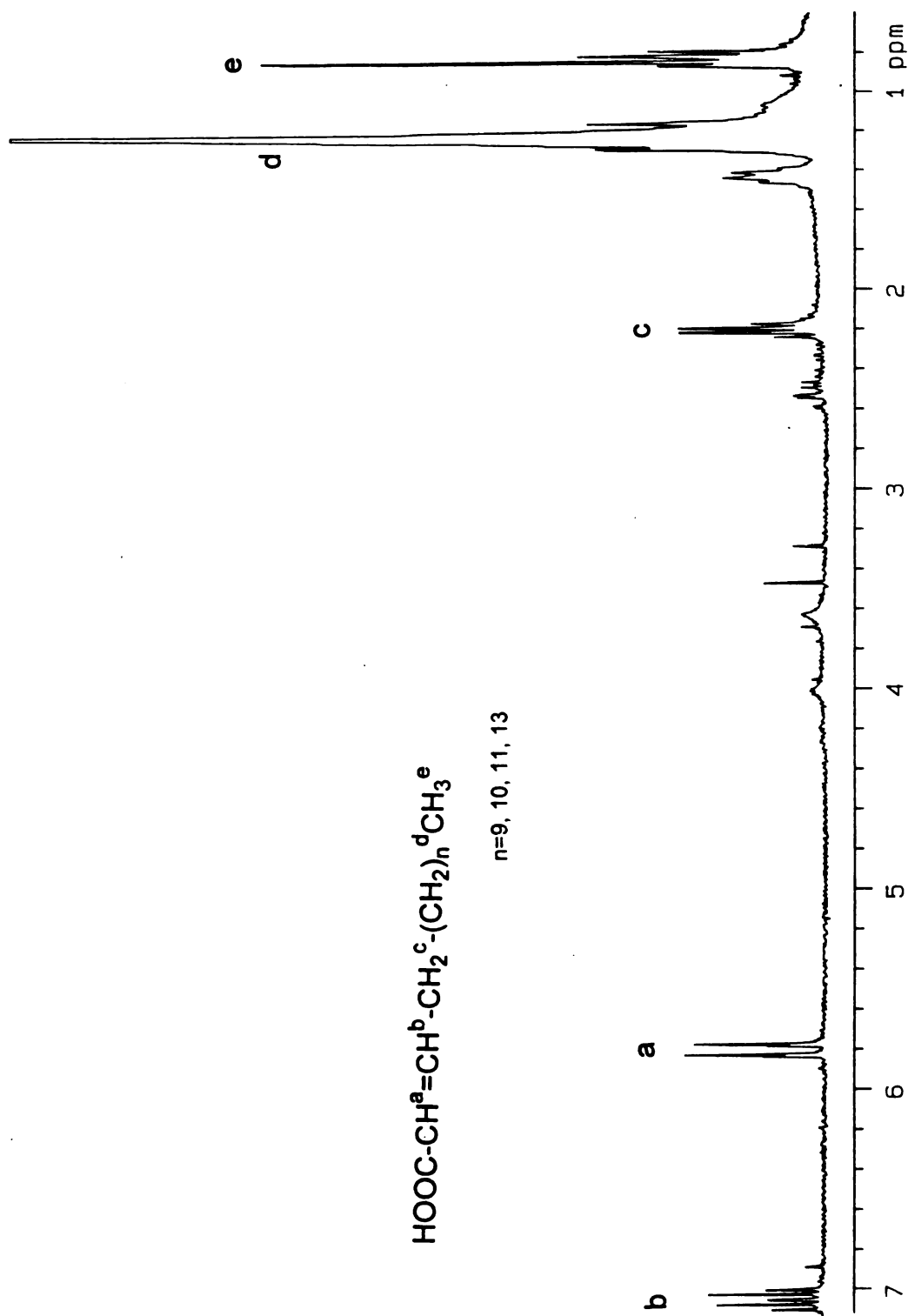
Figure 13. ^{1}H-NMR spectrum of the unsaturated fatty acids arising from the dehydration reaction.

another doublet of triplets at 7.05 ppm (J = 15.6, 7.8 Hz) were assigned to the α- and β-protons (respectively) of the unsaturated conjugated system. Signals at 2.2 ppm were assigned to the methylene group adjacent to the double bond. Proton signals at 0.8 and 1.2 ppm were assigned to the methyl and methylene protons, respectively, of the fatty acid chains.

A series of 1D and 2D NMR experiments were performed on the intact lipid A sample to confirm structural assignments and to obtain more information on the lipid A structure. The spectra were acquired in a perdeuterated solvent system containing pyridine, deuterium chloride solution, methanol and chloroform and were generally well resolved. The proton spectrum (Figure 14) contained several downfield signals that could be assigned to anomeric proton resonances. One of these signals at 5.10 ppm showed a correlation with a ^{13}C resonance (Figure 15) at 100 ppm in the proton / carbon HMQC spectrum (Figure 16) and was assigned to the α-anomeric proton of galacturonic acid. Another anomeric carbon cross peak was clearly evident in the HMQC spectrum at 101 ppm. This correlated with a signal at 4.70 ppm in the proton dimension that was obliterated by the left edge of the HOD line. Both proton and ^{13}C chemical shifts were consistent with an assignment as the anomeric proton of β-linked glucosamine. The ^{13}C / ^{1}H - HMQC spectrum (Figure 16) also provided a very convenient way of again definitively demonstrating the absence of a glucosamine aldonic acid residue. Secondary carbon atoms bearing N-acyl groups appear in the very limited range of 52 - 56 ppm. A sole signal at 52.6 ppm in the ^{13}C spectrum was assigned to C-2 of the glucosamine
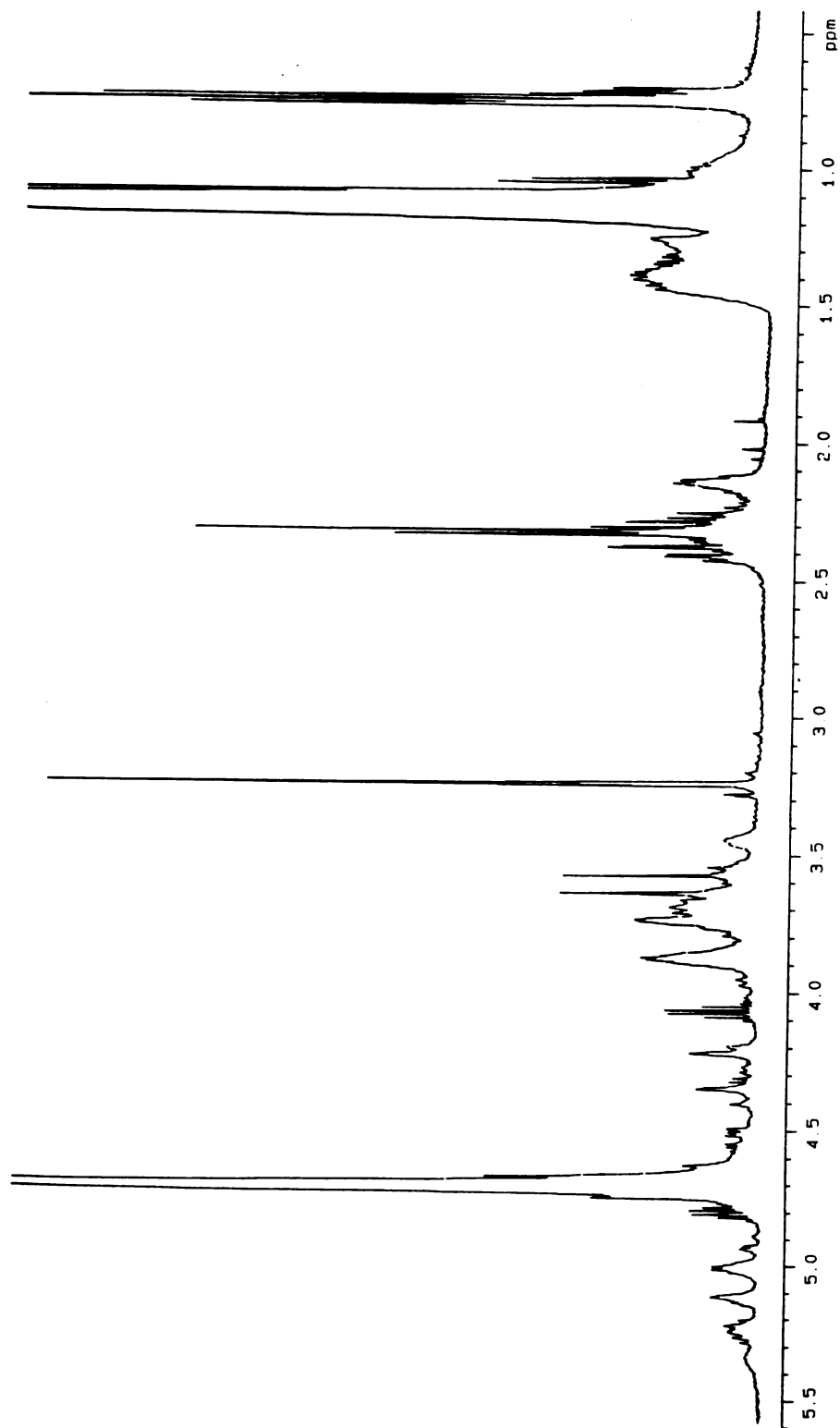
Figure 14. 500 MHz ^{1}H-NMR spectrum of the lipid A of *Rhizobium trifolii* ANU843.
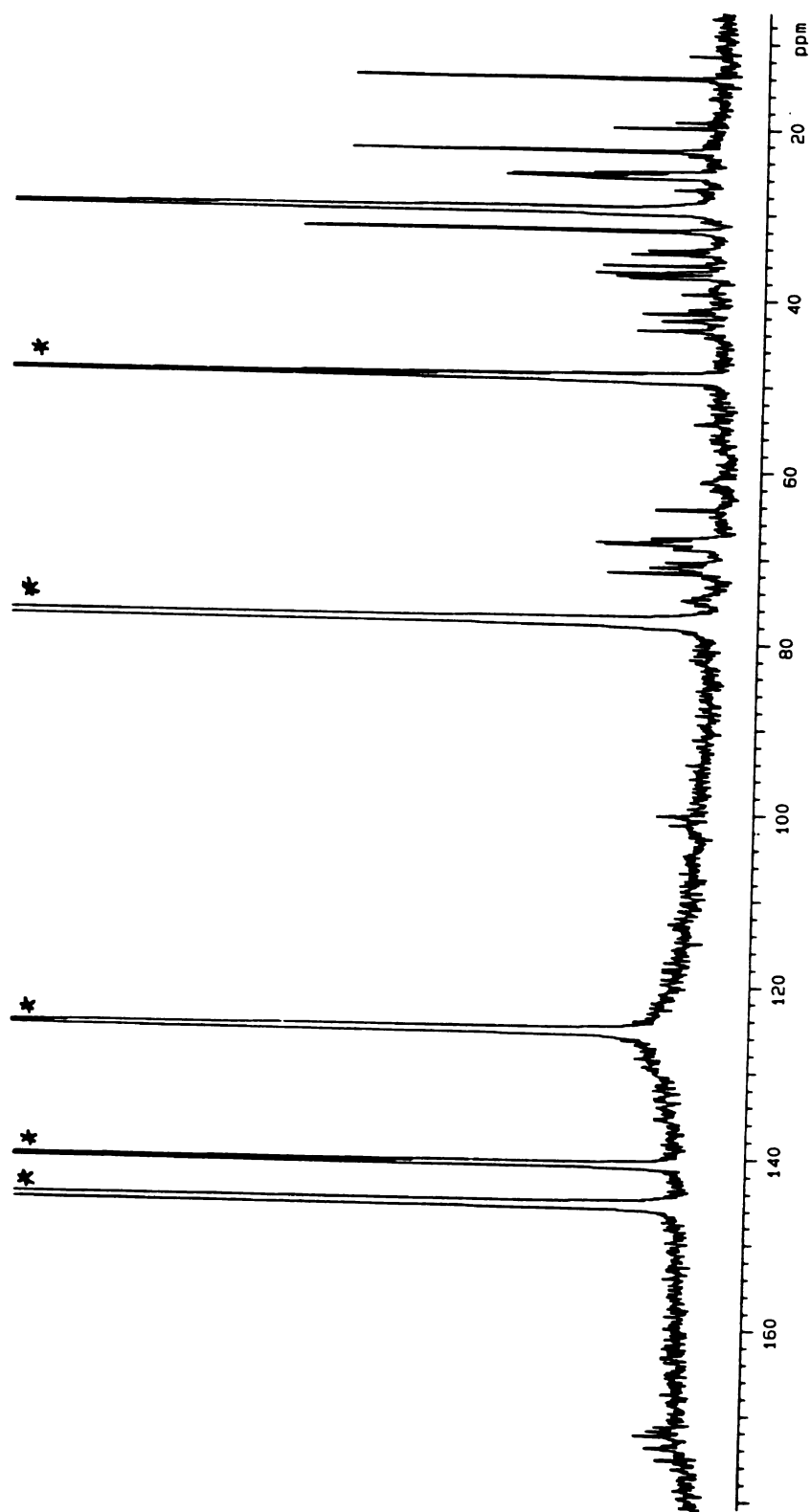
Figure 15. ^{13}C-NMR spectrum of the lipid A. Note that there are two anomeric signals at 100 and 101 ppm respectively. The (*) signals are from the solvent.

Figure 16. ^{1}H/^{13}C HMQC spectrum of *R. trifolii* lipid A showing key assignments. 1-6 refers to the ring protons of the GlcN residue and 1'-6' refers to the GalA residue.
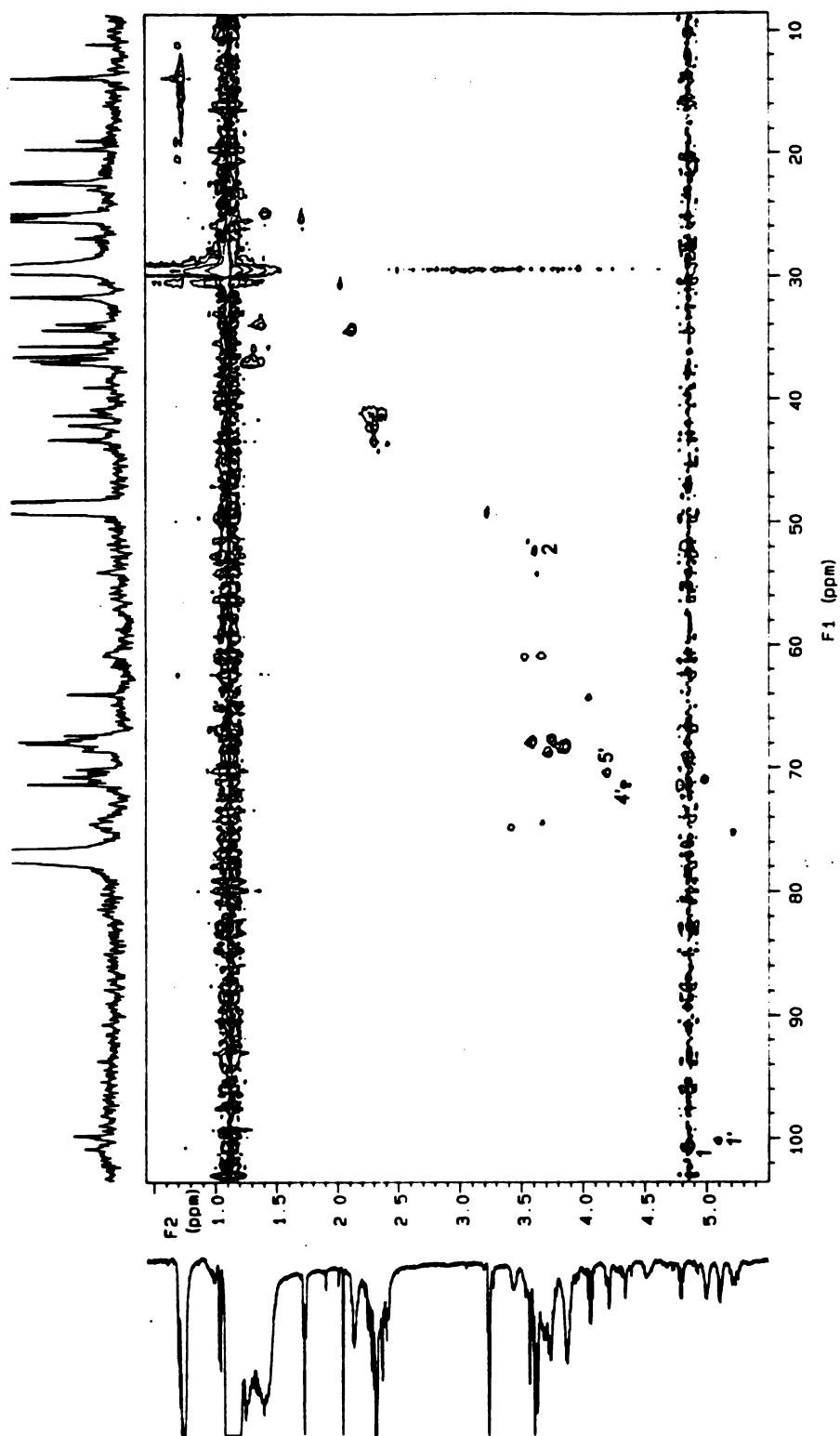
Figure 16

residue. This correlated with an ^{1}H signal at 3.66 ppm in the HMQC spectrum. There were no other signals between 50 and 60 ppm. The DQF-COSY spectrum (Figure 17) contained a cross peak showing a spin correlation between the proton signal at 3.66 ppm (H-2 of glucosamine) and a signal at 4.70 ppm (largely obliterated by the HOD line). This latter signal was assigned to H-1 of the glucosamine residue in accordance with the results from the HMQC study. The signal at 3.66 ppm (H-2 of glucosamine) in the DQF-COSY spectrum was also correlated with another downfield signal at 5.20 ppm which was assigned to H-3 of the glucosamine residue. The very downfield shift indicated that the site was acylated by a fatty acid. It was also possible to trace almost all of the connectivities for the H-1 to H-5 protons of the galacturonic acid residue in the DQF-COSY spectrum. The H-4 and H-5 protons appeared as a pair of mutually coupled signals with no detectible splitting in the 1-D spectra but with a clear correlation in the DQF-COSY spectrum. The H-4 signal appeared at 4.34 ppm and the H-5 signal appeared slightly more upfield at 4.22 ppm. These protons are unique and easiest to assign among the carbohydrate ring protons since they should display very small splitting (because of their *cis*-relationship with neighboring protons) and should appear as approximate singlets (actually narrow triplets). All other carbohydrate protons in the molecule (except for H-1 of galacturonic acid) has at least one *trans*-axial neighbor and should display at least one large ($\geq$ 7Hz) coupling. A correlation between H-4 and H-3 at 4.48 ppm was also very evident from the DQF-COSY spectrum. This signal, though visible, was partially obliterated by the water line. Its downfield position indicated that O-3 of the galacturonic acid residue bore an acyl group. This confirmed the selective elimination experiments we

Figure 17. Proton DQF-COSY NMR spectrum of *R. trifolii* lipid A showing key assignments. 1-6 refers to the ring protons of the GlcN residue and 1'-6' refers to the GalA residue. 3-OH FA stands for 3-hydroxyl fatty acyl residue. Unsub indicates that the fatty acyl residue bears a free 3-hydroxyl group. Sub indicates that the 3-hydroxyl position is esterified by another fatty acyl residue. The (*) signals denote impurities.
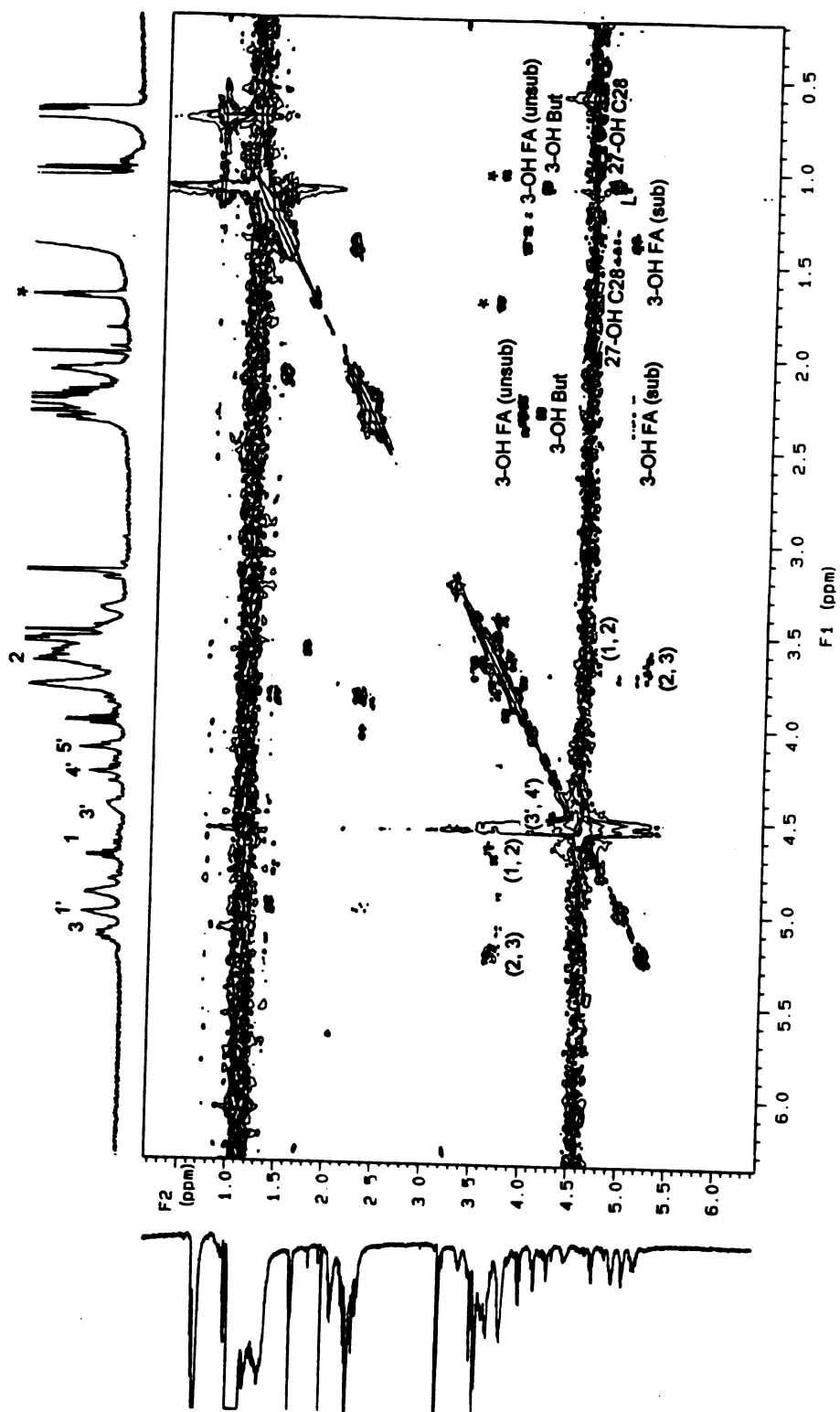
Figure 17

described earlier the results of which indicated that the 27-hydroxyoctacosanoyl group was attached to this position. It also disproved the conclusion of Bhat et al. that the galacturonic acid residue was not acylated. No connectivities were observed between H-3 and H-2 nor between H-1 and H-2 indicating that H-3 and H-2 both overlapped at the HOD line.

The DQF-COSY spectrum contained several other connectivities that were consistent with other structural details described earlier. Hence a quartet at 4.80 ppm showed a connectivity to a doublet at 1.08 ppm. These were easily assignable to the methine and methyl groups of a lactyl residue respectively. The connectivities for the 3-hydroxy fatty acids were also easily observable. There were two different sets of connectivities for 3-hydroxy fatty acids. One set corresponded to those in which the 3-hydroxy group was free, with the 3-methine proton resonance appearing at 3.82 ppm. Another set corresponding to acyloxyacyl residues in which the methine proton appeared significantly more downfield at 4.92 ppm was also observed. Assignments are given in Figure 17. The connectivities established above were also easily observable by proton TOCSY analysis (Figure 18).

One other major issue that remained to be settled was the position of linkage between the galacturonic acid residue and the glucosamine residue. This was very easily established by a ^{13}C - DEPT NMR spectroscopy experiment. This experiment also confirmed our earlier deduction that there is no glucosamine aldonic acid residue in the lipid A of this lipopolysaccharide. The subspectrum for primary carbons only is shown

Figure 18. TOCSY spectrum of *R. trifolii* lipid A.
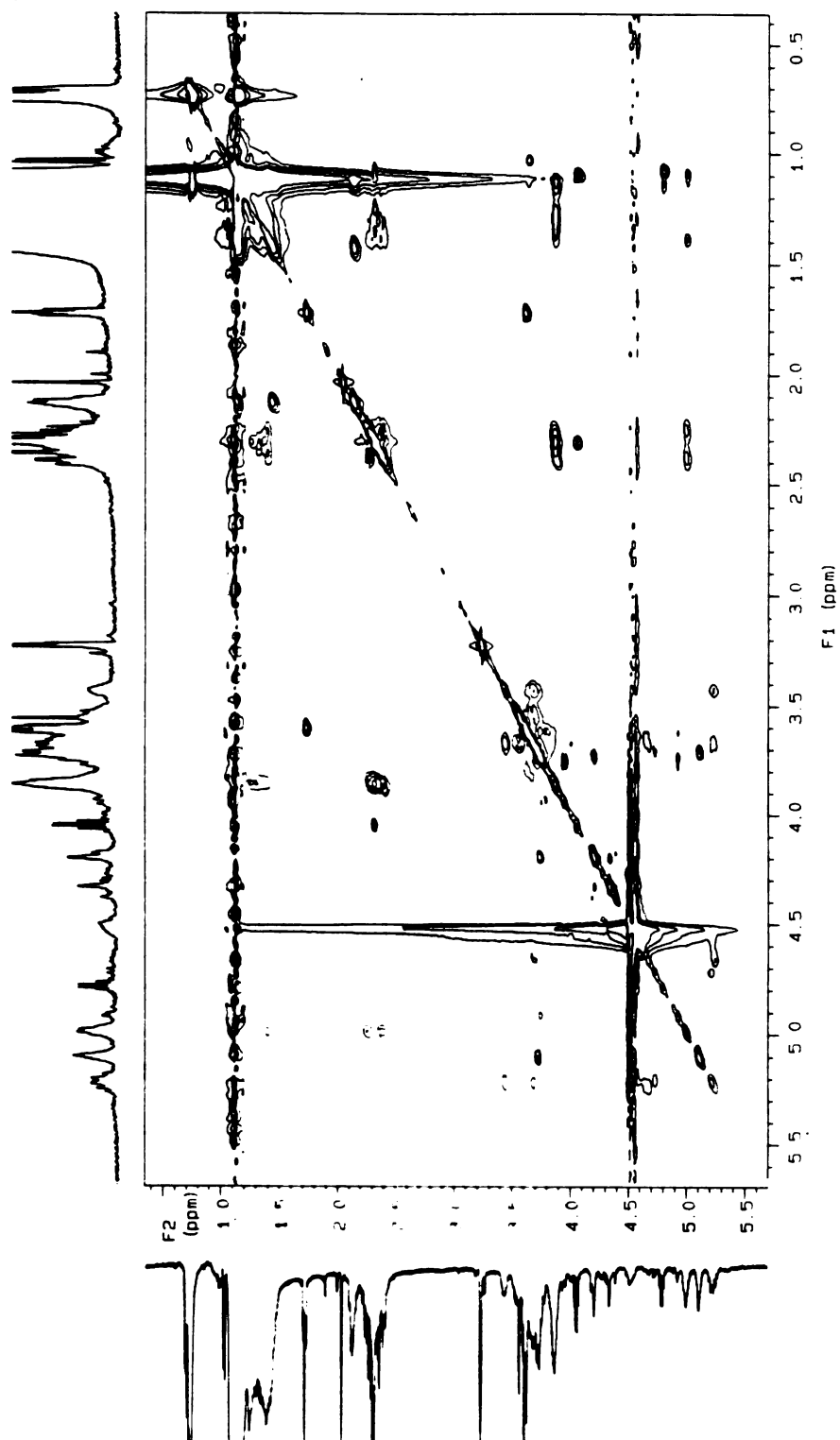
in Figure 19. Only one peak was observed and this had a chemical shift of 68 ppm. This result provided clear unambiguous proof that only one primary carbon (from C-6 of glucosamine) was present and that this position was substituted since it appeared downfield of the usual position for a primary alcohol ($\sim$ 61 - 62 ppm).

The revised structure was further confirmed by electrospray mass spectrometric analyses on the intact lipid A molecule (Figure 20). Fragment ions at m/z 487, 525, 713 and 921 were detected. These corresponded to the substructures shown in the figure. An ion at m/z 1961 $\pm$ 4 was assigned to the [M - H]⁻ pseudomolecular ion of the monosodium salt of a lipid A molecule containing 3-OH-C14, C15, C16, C18, C4 and 27-OH-C28 fatty acyl components. This molecule has a nominal molecular weight of 1959. Another pseudomolecular ion [M-H]⁻ was detected at m/z 1908. This was 28 (two $-CH_2$ units) less than the non-sodiated form of the first and could be attributed to heterogeneity. A cluster of peaks between m/z 3580 and 4300 (many of them separated by 56 mass units) corresponding to dimer species was also observed in the spectra.

The structure proposed here for the lipid A region of the LPS of *Rhizobium leguminosarum* biovars is radically different to the one proposed by Bhat et al. That prior study was seriously limited by the unavailability of NMR data. This is probably due to the intractability of these molecules to NMR analysis because of extensive aggregate formation in chloroform or dimethylsulfoxide solution as well as their insolubility in methanol. A problem we have surmounted here with the composite solvent system we describe. The inconsistencies in the linkage (1,4 vs 1,6) between galacturonic acid
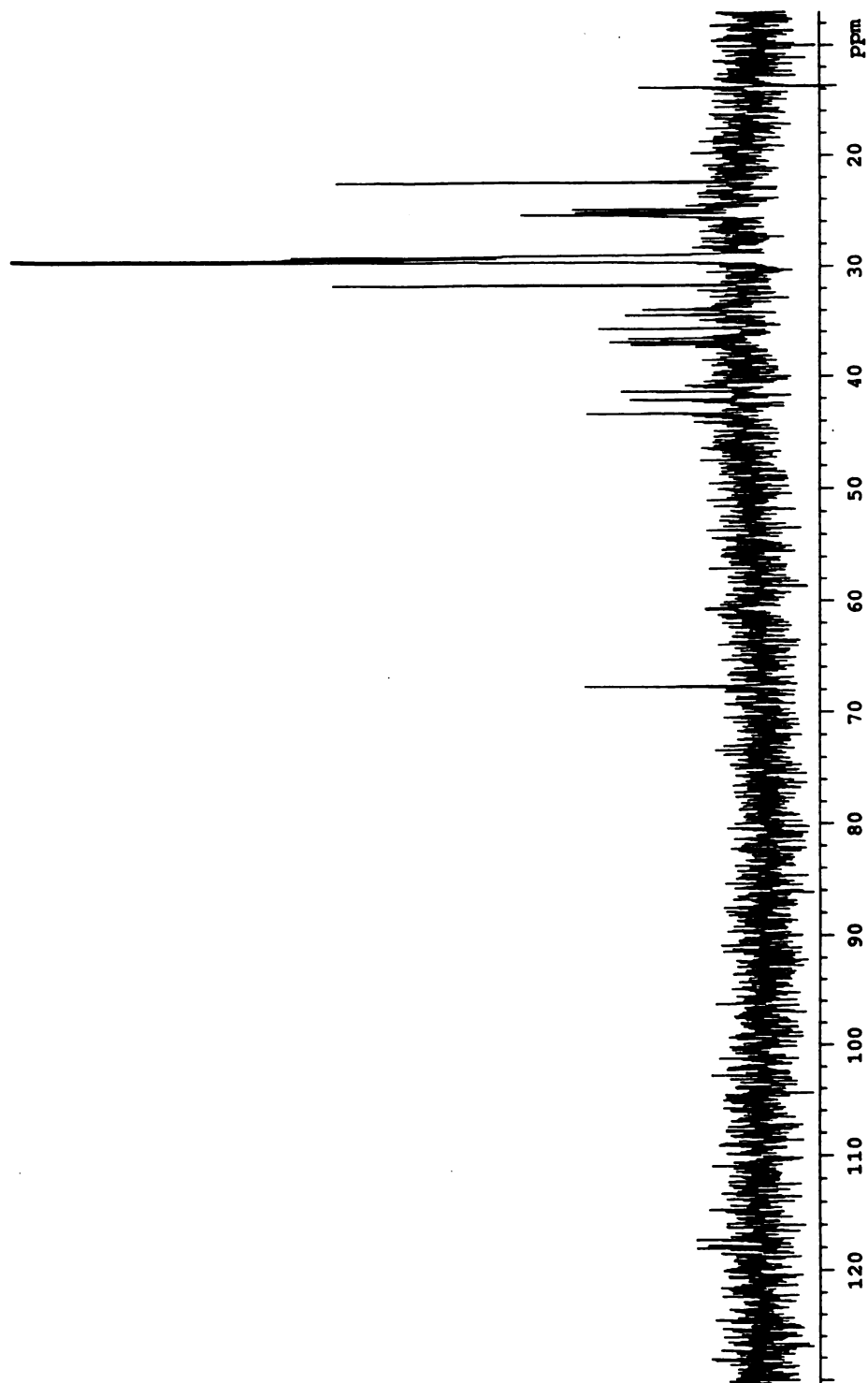
Figure 19. ^{13}C DEPT NMR spectrum of *R. trifolii* lipid A showing -CH$_2$ carbon signals only. Note the signal at 68 ppm is from C-6 of the GlcN residue, the downfield chemical shift indicates that it is the linkage position.

Figure 20. Negative ion ES-MS spectrum of the lipid A. The ion at m/z 525 can be assigned to carboxylate ions of 3-hydroxltutanoic acid esterified to 27-hydroxyoctacosanoic acid or to combinations such as C14+C18 or C16+C16. The ion at 1961±4 can be assigned to a sodiated molecular ion containing 3-OH c14, C15, C16, C18, C4 and 27-OH C28 fatty acyl components. The ion at 1907±2 can be assigned to a non-sodiated lipid A molecule with fatty acids two methylene units less. The ions between 3580 and 4300 correspond to dimer species.
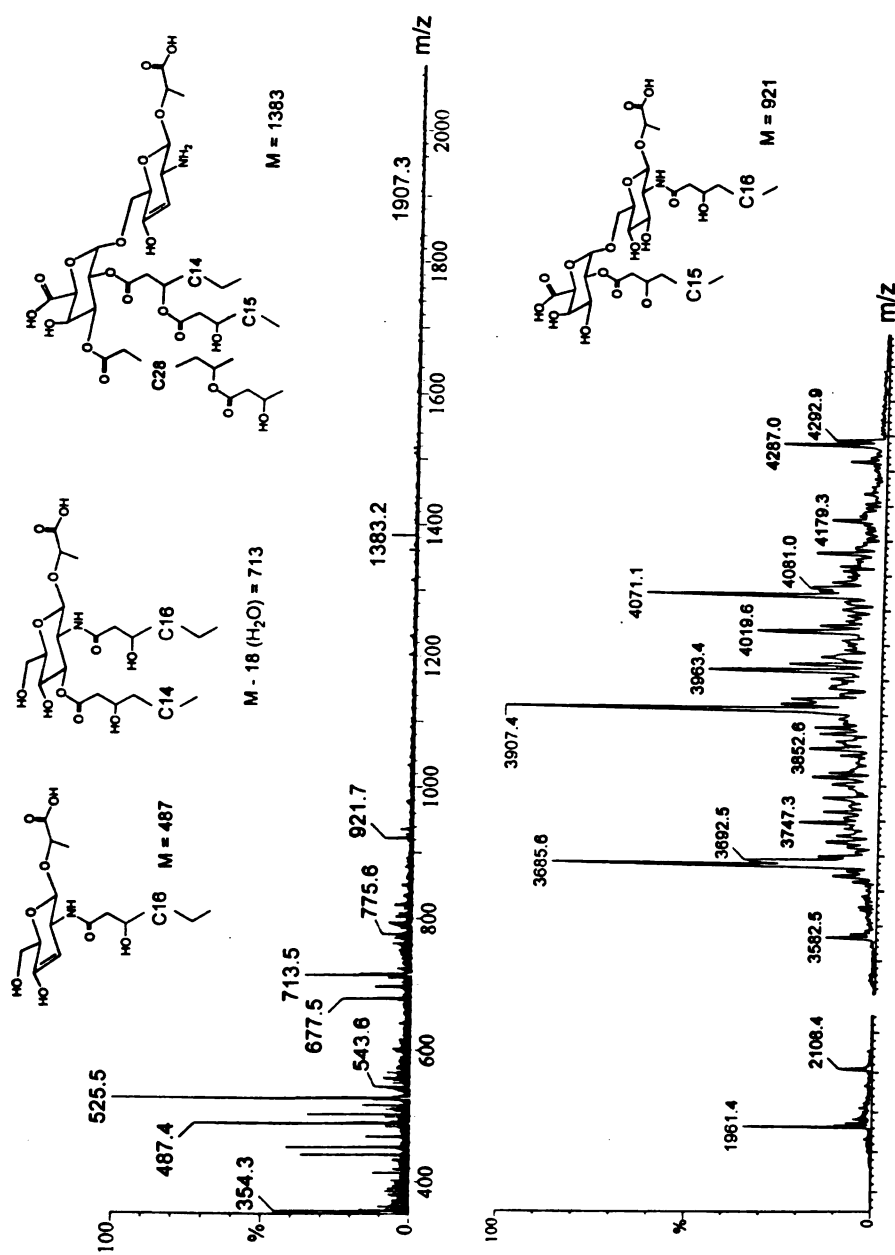
Figure 20

and glucosamine reported by the two studies may be due to the fact that β-elimination involving the galacturonic acid residue does not result in its loss from the lipid A structure as stated by Bhat et al. This process simply results in elimination of the 4-substituent of the galacturonic acid residue to form an α,β-unsaturated ester or acid that is quite stable. It still remains attached to the rest of the molecule as an α-L-*threo*-hex-4-enopyranosyl uronic acid residue. Such a residue is quite stable to acid hydrolysis but can be removed quite readily by ozonolysis (16). The galacturonic acid residue is not "eliminated" from the residue that it is attached to. Treatment of lipid A with the strong base required for the β-elimination will, however, lead to extensive deacylation. This would lead to easier alkylation by ethyl iodide of hydroxyl groups (such as the one at C-4 of glucosamine) that are located closer to the hydrophobic core. Another major conflict between these studies is that we definitively show that there is no glucosamine aldonic acid residue in the lipid A of this lipopolysaccharide. This is demonstrated here by mass spectrometry of a totally de-O-acylated product, and by very definitive NMR experiments that show the presence of only one primary carbon bearing an oxygen and only one carbon bearing a nitrogen. Several other lines of proof are also given. The study of Bhat et al. is compromised because, by the authors' own admission, they could never find a partially methylated glucosamine aldonic acid residue in their analyses. They explain this by saying that probably such a residue would be labile. In addition, the study seems to contain a flaw in the assignment of the structure of the product formed by total hydrazinolysis of the lipid A molecule. In this structure (based on a mass spectrum that is not shown) there is a glucosamine aldonic acid residue that is present in a lactone form. Lactones are extremely

reactive to nucleophiles such as hydrazine (even more so than acyclic esters) and that the glucosamine aldonic acid lactone residue cannot survive treatment harsh enough to remove all O-linked and N-linked fatty acids. One of the very frustrating aspects of structural analysis on rhizobial lipopolysaccharides using traditional chemical methods is their resistance to deacylation and methylation. Hence methods that are promoted by mild catalysts (such as silica catalyzed methylation with diazomethane) have yielded variable and very inconclusive results. Undermethylation is also a very common problem and we have had great difficulty obtaining reproducible results even from the same material. Aggregation has also made mass spectrometric studies extremely difficult. This might explain the absence of mass spectra for the intact lipid A in the work of Bhat et al. These problems have also been a major hindrance in our studies. Perhaps there are differences between the structures of the lipid A of *Rhizobium leguminosarum* biovars that would explain the other inconsistencies between these two studies. For the moment, this seems to be precluded by Bhat et al. who claim that theirs is a general structure.

REFERENCES

1.   Maier, R.; Brill, W. *J. Bacteriol.* **1978**, 133, 1295-1299.

2.   Noel, K.; Vandenbosch, K.; Kulpaca, B. *J. Bacteriol.* **1986**, 168, 1392-1401.

3.   Priefer, U. *J. Bacteriol.* **1989**, 171, 6161-6168.

4.   Hollingsworth, R. I.; Carlson, R. W.; Dazzo, F.B. *Carbohydr. Res.* **1988**, 176, 127-135.

5.   Hollingsworth, R. I.; Carlson, R. W.; Garcia, F.; Gage, D. *J. Biol. Chem.* **1989**, 264, 9294-9299.

6.   Hollingsworth, R. I.; Carlson, R. W.; Garcia, F.; Gage, D. *J. Biol. Chem.* **1989**, 265, 12752.

7.   Russa, R.; Luderitz, O.; Rietschel, E. Th. *Arch. Microbiol.* **1985**, 141, 284-289.

8.   Hollingsworth, R. I.; Carlson, R. W. *J. Biol. Chem.* **1989**, 264, 9300-9303.

9.   Hollingsworth, R.I.; Lill-Elghanian, D. In *Cellular and Molecular Aspects of Endotoxin Reactions*; Nowotny, A., Spitzer, J. J., Ziegler, E. J., Eds.; Elsevier Science Publishers. 1990; pp73-84,

10.   Hollingsworth, R. I.; Lill-Elghanian, D. *J. Biol. Chem.* **1989**, 264, 14039-14042.

11.   Urbanik-Sypniewska, T.; Seydel, U.; Greck, M.; Wechesser, J.; Mayer, H. *Arch. Microbiol.* **1989**, 152, 527-532.

12.   Bhat, U. R.; Forsberg, L. S.; Carlson, R. W. *Biochem.* **1994**, 269, 14402-14410.

13.   Wang, Y.; Hollingsworth, R. I. *Anal. Biochem.* **1995**, in press.

14.   Shyong, B. MS thesis, Michigan State University, East Lansing, Michigan, 1992.

15.   Dubois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F. *Anal. Chem.* **1956**, 28, 350-356.

16.   Hollingsworth, R. I.; Abe, M.; Dazzo, F. B.; Hallenga, K. *Carbohdr. Res.* **1984**, 134, C7-C11.

17.   Ernst, R. R.; Bodenhansen, G.; Wokaun, A. In *Principles of Nuclear Magnetic Resonance in One and Two Dimensions;* Oxford University Press: Oxford, 1987; pp431-440.

18.   Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, 65, 355-360.

19.   Hollingsworth, R. I.; Dazzo, F. B. *J. Microbiol. Methods* **1988**, 7, 295-302.

20.   Bhat, U. R.; Mayer, H.; Yokota, A.; Hollingsworth, R. I.; Carlson, R. W. *J.*

Bacteriol. **1991,** 173. 2155-2159.

21.   Hollingsworth, R. I.; Abe, M.; Sherwood, J. E.; Dazzo, F. B. *J. Bacteriol.* **1984,** 160, 510-516.

CHAPTER 5

THE MOLECULAR BASIS FOR THE SUPRAMOLECULAR STRUCTURE OF

LIPOPOLYSACCHARIDES

ABSTRACT

Lipopolysaccharides from gram negative bacteria interact with the mammalian immune system to trigger a cascade of physiological events leading to a shock syndrome which results in death in over 70% of cases. It is known that the supramolecular structure of lipopolysaccharide aggregates are critical contributors to its biological activities. Despite this, the molecular basis for the formation of the regular hexagonal plates and arrays observed in lipopolysaccharide films and suspensions is unknown. Since these structures are two dimensional, it is unlikely that X-ray crystallographic methods will shed much light on their detailed structure. Knowing this structure is important since it is becoming increasingly likely that insertion of the lipopolysaccharide hydrocarbon chains into the target host cell membrane may be involved in triggering host responses. This work describes the 3-dimensional structure of the lipopolysaccharide lipid A moiety. The structure was obtained by a combination of molecular mechanics calculations and nuclear magnetic resonance spectroscopy taking into account information from X-ray powder diffraction and electron microscopy studies.

INTRODUCTION

Lipopolysaccharides (LPSs) are complex lipid-linked carbohydrates which are found in the outer membranes of gram-negative bacteria. These molecules are composed of a polymeric carbohydrate region called the O-antigen, a short oligosaccharide region called the R-core and a fatty-acylated region (usually a glucosamine disaccharide) called the lipid A (1). The lipid A molecule is usually bisphosphorylated and, in the case of *E. coli* and most bacteria, has the structure shown in Figure 1.

Lipopolysaccharides are potent potentiators of host immune responses in mammalian systems. These responses include the production of interleukins, tumor necrosis factor and prostaglandins at the cellular level (2-4) and fever, shock and death at the level of the organism (5, 6). Despite evidence for an interaction between lipopolysaccharides and host immune cells via receptor-mediated pathways involving the carbohydrate headgroup (7, 8), it is becoming increasingly clear that the supramolecular structure and the fatty acid arrangement of lipopolysaccharides are critical. Several workers have demonstrated a dramatic change in host-cell membrane properties when treated with lipopolysaccharides by a variety of techniques including fluorescence depolarization and electron spin resonance (9-11). It has also been demonstrated that when lipopolysaccharides bind to macrophages, they do so by membrane insertion with the polar, carbohydrate regions facing away from the cell (12). Lipopolysaccharides which lack even one fatty acid group are not toxic (13). The biological activity of lipopolysaccharides is affected by their salt forms (14). These latter two facts indicate
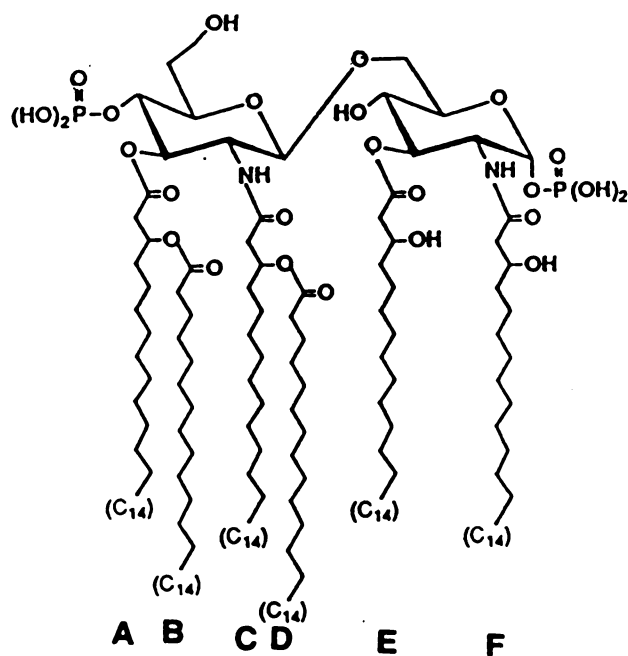
Figure 1. Structure of lipid A such as is found in *E. coli* strains. The disaccharide backbone is a β-1,6-linked D-glucosamine disaccharide. 3-Hydroxytetradecanoyl residues are attached to the 2, 3, 2′ and 3′ positions. The fatty acids at the latter two positions are esterified by C-12 and C-14 fatty acids, respectively.

that the supramolecular architecture (geometry and alkyl chain packing) of these molecules are critical determinants of their biological activity and that this activity might be mediated by some bulk colligative property of the lipopolysaccharide in which the fatty acid chain packing plays a critical role. It is, therefore, important to determine how this supramolecular structure may arise. The lipid A moiety is the anchor and structural basement of lipopolysaccharides. It also has the full spectrum of biological properties displayed by the intact LPS molecule. Any regularity in LPS structure must arise from lipid A organization.

MATERIALS AND METHODS

X-ray powder diffraction analyses were carried out on a RIGAKU diffractometer using a copper source. The K_β line was filtered out. A film of *E. coli* Re lipopolysaccharide was formed by spreading a 2% solution of lipopolysaccharide in 0.1 M calcium chloride or magnesium chloride onto carbon-coated grids. The calculations were performed on a Silicon Graphics 4D310 computer using MM2 (15) or DREIDING (16) force fields implemented in the BIOGRAF (17) molecular mechanics program. The full potential energy function was used. Interactions were summed over a 12 angstrom cut off. Minimizations were performed using the conjugate gradient method. Simulated annealed dynamics were used to avoid local minima.

The NMR sample of lipid A was prepared as 3 mg/ml in a mixture solvent (18) with

1:1:2:10 volume ratio of pyridine-D_5 / 37% deuterium chloride in deuterium oxide / methanol-D_4 / chloroform-D, respectively. All NMR spectra were acquired on a VARIAN VXR500 spectrometer operating at 500 MHz for protons and 202 MHz for ^{31}P. The NMR experiments were conducted using chloroform as the lock signal. The proton chemical shifts were reported using chloroform line at 7.24 ppm as the reference and the HOD signal was suppressed by presaturation. The ^{31}P chemical shifts were reported using 85% H_3PO_4 as an external reference. The DQF-COSY spectrum was recorded at phase sensitive mode using the VARIAN standard pulse sequence (19, 20). A total of 256 data sets with 16 transients of 2048 data points each with a relaxation delay of 2 s was used to resolve a spectral width of 5000 Hz in both the F2 and F1 dimensions. A Gaussian function was applied in both dimensions and zero-filling was applied in the F1 dimension to expand the data to a 2K×2K matrix. The ^{1}H/^{31}P HMQC (21) experiment was carried out using a data matrix of 256×2K with a spectral width of 4500 Hz in the F2 and 2000 Hz in the F1 dimension. A set of NOESY (22) experiments were performed using mixing times of 100, 200 and 300 ms.

RESULTS AND DISCUSSION

Several workers have demonstrated, using electron microscopy, that lipopolysaccharides form structures such as ribbons and sheets in solution (23, 24). Analysis of transmission electron micrographs of films prepared from such solutions containing calcium or magnesium ions shows that the area of electron density forms a highly ordered hexagonal

array mandating that the individual lipopolysaccharide molecules must form structures with a regular geometry. X-ray powder diffraction analyses of hydrated lipopolysacchrides (25) show maxima indicating a hexagonal array of lipid molecules. Our studies using X-ray powder diffraction also show that lyophilized lipopolysacchardies are microcrystalline and the relative intensities of the maxima indicate that the hydrocarbon chains are hexagonally packed. Recently, various lipopolysaccharides with lipid A structures similar to that shown in Figure 1 were crystallized and found to form perfectly hexagonal plates (26). However, it is still not known how this might be possible given the structure of the lipid A molecule (Figure 1).

Molecular mechanics calculations were performed in order to arrive at possible symmetrical structures for a bis-phosphorylated lipid A molecule bearing six fatty acid chains as shown in Figure 1. Specific structural requirements had to be met for any proposed structure. The six fatty acid chains must be arranged so as to form the most densely packed system since the hexagonal array is the most densely packed array. This system is one with a triangular cross section perpendicular to the direction of the extended alkyl chains. The sides of the triangle are all equal in length. This gave a maximum value of the van der Waals energy for the interaction of the extended alkyl chains. This arrangement is obligatory and only possible with this arrangement of the six fatty acid chains. It is also the only one which can give rise to a hexagonal arrangement of alkyl chains in the extended array, since an equilateral triangle is the fundamental regular geometric unit of a regular hexagon. The disaccharide head group conformation which allowed this fatty acid arrangement was one in
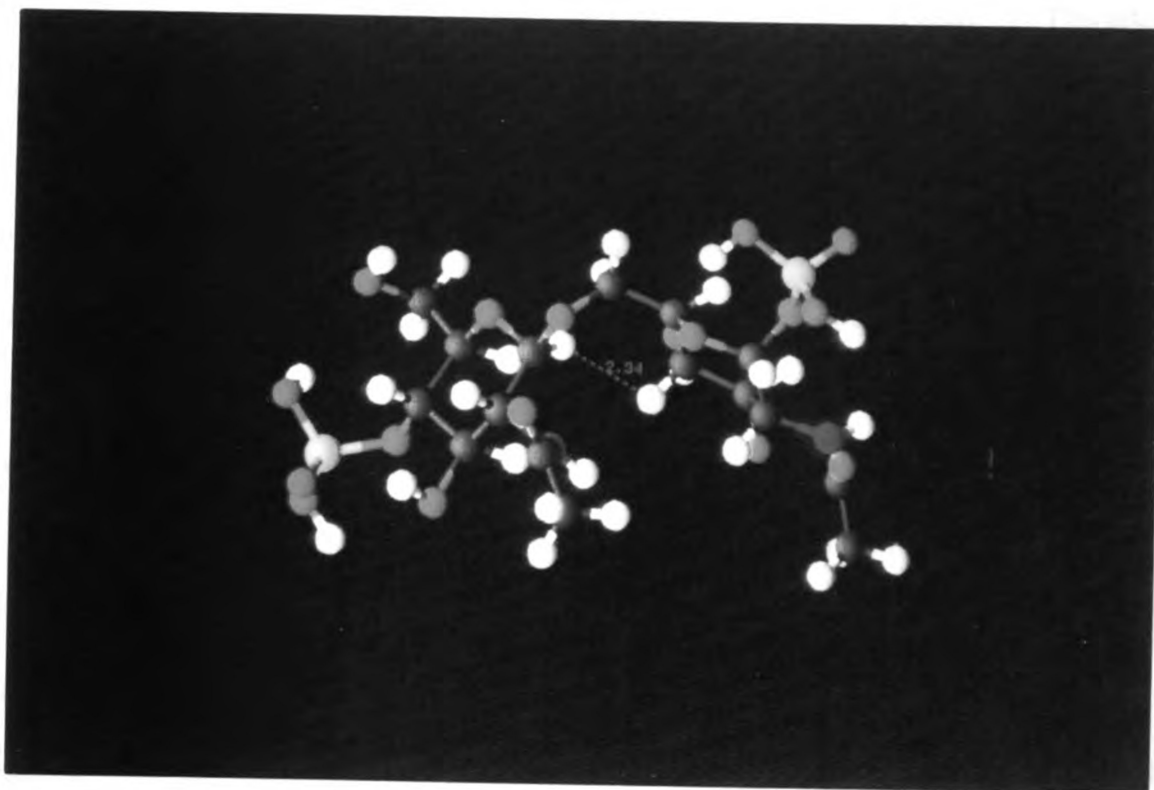
Figure 2. Computer model of the bis-phosphorylated glucosamine headgroup of the lipid A molecule as it is proposed to be present in the intact structure. The yellow atoms are phosphorus, dark blue are nitrogen, blue-green are carbon, red are oxygen and white are hydrogen. The N-linked fatty acids in lipid A have been replaced by N-acetyl groups. Note that the planes of the two glycosyl rings are orthogonal. The separation between H-1' (the anomeric proton of the distal glucosamine ring) and H-4 (the proximal ring) is indicated.

which the planes of the two carbohydrate rings were orthogonal (at right angles) and the

phosphate groups at a maximum separation but symmetrically displaced (Figure 2). This also

corresponded to a favorable torsional interaction term about the interglycosidic bond. There

were also some mandatory structural requirements which, once substantiated, would verify

the model. The first of these requirements was the distance between the anomeric proton of

the distal glucosamine ring and the C-4 proton of the proximal ring. This was 2.34 Å and

represented the only possible inter - ring proton - proton interaction which could give rise to

a nuclear Overhouser effect (n.O.e) in NMR spectroscopy. The second condition was the

dihedral angles about the bond between C-5 and C-6 of the proximal glucosamine ring. The

H-5 - C-5 - C-6 - H-6$_a$ angle obtained from the model was 86.17° and the H-5 - C-5 - C-6 -

H-6$_b$ angle was -29.16°. An analysis of the conformational space about the interglycosidic

bond also indicated that this orientation of the disaccharide rings is the only one which will

allow the fatty acid chains to be parallel and so attain the closest packed configuration. These

conditions, therefore, constituted a unique set of necessary and sufficient conditions which,

if demonstrated practically, would lead to the correct 3-dimensional structure of diphosphoryl

lipid A.

The distance and angular requirements of the proposed structure could easily be

verified by NMR spectroscopy. Lipopolysaccharides and lipid A molecules do not ordinarily

give well resolved NMR spectra. This was facilitated by the use of a new solvent system that

gives well resolved spectra for many different classes of lipid molecules. The required

interproton distance was established by the demonstration of a strong n.O.e. signal between

the two nuclei in question. A combination of homo and hetero-2-dimensional NMR experiments were performed in order to allow the assignments of the signals for the two protons. The signal for H-1 of the proximal glucosamine residue was readily assignable from the 500 MHz proton spectrum (Figure 3). It appeared at 5.34 ppm as a doublet of doublets (J= 7.5 and 3.0 Hz) split by both H-2 and the phosphorous atom in the phosphate group. This assignment could readily be confirmed by a 2-dimensional $^1H/^{31}P$ heteronuclear multiquantum coherence spectroscopy (HMQC) experiment which also indicated the identity of H-4 of the distal glucosamine residue which is also phosphorylated (see inset to Figure 3). The next obvious assignments were the signals for the two C-3 protons. These appeared in the proton spectrum in the region of 4.9 - 5.1 ppm. This relatively downfield location was to be expected since the carbon atoms to which they are attached also bear electron withdrawing acyloxy groups. Confirmation of these assignments and assignment of the other ring resonances as well as critical fatty acid resonances was obtained from 2-dimensional total correlated spectroscopy (TOCSY) (Figure 4) and double quantum filtered J-correlated spectroscopy (DQF-COSY) experiment (Figure 5). Once the assignments were made, it was then possible to demonstrate a substantial n.O.e between the anomeric proton of the distal glucosamine residue and H-4 of the proximal residue as required by the model (Figure 6).

The next step was to establish the magnitude of the dihedral angles in question. This could be established by determining the magnitudes of the vicinal coupling constant between H-5 and H-6$_a$ and the one between H-5 and H-6$_b$. These coupling constants could then be related to the dihedral angles between the bonds involving them by a Karplus-type

Figure 3. **500 MHz ^{1}H-NMR spectrum of lipid A in 1:1:2:5 pyridine-D_5 / 37% DCl in D$_2$O / CD$_3$OD / CD$_3$Cl.** The HOD line at 5.0 ppm was suppressed by presaturation. 1-6 refer to the proximal ring protons and 1'-6' refer to the distal ring protons. The insets show partial ^{1}H/^{31}P HMQC spectra. Note the correlations to the anomeric proton of the proximal ring residue (H-1) and to H-4'.

Figure 3

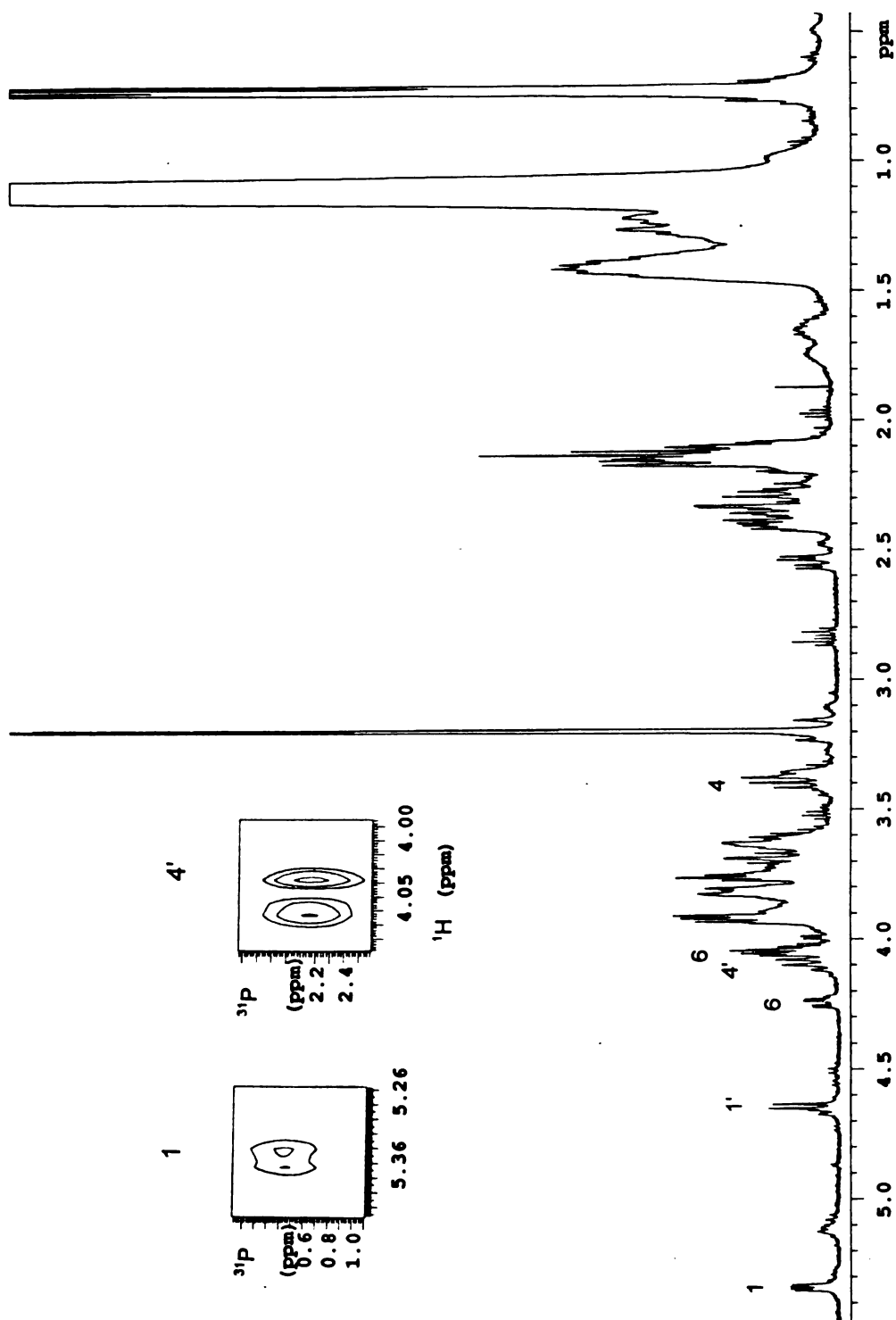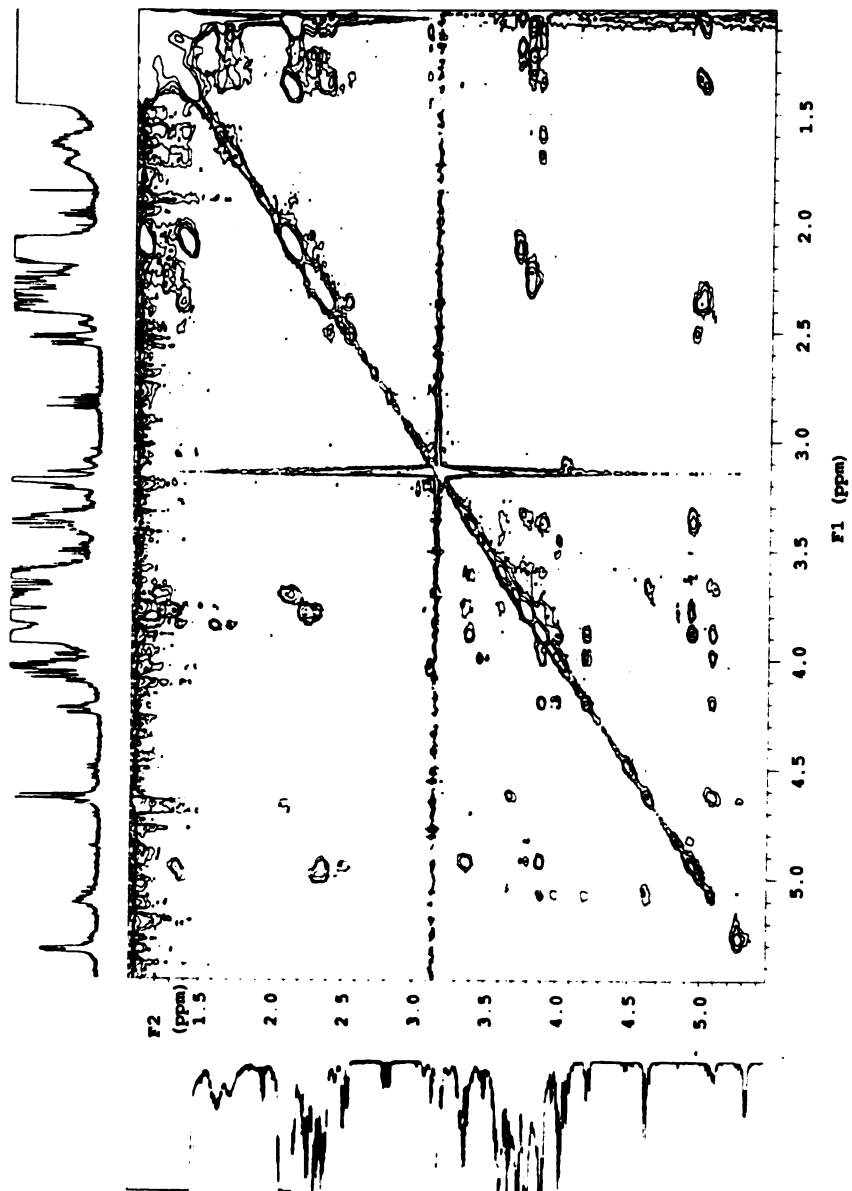**Figure 4. TOCSY spectrum of lipid A.**

Figure 5. DQF-COSY spectrum of lipid A. Cross peaks are labeled so that the unprimed numbers refer to the proximal glucosamine ring, the primed numbers refer to the distal ring and the double primed numbers refer to the fatty acid chains. 3"-OH refer to the 3-position of a fatty acyl chain in which the hydroxyl group is free and 3"-OR refer to one bearing an acyloxy substituent.
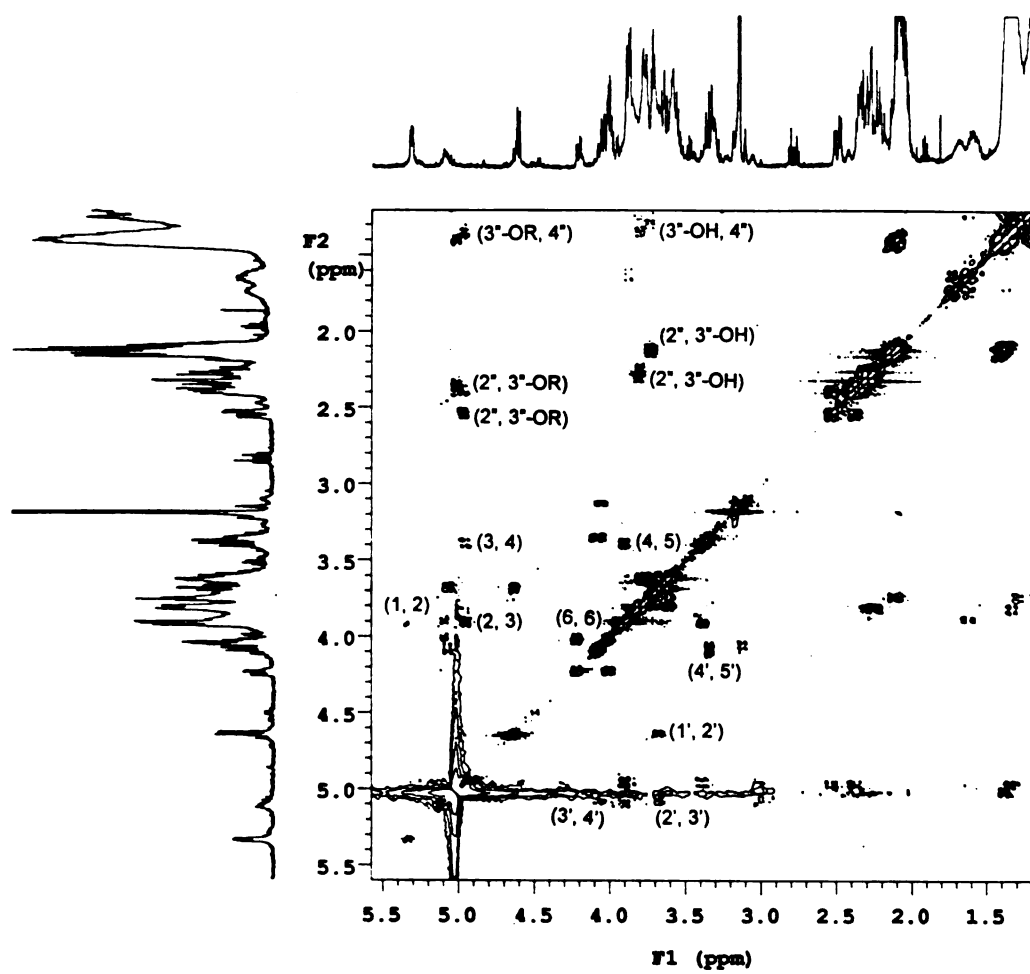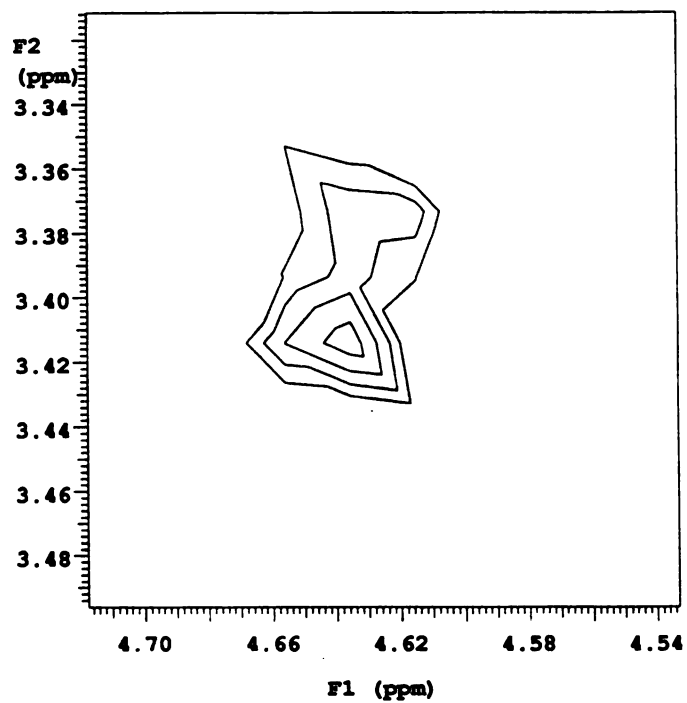
Figure 5

Figure 6. A partial ¹H/¹H NOESY spectrum of lipid A showing a strong cross peak between H-1' and H-4.

relationship (27). Signals for the two H-6 protons in question were easily discernible in the proton spectrum. The two separate sets of signals were mutually coupled with a coupling constant of 12 Hz, a typical value for germinal protons. The most downfield H-6 signal had a further splitting of 3.5 Hz and the most upfield one had no discernible splitting. Attempts were made to find a parametrized relationship which would allow the reliable calculation of dihedral angles from vicinal coupling constants. Such parametrized relationships treat the bond about which the dihedral angle is being calculated as substituted ethane fragments. The relationship should take the electronegativities of the ethane substituents, the number of substituents on the ethane fragment, the effect of the angular position of the substituents on their electronegativities and the special electronegativities of the pyran ring oxygen and the glycosidic (acetal) oxygen of the distal glucosamine residue into account. Unfortunately, no such parametrized relationship yet exists. However, it is well known that coupling constants of protons attached to carbon atoms adjacent to pyran and furan oxygens are usually ~2 ppm less than the values calculated from simple substituted alkanes in instances when the calculated coupling constant is intermediate to large. The original Karplus relationship was therefore employed and the values corrected by an increment which was in line with those observed in model compounds from the literature (28). For two alkoxy substituents, one of which is in a pyran ring, the expected reduction in coupling constant for a moderate to large value of J is ~2-3 Hz. The measured (absolute) values of J were between 0 and 1Hz (immeasurably small) for one of the H-6 protons and 3.5 Hz for the other. The calculated values were 0.2 and 6.2 Hz. The corrected calculated coupling constant for the larger splitting should then be of the order of 3.2 to 4.2 Hz, in

excellent agreement with the measured value.

The calculated structures (Figure 7-9) can easily form extended hexagonal structures by van der Waals interactions between their fatty acid chains and electrostatic interactions by magnesium or calcium/phosphate salt-bridges. Such extended hexagonal arrays have several possible configurations based on the same starting lipid A structure. These and their relative energies are being explored by Monte Carlo Simulations. The simple geometrical symmetry of the monomer unit is being exploited in these analyses.

The results described here advances a clear molecular basis for the observed structure and periodicity of lipopolysaccharide and lipid A aggregates. Other computational studies have appeared in the past (29, 30) but none have identified regular structural elements which explain the observed geometries of lipid A aggregates. These also do not incorporate any information from physical measurements. The results shown here identifies the origins of the forces and geometric attributes which make lipopolysaccharides the highly efficient, self-assembling systems that they are. The interchain forces for the alkyl groups were in agreement with those obtained using Salem's approximation (31). These forces amounted to ~70 Kcal/mole for a single isolated lipid A molecule. This is the value expected from Salem's approximation if the interactions are summed over a nearest and 6 next-nearest neighboring chains. These results should provide a firm molecular basis for rationalizing the effects of lipopolysaccharides on the dynamics and organization of membrane lipids as well as a structural basis for understanding the interaction between lipopolysaccharides and lipopolysaccharide binding molecules.

Figure 7. (A) Vector model showing van der Waals surfaces of the calculated structure of the lipid A molecule shown in Figure 1. (B) Structure seen from on top. Note the extended alkyl chains. Note the triangular cross section for the most efficient packing arrangement of the hydrocarbon chains. The 1-phosphate group is shown to the right. The leftmost alkyl chain is "A" in Figure 1. The topmost is E, the bottom is C and the one closest to the 1-phosphate group is F. Of the two remaining chains, the topmost is B and the other D.
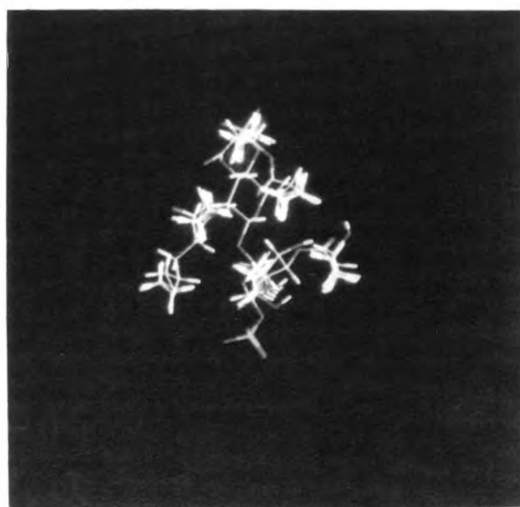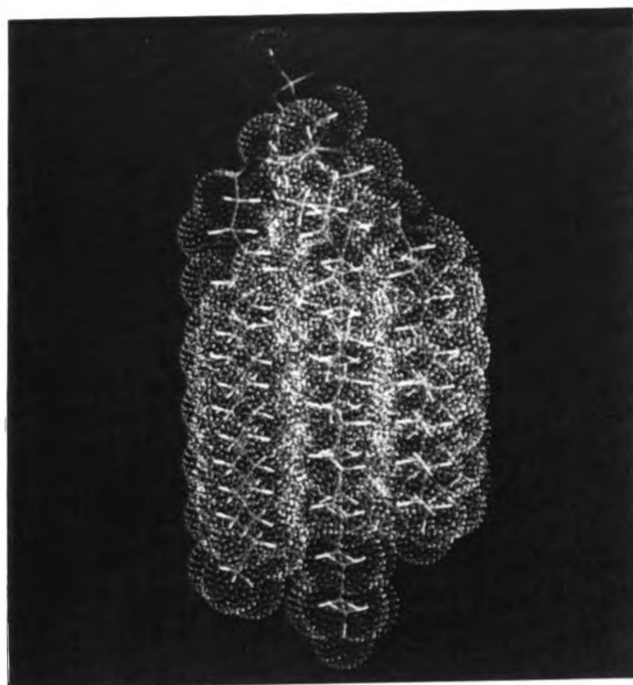
Figure 7

Figure 8. (A) A hexagonal array formed from six lipid A molecules with phosphate groups in the center bridged with calcium (white dots) ions. (B) A larger array formed from seven of the units shown in (A). The units are held together by electrostatic bridges involving calcium and the six phosphate groups around each unit. The separation between the units is slightly exaggerated for clarity.
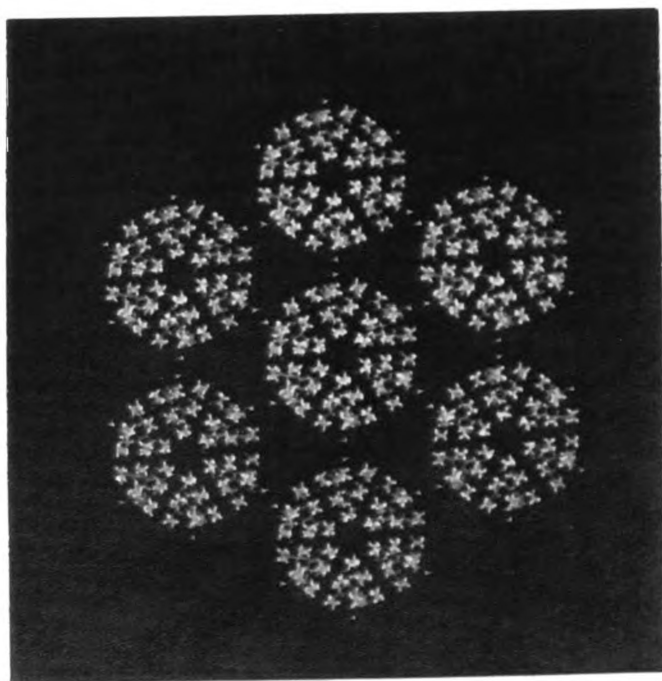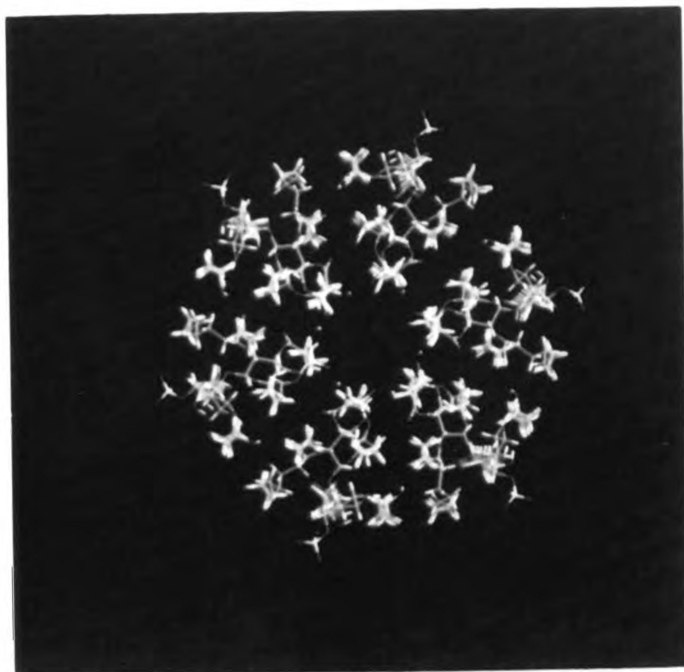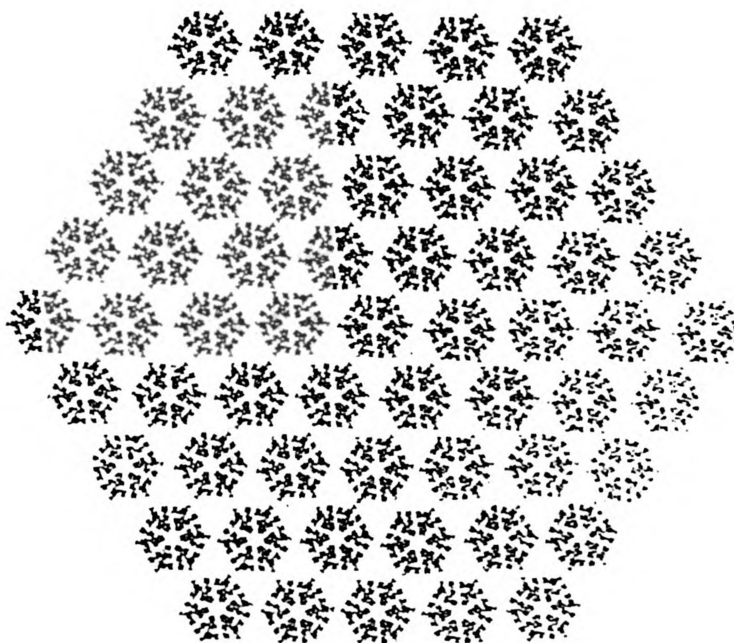
Figure 8

Figure 9. A very large hexagonal array. There are other possible hexagonal arrangements of the basic structure. These are being explored by Monte Carlo Calculations.

REFERENCES

1. Rietschel, E. T.; Brade, L.; Holst, O., Kulsin, V. A.; Linder, B., Morgan, A., Schade, U. F., Zähringer, U., Brade, H. In *Cellular and Molecular Aspects of Endotoxin Reactions Vol. 1*; Nowotny, A., Spitzer, J. J., Ziegler, E. J., Eds.; Elsevier: Amsterdam, 1990; pp 15-32.

2. Gery, I.; Waksman, B. J. *J. Exp. Med.* **1972**, 136, 143-155.

3. Turner, M.; Chantry, D.; Bochan, G.; Barrett, M.; Feldmall, M. *J. Immunol.* **1989**, 143, 3556-3561.

4. Morrison, D. C.; Ryan, J. L. *Adv. Immunol.* **1979**, 28, 293-341.

5. Schechmeister, I. L.; Bond, V. P.; Swift, M. N. *J. Immunol.* **1952**, 68, 87-93.

6. Nowotny, A. *Bacteriol. Rev.* **1969**, 33, 72-88.

7. Raetz, C. R. H. *J. Bacteriol.* **1993**, 175, 5745-5753.

8. Wright, S. D.; Ramos, R. A.; Tobias, P. S.; Ulevitch, R. J.; Mathison, J. C. *Science* **1990**, 249, 1431-1433.

9. Price, R.; Jacobs, D. *Biochim Biophys. Acta* **1986**, 859, 26-32.

10. Larsen, N.; Enelow, R.; Simons, E.; Sullivan, R. *Biochim. Biophys. Acta* **1985**, 815, 1-8.

11. Jackson, S. K.; James, P. E.; Rowlands, C. C.; Evans, J. C. *Free Radic. Res. Commun.* **1989**, 8, 47-53.

12. Jackson, S. K.; James, P. E.; Rowlands, C. C.; Mile, B. *Biochim. Biophys. Acta* **1992**, 1135, 165-170.

13. Myers, K. R.; Truchot, A. T.; Ward, J.; Hudson, Y.; Ulrich, J. T. In *Cellular and Molecular Aspects of Endotoxin Reactions Vol. 1*; Nowotny, A., Spitzer, J. J., Ziegler, E. J., Eds.; Elsevie: Amsterdam, 1990; pp 145-156.

14. Galanos, C.; Lüderitz, O. In *Handbook of Endotoxin Vol. 1*; Rietschel, E. T., Ed.; Elsevier: Amsterdam, 1984; pp 46-58.

15. Liljefors, T.; Tai, J. C.; Li, S.; Allinger, S. N. *J. Comp. Chem.* **1987**, 8, 1051-1056.

16. Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, 94, 8897-8909.

17. Molecular Simulations, Inc. Waltham, MA 02154 U. S. A.

18. Wang, Y.; Hollingsworth, R. I. *Anal. Biochem.*, in press.

19. Piantini, U.; Sorensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, 104, 6800-6801.

20. Rance, M. *Biochem. Biophys. Res. Comm.* **1983**, 117, 479-485.

21. Summers, M. F.; Marzilli, L. G.; Bax, A. *J. Am. Chem. Soc.* **1986**, 108, 4285-4294.

22. States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, 48, 286-292.

23. Shands, J. W.; Graham, J. A.; Nath, K. *J. Molec. Biol.* **1967**, 25, 15-21.

24. Shands, J. W. In *Microbial Toxins Vol. 4*; Weinbaum, G., Kadis, S., Ajl, S., Eds.; Academic Press: New York, 1971; pp 127-144.

25. Kastowsky, M.; Gutberlet, T.; Bradaczek, H. *Eur. J. Biochem.* **1993**, 217, 771-779.

26. Kato, N. *Micron* **1993**, 24, 91-114.

27. Karplus, M. *J. Chem. Phys.* **1959**, 30, 11-15.

28. Hall, L. D.; Manville, J. F. *Can. J. Chem.* **1969**, 47, 1-17 (1969).

29. Katowsky, M.; Sabisch, A.; Butberlet, T.; Brodaczek, H. *Eur. J. Biochem.* **1991**, 197, 707-716.

30. Labischinski, H.; Barnickel, G.; Bradaczek, H.; Naumann, D.; Rietschel, E. T.; Giesbrecht, P. *J. Bacteriol.* **1985**, 162, 9-20.

31. Salem, L. *Can. J. Biochem. Physiol.* **1962**, 40, 1287-1298.

CHAPTER 6

SUMMARY AND PERSPECTIVE

Throughout this research project, much has been learned about the chemical structure of the lipopolysaccharide of *Rhizobium trifolii*. Strains of *R. trifolii* 4S and ANU 843 have been studied. The structures of all three components of the *R. trifolii* 4S LPS, namely, the O-specific polysaccharide, the core oligosaccharide, and the lipid A, have been established employing a combination of chemical and spectroscopic methods as discussed earlier in this thesis. The O-polysaccharide consists of a pentasaccharide repeating unit containing rhamnose, N-acetylglucosamine and N-acetylmannosamine with the structure as follows:

α-D-ManNAc

↓

2

→3)-α-L-Rha-(1→3)-α-L-Rha-(1→4)-β-D-GlcNAc-(1→3)-α-L-Rha-(1→
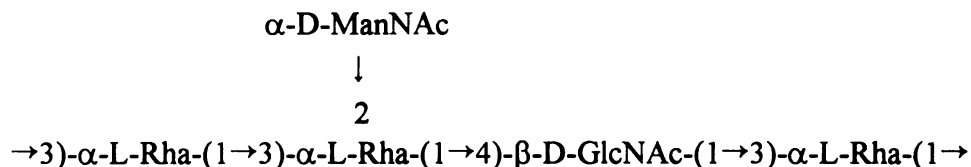
The core is found to contain a trisaccharide component composed of a 3-deoxy-D-manno-2-octulosonic acid (Kdo) and two galacturonic acid residues with the structure of α-D-GalA-(1→4)-Kdo-(5←1)-α-D-GalA. The very same trisaccharide core component has been found earlier in the LPS of *Rhizobium trifolii* ANU843 (1-3). The structure of the lipid A region of the *Rhizobium leguminosarum* has been recently proposed by Bhat at al. (4), however, their study was seriously limited in several aspects. We have proposed a rather different structure for the lipid A region of the LPS of *Rhizobium trifolii*, and indeed of all *Rhizobium leguminosarum* biovars. Our proposed structure contains an α-D-GalA-(1→6)-β-D-GlcN-(1→Lactyl ether disaccharide headgroup. The 2- and 3- positions of both the glucosamine

and galacturonic acid residues are acylated by fatty acids with acyloxyacyl groups present. The detailed structure is shown in Figure 1.

With the knowledge of the structures of all three components we can propose a complete structure for the *Rhizobium trifolii* 4S LPS as shown in Figure 1. The O-antigen and lipid A regions are both connected to the core region. The trisaccharide core is proposed to be linked to the 4'- position of the lipid A via the Kdo reducing end. This linkage is labile even to mild acid hydrolysis. The reducing end of the O-antigen is connected to the core, however, it is not clear so far which hydroxyl group of the trisaccharide is the linkage position.

The studies described herein show that, unlike most acidic extracellular polysaccharides of *Rhizobium trifolii* reported so far, the lipopolysaccharides from *R. trifolii* 4S do not contain D-galactose. The EPS of this strain also does not contain galactose (5-8). Biosynthesis studies have been initiated to determine whether this organism can synthesize galactose-containing polysaccharides if it is provided with uridine diphosphate galactose (9-11).

Another area we address in this work is the molecular basis for the form (supramolecular architecture) and function of lipopolysaccharides in bacterial membranes. Different bacteria living under different conditions, could arrive at the same functional protective molecular matrix membrane by different routes, using different molecules. These molecules may have chemical groups that bear some resemblance in molecular size, shape and charge etc., and as a consequence, can perform the same biological function (12). Although *R. trifolii* lipid A molecule has an unusual chemical structure, it has the same functional

Figure 1. **The proposed structure of the lipopolysaccharide of *Rhizobium trifolii* 4S.**
The top substructure is the O-antigen, the middle is the core, and the bottom is the lipid A.
The lipid A is linked to the reducing end of the core through its 4'-position. The binding site
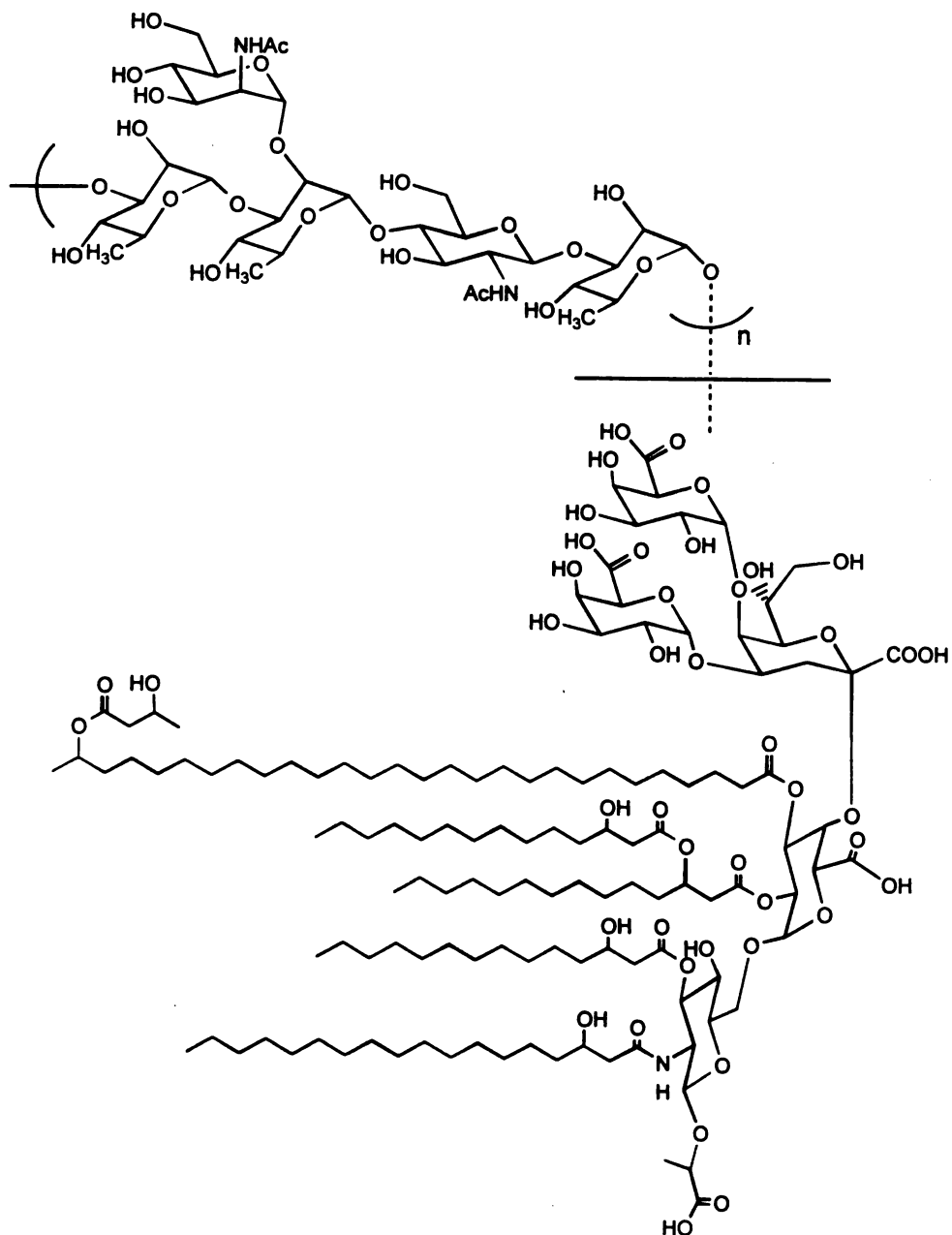of the core for the O-antigen is not known.

Figure 1

architecture as the typical lipid A and shares some common structural features. The lipid A

of *R. trifolii* like the *E. coli* lipid A contains a disaccharide headgroup with negatively charged

substitute groups. These substituents are carboxylate groups and their orientation closely

mirrors the phosphate groups in the *E. coli* lipid A. Computational models of these two lipid

A headgroups clearly shows their striking similarities. In both cases, the planes of the two

saccharide rings are orthogonal and the two negative charges are at maximum separation.

Membranes formed by these two different lipid A molecules could, therefore, have the same

surface topography with respect to the more important aspects such as charge, geometry, and

polarizability.

LPS and lipid A, like phospholipids, spontaneously organize into noncovalently

bonded, high ordered complex structures such as monolayers, vesicles, or bilayers with a

hexagonal close packing of their hydrocarbon chains perpendicular to the membrane surface.

Compared with phospholipids, LPS and lipid A are much more complex molecules with

various functional groups, and the self-assemblies formed are much higher ordered structures.

Phospholipid self-assembly derived biomaterials have been developed by several research

groups (13-17). LPS / lipid A self-assembly derived biomaterials have not been investigated

yet. However, the very same strategies used in the phospholipid materials can, potentially,

be applied to LPS and lipid A aggregates as well. The highly ordered LPS and lipid A

aggregates are attractive candidates for various applications. The hexagonal lipid A

membrane itself is a special surface which can be modified to form biosensors for use in areas

such as medical diagnostics. Its carbohydrate surface can be modified or coated with different

groups, which would dramatically change the properties of the materials. Nonlamellar lipid

A assemblies (e. g. cubic phase or vesicles) contain natural channels that might be chemically modified for different uses. In addition, lipid A assemblies can be used as templates for forming more stable structures which can be used in electronics as miniaturized conductive systems, as field-emimitting cathodes to produce high current electron beams, and as rigid microvials for controlled drug release. Self-directed LPS and lipid A supramolecular structures, whose final architectures can be precisely controlled at the molecular level, far surpass those that can be achieved in synthetic or chemical polymerization processes. In the future we expect to see much attention and development in LPS and lipid A derived materials.

REFERENCES

1. Carlson, R. W.; Hollingsworth, R. I.; Dazzo, F. B. *Carbohydr. Res.* **1988**, 176, 127-135.

2. Carlson, R. W., unpublished data.

3. Zhang, Y.; Hollingsworth, R. I.; Priefer, U. B. *Carbohydr. Res.* **1992**, 231, 261-271.

4. Bhat, U. R.; Forsberg, L. S.; Carlson, R. W. *Biochem.* **1994**, 269, 14402-14410.

5. Amemura, A.; Harada, T. *Carbohydr. Res.* **1983**, 115, 165-174.

6. Hollingsworth, R. I.; Abe, M.; Sherwood, J. E.; Dazzo, F. B. *J. Bacteriol.* **1984**, 160, 510-516.

7. Philip-Hollingsworth, S.; Hollingsworth, R. I.; Dazzo, F. B. *J. Biol. Chem.* **1989**, 264, 1461-1466.

8. Philip-Hollingsworth, S.; Hollingsworth, R. I.; Dazzo, F. B.; Djordjevic, M. A.; Rolfe, B. G. *J. Biol. Chem.* **1989**, 264, 5710-5714.

9. Tolmasky, M. E.; Staneloni, R. J.; Ugalde, R. A.; Leloir, L. F. *Arch. Biochem. Biophys.* **1980**, 203, 358-364.

10. Lelpi, L.; Couso, R. O.; Dankert, M. A. *Biochem. Biophys. Res. Commun.* **1981**, 102, 1400-1408.

11. Gardiol, A. E.; Dazzo, F. B. *J. Bacteriol.* **1986**, 168, 1459-1462.

12. Hollingsworth, R. I.; Lill-Elghanian, D. In *Cellular and Molecular Aspects of Endotoxin Reactions*; Nowotny, A., Spitzer, J. J., Ziegler, E. J., Eds.; Elsevier Science Publishers, 1990, pp 73-84.

13. Tirrell, J. G.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *C&EN* **1994**, December 19, 40-51.

14. Schnur, J. M. *Science* **1993**, 262, 1669-1676.

15. Charych, D. H.; Nagy, J. O.; Spevak, W.; Bednarski, M. D. *Science* **1993**, 261, 585-588.

16. O'Brien, D. F. *Trends in Polymer Science* **1994**, 2, 163.

17. Ghadiri, M. R.; Granja, J. R.; Buehler, L. K. *Nature* **1994**, 369, 301-304.

APPENDIX

ISOLATION OF THE CORE COMPONENT OF THE LIPOPOLYSACCHARIDE

OF *RHIZOBIUM TRIFOLII* 4S

INTRODUCTION

Rhizobia are gram-negative bacteria which form a nitrogen-fixing symbiotic relationship with legume plants. Lipopolysaccharides and capsular polysacchrides are the major surface polysaccharides, and they have all been suggested to be involved in the specific adhesion of the bacteria to the legume host and, therefore, in the symbiosis (1-3). In order to define the roles of LPS in symbiosis, it is important to characterize their structures.

Two core oligosaccharides have been isolated and characterized from the LPS of *R. trifolii* ANU843 (4-8). A tetrasaccharide fragment is composed of a D-mannose residue α 1, 5- linked to the 3-deoxy-D-manno-2-octulosonic acid, and mannose is substituted at 4-OH by D-galacturonic acid and at 6-OH by D-galactose (4, 5). All of the aldoses are in the pyranose form. The other oligosacchride is a trisaccharide consisting of two galacturonic acid residues linked to the 4-OH and 5-OH of the 3-deoxy-D-manno-2-octulosonic acid residue (6-8). We now report here the isolation of the core components for the LPS of another related strain *R. trifolii* 4S.

MATERIALS AND METHODS

Bacterial Cultures

Rhizobium trifolii 4S was grown in shaken broth culture (41) of modified Bergensen's (BIII) medium at 30°C. The medium was made of 0.23g potassium phosphate (dibasic), 0.10g magnesium sulfate (heptahydrate), 1.10g sodium glutamate, 2g mannitol, and 1 ml trace element stock per liter. The medium was adjusted to pH 6.9 with hydrochloric acid and autoclaved. 1 ml of vitamin stock was then added.

Isolation of LPS

Bacteria were grown up to early stationary stage (9) and harvested by centrifugation (8000 rpm for 40 min). The medium was discarded, and the cells were collected and extracted with a mixture of 100 g of phenol and 100 ml of water containing 0.5 g of sodium bicarbonate. The suspension was vigorously shaken with heating for 30 min, followed by sonication for 2 min, and then shaken with heating for another 10 min. The suspension was then centrifuged for 30 min at 8000 rpm, and separated into 2 layers, a water layer and a phenol layer with an insoluble interfacial residue. The aqueous layer was decanted and saved. To the phenol layer was added 100 ml of water. The mixture was then shaken vigorously for 10 min and the layers were separated by centrifugation. The aqueous fraction was combined with the previous aqueous extract. The extraction was repeated several times to ensure maximum recovery of LPS. The combined water extracts were exclusively dialyzed against distilled water. The aqueous solution was

reduced to about 10 ml by rotary evaporation and treated with 0.2 ml of a solution containing 16 units/ml each of DNase and RNase. The solution was then divided into 3 portions and subjected to gel filtration on a Sepharose 4B column eluting with 0.05M ammonium formate buffer at pH 5.5. Each 10 ml fraction was collected, and assayed for carbohydrate content using the phenol-sulfuric acid assay (10). In a typical phenol/sulfuric acid assay, 100 µl of solution was taken from each fraction tube, and 5 ml of phenol and 15 ml of concentrated sulfuric acid was added. After vigorous mixing for 1 min, the absorbance of each fraction was measured at 490 nm in a Spectronic 21 UV-Vis spectrometer. The absorbance of each assay was plotted vs. fraction number. From this plot two major peaks could be discerned. The first peak containing LPS was poured into a Spectra-Por membrane tubing (molecular weight cut-off 1200-1400), dialyzed extensively and lyophilized to give a white fluffy powder.

Isolation of Core Components

The LPS sample was hydrolyzed with 10 ml of 1% acetic acid for 2 hrs at 100°C, and then extracted with 2 volumes of 5 : 1 methanol-water to remove lipid A. The aqueous layer was lyophilized and separated on a Biogel P2 column eluting with 0.1% formic acid solution. The fractions were assayed for carbohydrate content as already describes and those containing carbohydrate were lyophilized and subjected to ^{1}H NMR analysis.

NMR Spectroscopy

All ^{1}H NMR spectra were recorded at 500 MHz for protons on a VARIAN 500 spectrometer in deuterium oxide at room temperature. The chemical shift was reported relative to the HOD line at 4.65 ppm which was surpressed by presaturation.

PRELIMINARY RESULTS AND DISCUSSION

The crude lipopolysaccharide extract was purified on a Sepharose 4B column and this yielded two major fractions (Figure 1). The first one corresponded to the LPS and the second one contained low molecular weight cellular molecules. After removal of the lipid A portion by 1% acetic acid hydrolysis, the carbohydrate region of the LPS was separated on a P2 Bio-Gel column to three major peaks as seen in Figure 2. The first eluting peak corresponded to the O-antigenic polysaccharide which contained a pentasaccharide repeating unit consisting of rhamnose, N-acetylmannosamine and N-acetylglucosamine as discussed in Chapter 2 (11).

The second peak was divided into 4 fractions (see Figure 2, fractions A-D), and each fraction was subjected to ^{1}H NMR analysis (Figures 3A-C). Fraction D was purified on a P2 Bio-Gel column again, two major peaks were obtained and their ^{1}H NMR spectra are shown in Figure 4. All the fractions from this eluting peak were similar as indicated by their proton NMR spectra. This peak remains to be further purified and characterized but appears to be from contaminating DNA. No galactose-containing tetrasaccharide component, as the one found in *R. trifolii* ANU843, was observed at this point.
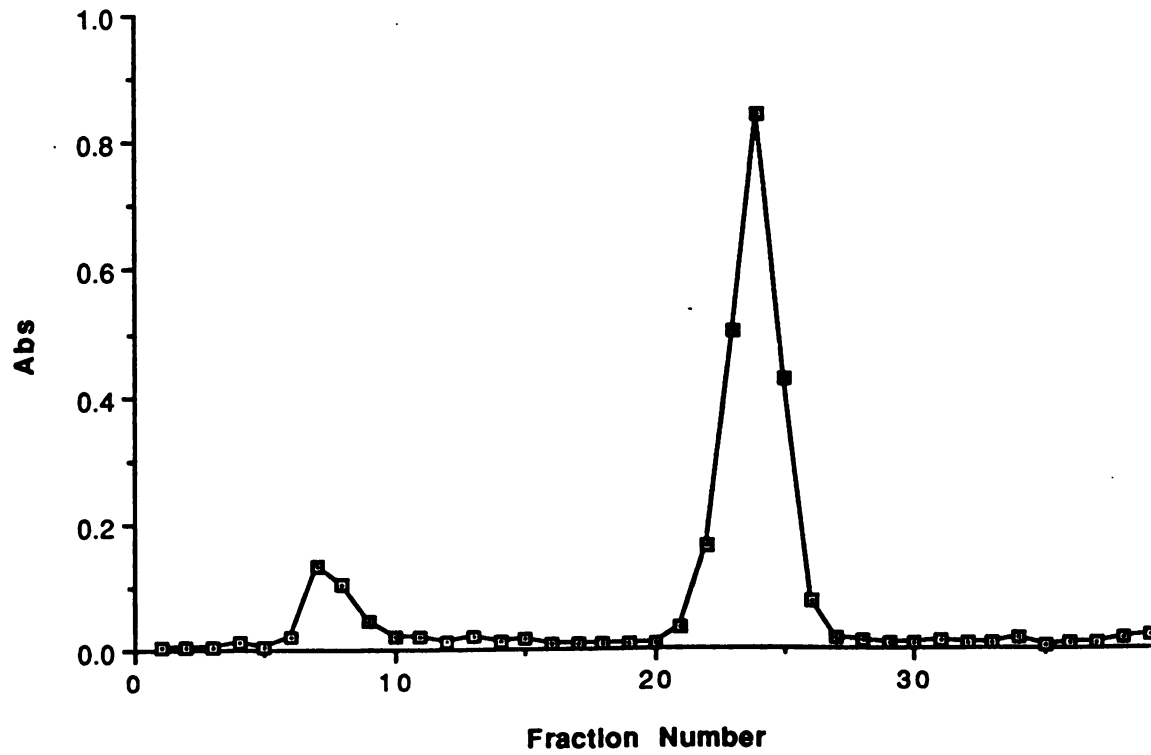
Figure 1. **Gel filtration profile of the crude LPS extract of *R. trifolii* 4S.** The first peak corresponded to the LPS. and the second peak contained low molecular weight cellular molecules.
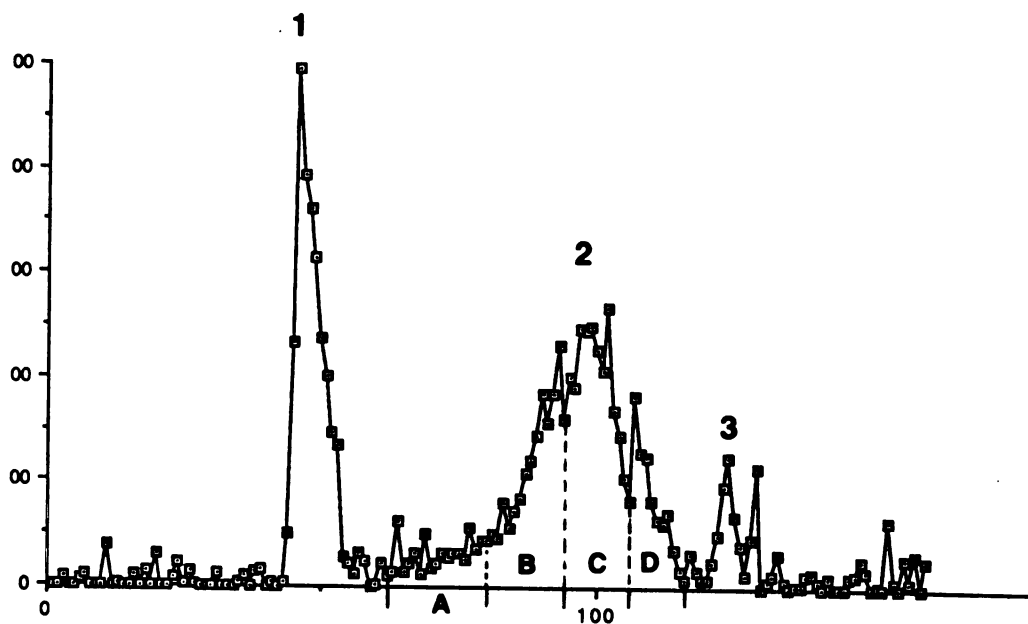
Figure 2. Elution profile of carbohydrate regions of *R. trifolii* **4S LPS separated on Biogel P2 column.** The first peak corresponds to the O-antigenic polysaccharide. The second one is further divided into 4 fractions labeled A-D. The third peak corresponds to a trisaccharide.

Figure 3. ^{1}H NMR spectra of fractions 2A-C. Spectra A-C cossecpond to fractions 2A-C respectively.
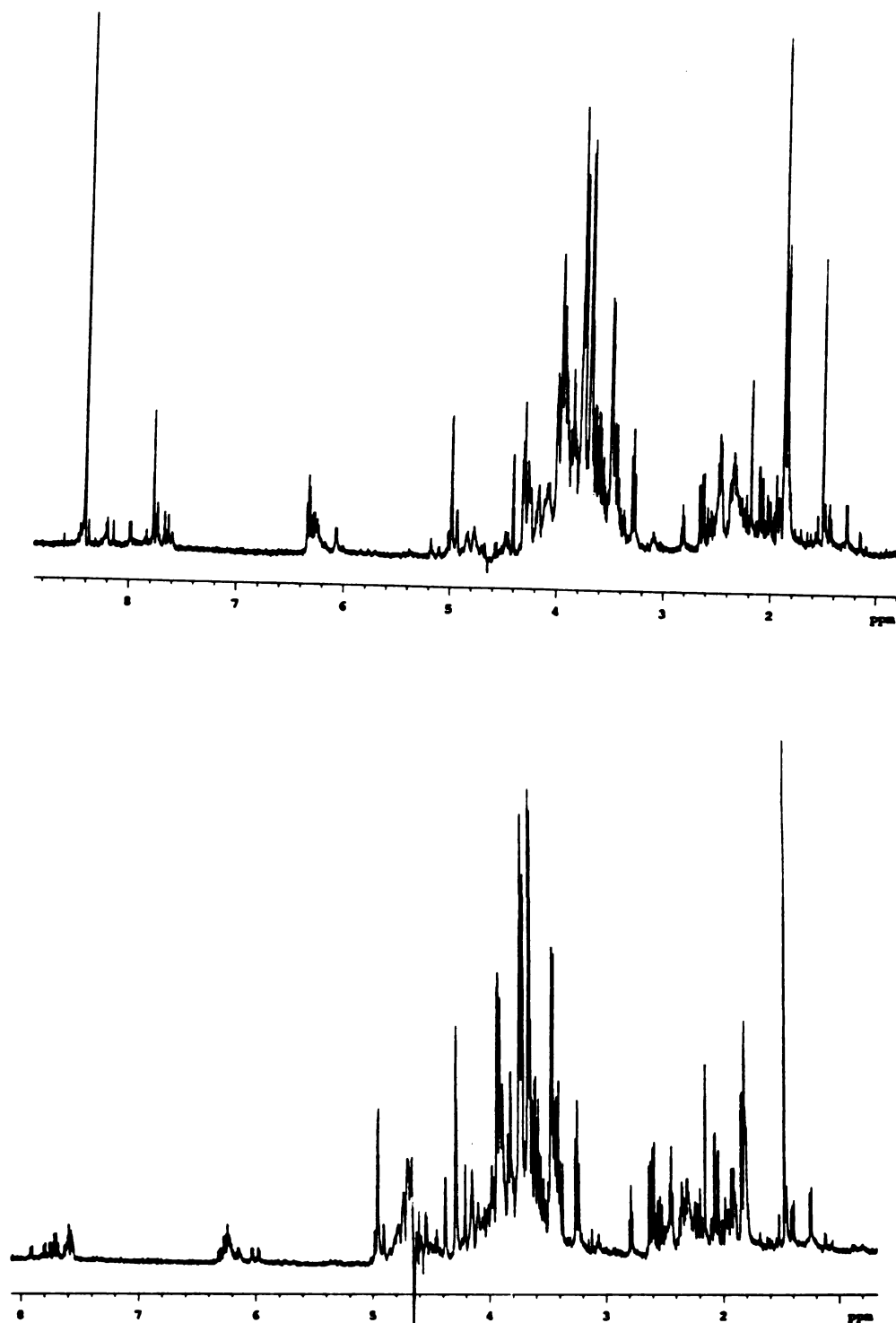
A

B

C
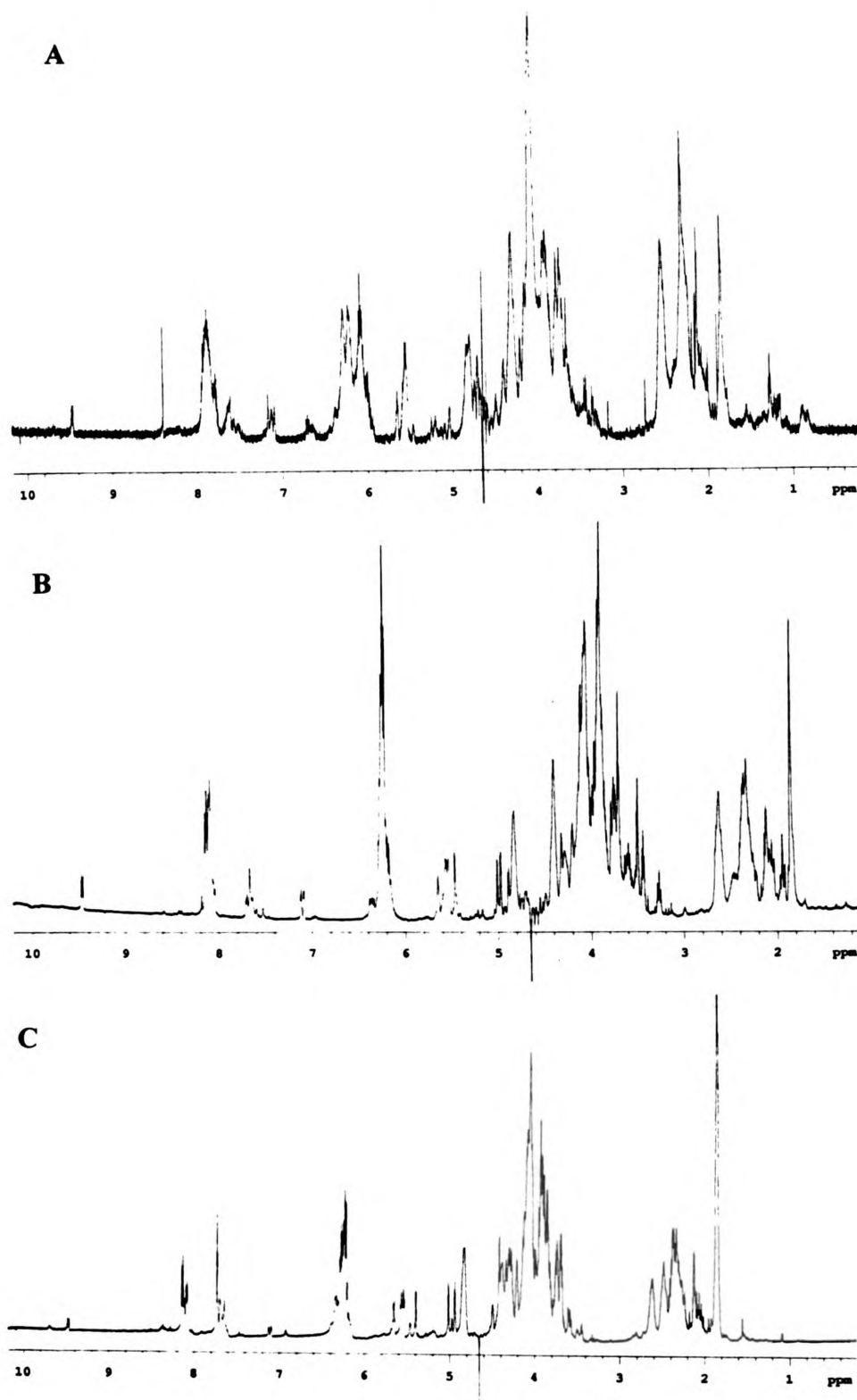
Figure 3

Figure 4. ^{1}H NMR spectra of the two fractions purified from fraction 2D.

The third peak corresponded to a trisaccharide core component, which was identical to the one found in *R. trifolii* ANU843 LPS (6, 7) as revealed from the ^{1}H NMR result. It was composed of a Kdo and two D-galacturonic acid residues with the structure of 3-deoxy-4,5-di-O-(α-D-galactopyranosyluronic acid)-β-D-manno-2-octulopyranosic acid. Two doublets at 5.18 ppm (J 3.8 Hz) and 5.12 ppm (J 3.8 Hz) in the ^{1}H NMR spectrum (Figure 5) were assigned to the anomeric protons of the two galacturonic acid residues respectively. The methylene proton of the Kdo residue appeared at 1.92 ppm as a doublet of doublets (J 12.5 and 2.9 Hz) and 2.18 ppm as a triplet (J 12.5 Hz). The triplet was assigned to the axial proton (H-3a) which should have a large splitting from the neighboring trans-vicinal proton and another large splitting from the other germinal proton. The equatorial proton (H-3e) should have a comparatively small vicinal coupling, as it is cis to the proton on C-4. The structure of this trisaccharide is shown in Figure 6.

In summary, the LPS of *R. trifolii* 4S consist of a trisacchardie core, but no tetrasaccharide component was observed as in the LPS of *R. trifolii* ANU843. The result also suggest that this organism, unlike other *R. trifolii* strains, might not synthesize galactose since no galactose is found in any of the surface polysaccharides.
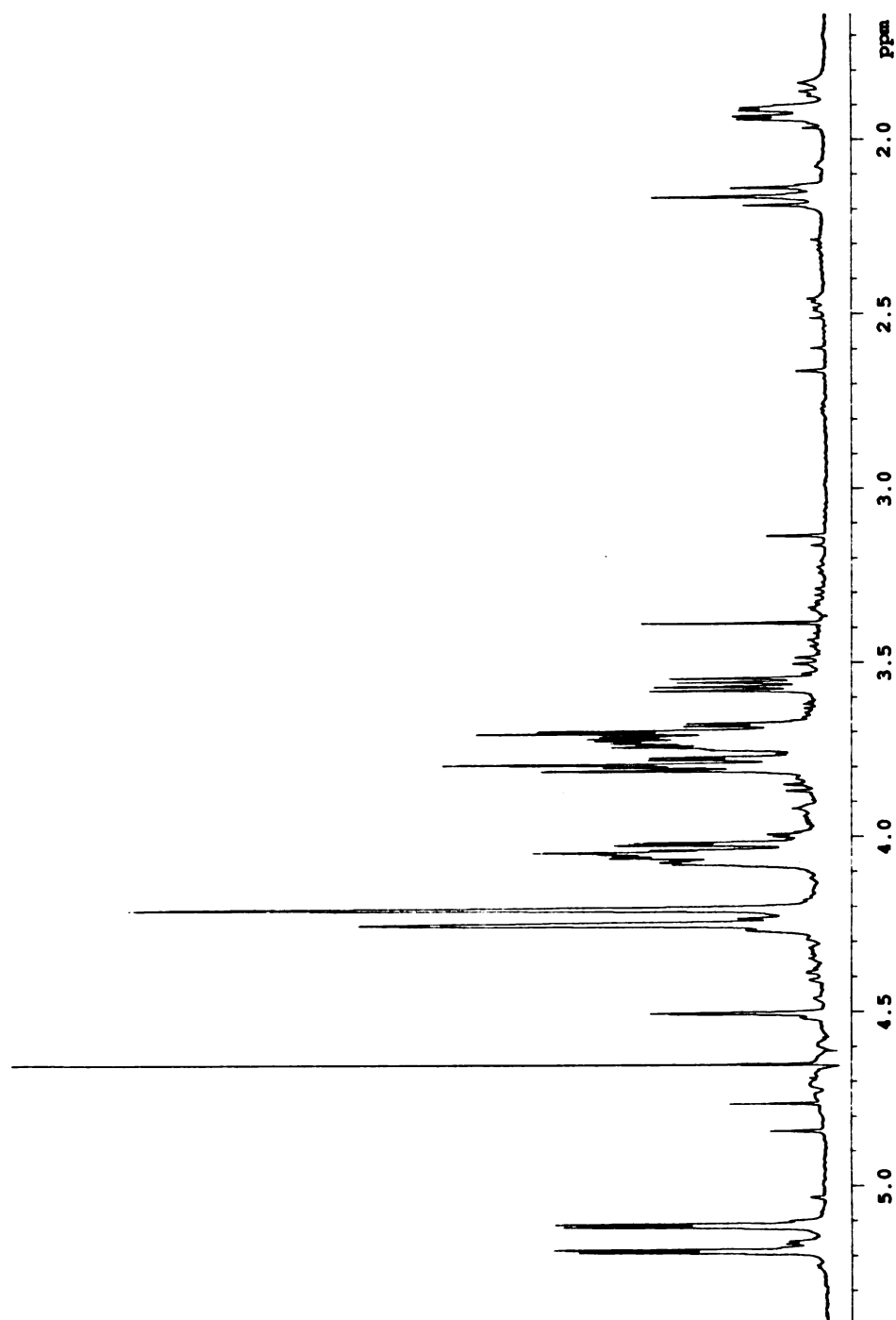
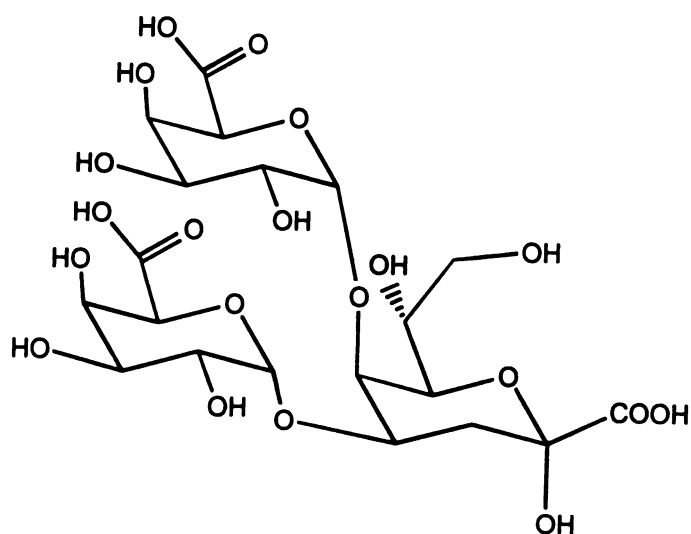Figure 5. ^{1}H NMR spectrum of the trisaccharide.

Figure 6. Structure of the trisaccharide. (After Ref. 6-8.)

REFERENCES

1. Bauer, W. D. *Annu. Rev. Plant Physiol.* **1981**, 32, 407-449.

2. Carlson, R. W. In *Surface Chemistry in Nitrogen Fixation: Rhizobium*; Broughton, W. J., Ed.; Clarendon Press: Oxford, 1982; pp. 199-234.

3. Dazzo, F. B.; Hubbel, D. H. In *Control of Root Hair Infection in Nitrogen Fixation: Rhizobium*; Broughton, W. J., Ed.; Clarendon Press: Oxford, 1982; pp. 274-310.

4. Hollingsworth, R. I.; Carlson, R. W.; Garcia, F.; Gage, D. A. *J. Biol. Chem.* **1988**, 264, 9294-9299.

5. Hollingsworth, R. I.; Carlson, R. W.; Garcia, F.; Gage, D. A. *J. Biol. Chem.* **1988**, 265, 12752.

6. Calson, R. W.; Hollingsworth, R. I.; Dazzo, F. B. *Carbohydr. Res.* **1988**, 176, 127-135.

7. Carlson, R. W., unpublished data.

8. Zhang, Y.; Hollingsworth, R. I.; Priefer, U. B. *Carbohydr. Res.* **1992**, 231, 261-271.

9. Bishop, P. E.; Huevara, J. A.; Engelke, J. A.; Evans, H. J. *Plant Physiol.* **1976**, 57, 542-546.

10. Dubois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F. *Anal. Chem.* **1956**, 25, 350-356.

11. Wang, Y.; Hollingsworth, R. I. *Carbohydr. Res.* **1994**, 260, 305-317.