STRUCTURAL STUDIES ON THE CAPSULAR ANTIGENS OF SOME *Escherichia coli* SEROTYPES

THESIS

Submitted in Fulfilment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

of Rhodes University

by

Margaret Ruth Leslie

November 1994

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Professor H. Parolis, for his advice, encouragement and support for the last three years.

I would also like to thank the following people:

Dr L.A.S. Parolis, for her advice and assistance throughout this work.

Mr D.L. Morley for valuable technical assistance and help with the production of this thesis.

Mr A Soper, for technical assistance.

Mr A.W. Sonemann for his assistance with mass spectrometry.

Drs F. and I. Ørskov, Statens Seruminstituut, Copenhagen, for cultures of the bacteria.

Dr Y.M. Choy of the Chinese University of Hong Kong for performing the ß-elimination studies on the *E. coli* K43 polysaccharide.

The Foundation for Research Development, the Loewenstein Educational Trust and Rhodes University for financial support.

My colleague, Dr D.V. Whittaker, for his advice and support.

Dr S. Robertson and Mr G. Hooper, for time spent proof reading the manuscript.

My husband, Julian Grue, for continued advice and support.

TABLE OF CONTENTS

ABBREVIATIONS AND SYMBOLS

BIRD	Bi-rotational
CI	chemical ionization
COSY	correlation spectroscopy
CTAB	cetyltrimethylammonium bromide
δ	chemical shift in parts per million
DCI	direct chemical ionization
DEAE	diethylaminoethyl
DMF	dimethylformamide
DMSO	dimethylsulphoxide
EI	electron impact
eV	electron volts
f	furanose
FAB	fast atom bombardment
FID	free induction decay
Fru	fructose
FT	Fourier transform
Fuc	fucose
FucNAc	*N*-acetylfucosamine
g	gram
Gal	galactose
GalA	galacturonic acid
GalNAc	*N*-acetylgalactosamine
GLC	gas liquid chromatography
Glc	glucose
GlcA	glucuronic acid
GlcNAc	*N*-acetylglucosamine
GPC	gel permeation chromatography
h	hour
HETCOR	heteronuclear correlated spectroscopy
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum coherence
HOHAHA	homonuclear Hartman Hahn
HPAEC	high performance anion exchange chromatography
HPLC	high performance liquid chromatography
Hz	Hertz
IR	infra-red

kd	kilo dalton
KDO	2-keto-3-deoxy*manno*-octulosonic acid
L	litre
M	molar
Man	mannose
ManNAc	*N*-acetylmannosamine
ManNAcA	*N*-acetylmannosaminuronic acid
mg	milligram
min	minute
mL	millitre
mM	millimolar
mol	mole
M_r	average molecular mass
MS	mass spectrometry
m.u.	mass unit
m/z	mass-to-charge ratio
µL	microlitre
Neu5Ac	neuraminic acid
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser enhancement spectroscopy
p	pyranose
PAD	pulsed amperometric detection
PAAN	peracetylated aldononitrile
PMAA	partially methylated alditol acetate
ppm	parts per million
Pyr	pyruvate
Rha	rhamnose
Rib	ribose
ROESY	rotating frame overhauser enhancement spectroscopy
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TMS	tetramethylsilane
TPPI	time proportional phase increments
TOCSY	total coherence transfer spectroscopy
TSR	total sugar ratio

Symbols and abbreviations not listed here are defined in the text

ABSTRACT

The research presented in this thesis forms part of an on-going collaborative programme concerned with the determination of the chemical structures of the surface antigens of bacteria belonging to genera within the family *Enterobacteriaceae*. Bacteria of this family are opportunistic pathogens and are frequently responsible for serious infections in animals and man. Surface antigens produced by virulent strains are largely polysaccharides and occur as lipopolysaccharides (the O-antigens) and capsular polysaccharides (the K-antigens) respectively. The extracellular polysaccharide antigens expressed by strains of the species *Escherichia coli* are of considerable interest due to their effect on immunological processes and the relationship which exists between their chemical structure and virulence. To date, some seventy-four K-antigens have been distinguished serologically within the species *E. coli* and structures have been determined for most of these. The K-antigens of *E. coli* are structurally diverse and exhibit serological cross-reactivity with other pathogenic bacteria.

The structures of five previously unstudied *E. coli* K-antigens, *viz.* those produced by serotypes O20:K101:H$^-$, O8:K45:H9, O8:K50:H$^-$, O101:K103:H$^-$, and O8:K43:H11, are presented in this thesis. A variety of chemical techniques has been employed in the structural analysis, and these are discussed. Two-dimensional NMR spectroscopic techniques proved invaluable for the structural elucidation of these complex carbohydrates, and high-field NMR spectroscopy alone was used in the analysis of the K43 antigen. Structural analysis of the K103 antigen was facilitated by specific enzymatic degradation, using a bacteriophage-borne endoglycanase. The K45 antigen was found to contain the unusual sugar 3-acetamido-3,6-dideoxygalactopyranose, while the K50 and K103 antigens join a minority group of polysaccharides which contain pyruvate as their only acidic component.

CHAPTER ONE: INTRODUCTION

1.1 BACTERIAL POLYSACCHARIDES — AN OVERVIEW

Bacterial polysaccharides comprise a large group of polymers which exhibit almost unlimited structural variation, and include several which are commercially important. Xanthan gum, obtained from *Xanthamonas campestris*, is used in the petroleum oil, food, pharmaceutical, textile, and printing industries due to its unique rheological properties in solution[1-3], while dextrans are utilized extensively in the pharmaceutical and fine-chemicals industries[1,4]. A large number of bacterial polysaccharides are of considerable medical importance since they are immunogenic and are responsible for the pathogenicity and disease specificity of the organism in which they occur. Some have been used to develop bacterial vaccines effective against a number of diseases, including pneumococcal pneumonia and meningococcal meningitis[2,5-7]. In addition, the dextrans from *Leuconostoc mesenteroides* are used medically as blood plasma substitutes[3,8].

Many types of polysaccharide occur in bacteria. The majority exhibit a high degree of immunological specificity and are produced by only one bacterial type, species or group. Specific polysaccharides are usually heteropolysaccharides comprising repeating units of between two and six monosaccharides. Specific polysaccharides containing significant non-carbohydrate portions are commonly encountered, in addition to such non-specific homopolysaccharides as glycogen, cellulose, glucans, fructans and mannans[1].

The systematic study of bacterial polysaccharides is complicated by the ongoing reclassification of bacteria, but broadly speaking bacteria belonging to the Kingdom *Procaryotae* can be divided into three main groups, *viz.* Gram-positive bacteria (Division *Firmicutes*), Gram-negative bacteria (Division *Gracilicutes*) and bacteria which lack rigid cell walls (Division *Tenericutes*), although a recent phylogenetic classification of bacteria places this third group with the Gram-positive bacteria[9]. Bacteria belonging to the Kingdom *Archaebacteria* possess unusual cell walls, although little is known about their composition. The specific polysaccharides produced by Gram-positive and Gram-negative bacteria are discussed in the following sections.

1.1.1 POLYSACCHARIDES PRODUCED BY GRAM-POSITIVE BACTERIA

1.1.1.1 The peptidoglycan layer

The peptidoglycan layer confers rigidity on the cell and comprises approximately 50% of the cell wall in Gram-positive bacteria[2,3]. The basic structure of this layer is consistent for all bacteria and is composed of linear polysaccharide chains, each of approximately 200 disaccharide repeating units. The repeating unit comprises alternating 1→4 ß-linked units of *N*-acetyl-glucosamine and *N*-acetylmuramic acid, the latter being a derivative of *N*-acetylglucosamine which is unique to peptidoglycan[2,3]. Tetrapeptide side-chains are linked to each *N*-acetylmuramic acid residue via a peptide link at the carboxyl group, and the amino acids are cross-linked at various points to build up the structured peptidoglycan lattice. The degree of cross-linking varies from as little as 20% in *Micrococcus luteus* to as much as 90% in organisms such as *Staphylococcus aureus*[2,3,10]. Figure 1.1 shows two disaccharide units with cross-linking between the peptide chains.

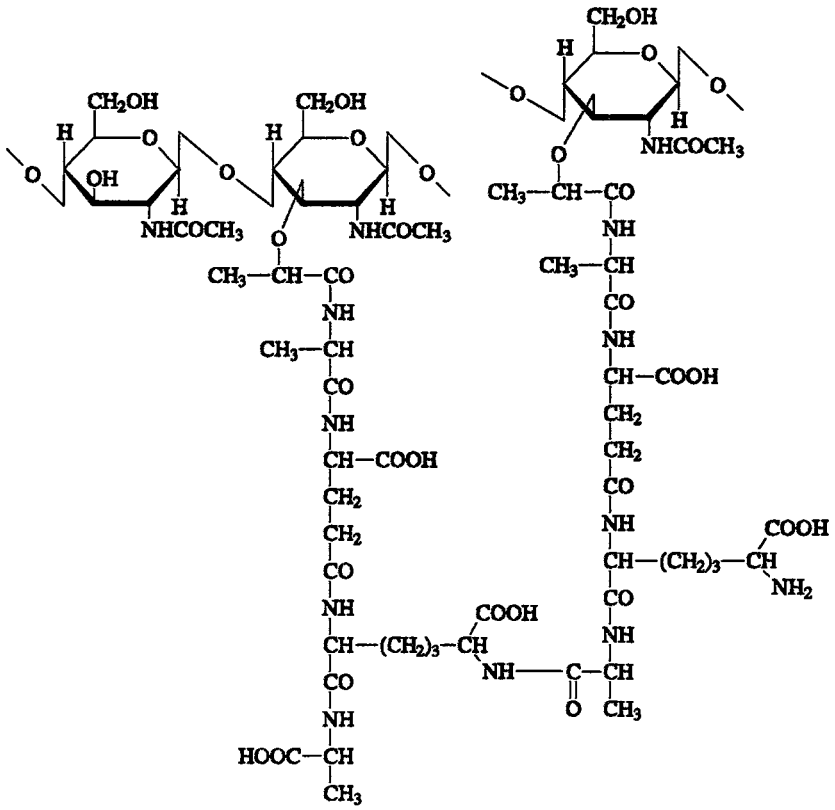


Figure 1.1 Disaccharide unit of peptidoglycan showing cross-linking to an adjacent chain[2]

The peptidoglycan layers of different bacterial species may vary with regard to the nature of the amino acids comprising the tetrapeptide chains, the type and position of cross-linking between them and the presence of substituents on the *N*-acetylmuramic acid residue. In *Staphylococcus aureus* approximately half of the *N*-acetylmuramic acid residues bear an acetyl group at position 6[11,12]. A high degree of *O*-acetylation in *Lactobacillus acidophilus* causes the peptidoglycan to be resistant to lysozyme[13]. This resistance is also observed in *Bacillus cereus* and is ascribed to the presence of free amino groups in the polymer[14].

1.1.1.2 Teichoic and Teichuronic Acids

These two polymer types, along with protein constituents, comprise the remaining 50% of the Gram-positive cell wall[2]. Teichoic acids are polymers comprising repeating units of sugars such as glycerol phosphate and ribitol phosphate which are joined through phosphodiester linkages[1,10]. Other sugar residues may also be incorporated into the repeating unit. In addition to the sugar residues there may be D-alanyl groups present, attached by an ester linkage to a free hydroxyl group on one of the sugar residues. These alanyl ester groups are extremely labile and their concentration is critically dependent on pH[2]. The teichoic acid molecule is covalently linked to the peptidoglycan layer via a phosphodiester linkage unit which couples one end of the chain to position 6 of a muramic acid residue on the peptidoglycan[15,16]. Teichoic acids are exclusive to Gram-positive bacteria and are found in nearly all species belonging to this Division[10]. A typical cell wall teichoic acid is shown in Figure 1.2.

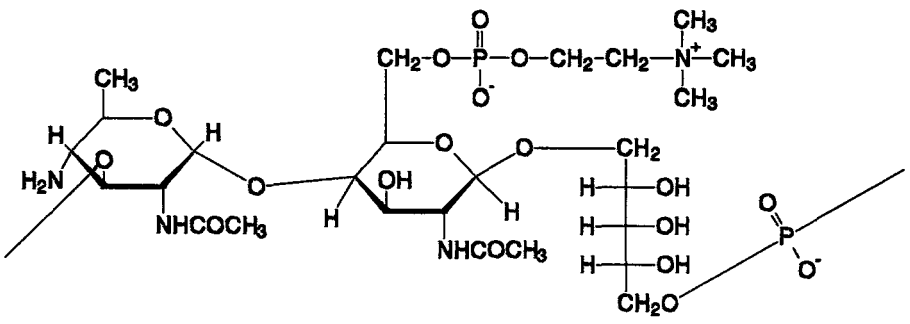


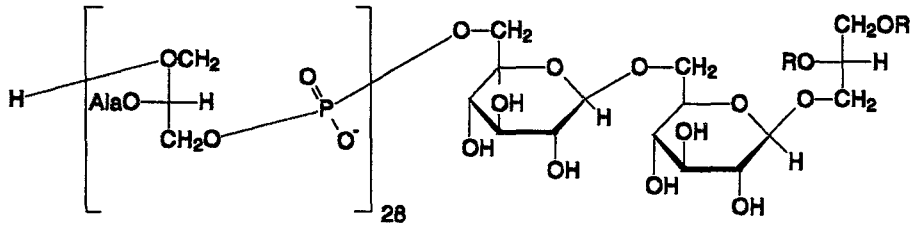
Figure 1.2 The repeating unit of the wall teichoic acid of *Streptococcus pneumoniae*[2]

In many Gram-positive bacteria which are grown in phosphate-deprived conditions, wall teichoic acids disappear and are replaced by teichuronic acids, which have similar properties[1,2,10]. Teichuronic acids are linear polysaccharides containing uronic acids and are covalently linked to the peptidoglycan layer. A phosphodiester linkage unit has been proposed[17,18]. In contrast to the teichoic acids, this is the only phosphate group found in the polysaccharide.

Teichoic acids are believed to mediate access of divalent cations, particularly Mg^{2+}, to the cell[19,20], and to play a role in controlling autolytic enzymes within the cell[10]. Some wall teichoic acids are also antigenic and are therefore responsible for the immunological properties of the bacteria[21].

1.1.1.3 Lipoteichoic acids

Gram-positive bacteria which have cell-wall teichoic acids also produce polyglycerolphosphate teichoic acids which are covalently linked to glycolipid, which serves to anchor the molecule to the cell membrane[22]. The glycerol units of the molecule are linked in a 1,3 manner by phosphodiesters, as shown in Figure 1.3. The carbohydrate portion of the molecule protrudes through the wall to the outer surface of the cell and is antigenic in nature[23].



Ala = Alanyl R = long chain fatty acid

Figure 1.3 The structure of the lipoteichoic acid of *Staphylococcus aureus*[2]

1.1.1.4 Neutral polysaccharides

The cell walls of some Gram-positive bacteria contain neutral polysaccharides which are covalently linked to the peptidoglycan. There is evidence to suggest that this linkage is again via a phosphodiester group[24], and like the lipoteichoic acids these polysaccharides are antigenic.

The antigens from *Streptococcus pneumoniae*, some of which are teichoic acids and some of which are neutral wall polymers, have been studied extensively. The species has been divided into 83 different types based on the type-specific antigen produced by each strain[25]. A number of neutral cell wall polysaccharides have been identified in strains of *Eubacterium saburreum*. The antigen of strain 49 (Figure 1.4) contains an unusual sugar occurring in the furanose form which was tentatively identified as 6-deoxy-D-*altro*-heptose[28].

Neutral antigenic polysaccharides from several species of *Bifidobacterium* have also been studied[27,28].

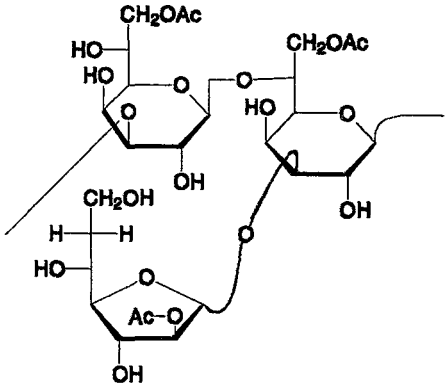


Figure 1.4 The structure of the antigen of *Eubacterium saburreum strain 49*[28]

1.1.1.5 Capsular polysaccharides

Some Gram-positive bacteria produce capsular polysaccharides which are not covalently attached to the peptidoglycan layer but are loosely associated with it. Over 85 immunologically distinct capsular polysaccharides have been identified in the species *Streptococcus pneumoniae* and are used as a basis for typing of strains[2]. Structures have been determined for many of these capsular types. Other Gram-positive organisms, for example some strains of *Staphylococcus aureus*, also possess capsular polysaccharides but relatively few structures have been established. Most Gram-positive capsules are heteropolysaccharides composed of repeating units, but a few

homopolysaccharide capsules also occur. Examples are the dextrans produced by *Leuconostoc mesenteroides* and *Streptococcus mutans*, which aid adherence of the organism to teeth with the resulting production of cavities.

1.1.2 POLYSACCHARIDES PRODUCED BY GRAM-NEGATIVE BACTERIA

1.1.2.1 The peptidoglycan layer

The peptidoglycan layer of the Gram-negative bacteria has the same basic structure as that found in Gram-positive bacteria, but the degree of crosslinking is much lower, commonly dropping to as little as 20% in *Escherichia coli*[2,10]. The type of cross-linking is also simpler, with fewer peptide chains being present[10]. The peptidoglycan layer comprises a much lower percentage of the total cell wall than in Gram-positive bacteria, resulting in a less rigid structure. A major difference between Gram-positive and Gram-negative bacteria, as shown in Figure 1.5, is the presence in the latter of an outer membrane which encloses the cell wall. The components of this membrane are of more relevance to the interaction of the Gram-negative bacterium with its environment, and hence its pathogenicity, than the wall itself, and will be considered in greater detail.

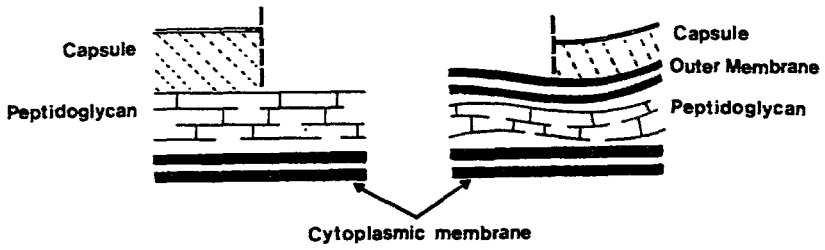


Figure 1.5 **Diagrammatic representation of the cell surface of Gram-positive and Gram-negative bacteria[2]**

1.1.2.2 Lipopolysaccharides

Although the term lipopolysaccharide indicates any molecule containing both carbohydrate and lipid in significant amounts, in bacteriology it denotes the specific endotoxic O-antigens obtained from

"smooth" Gram-negative bacteria and the corresponding products from "rough" mutant strains. All Gram-negative bacteria contain lipopolysaccharides (LPS) in this restricted sense. Such molecules consist of three parts, *viz.*, Lipid A, the core and the O-specific side-chain (Figure 1.6).

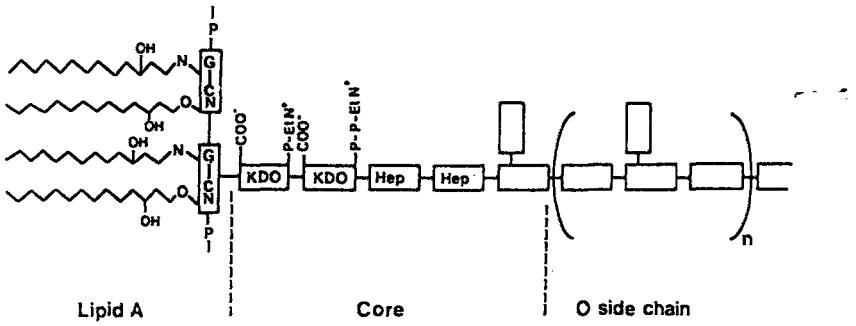


Figure 1.6 Generalised structure of bacterial lipopolysaccharide[2]

Lipid A is the most conserved part of the molecule, with the main conserved entity being a disaccharide of glucosamine, to which fatty acids are linked[1,2]. The terminal residue of lipid A is linked to the core region by 3-deoxy-D-manno-octulosonic acid (KDO). The core region is a complex oligosaccharide which contains a KDO dimer as an indispensable part of the structure[1,2]. The core regions of different species and families have the same general structure, and most have a terminal glucose residue. In smooth wild strains of bacteria, the core region of the LPS is linked to an O-specific side-chain, which is the immunologically dominant part of the molecule and is usually referred to as the O-antigen. The O-antigen is composed of oligosaccharide repeating units, often containing rare sugars. The variations in sugar composition, linkage positions, configuration of glycosidic linkage, branching of oligosaccharide chains, and the presence of various non-carbohydrate substituents accounts for the large number of O-serotypes in Gram-negative bacteria. Up to one hundred different serotypes have been described in *Salmonella* alone[2] and structures of a large number of the O-antigens of *Salmonella* and *Escherichia coli* have been determined.

1.1.2.3 Enterobacterial Common Antigen

Enterobacterial common antigen (ECA) is a component found in most strains of *Enterobacteriaceae*. It is thought to be located in the exterior leaflet of the outer membrane[2], and occurs in two forms, *viz.* an immunogenic form present in R-mutants where it is believed to be linked to the

lipopolysaccharide, and a hapten form present in wild-type strains which is linked to an unidentified carrier molecule[29]. The immunodeterminant portions of the two forms are identical. Early work performed on ECA isolated from various *Salmonella* species indicated that the antigen was composed of oligosaccharide repeating units and also appeared to contain palmitic and acetic acids[29]. More recent studies performed on ECA isolated from *Shigella sonnei* permitted identification of the repeating unit of the carbohydrate portion of the antigen[30,31]. The antigen was shown to be a cyclic polysaccharide containing four, five or six trisaccharide repeating units, this being the first example of a cyclic heteropolysaccharide[31]. It was found that the small amount of fatty acid present was not chemically linked to the antigen. ECA isolated from other sources has been shown to have the same structure as that from *Shigella sonnei* (Figure 1.7), with the modification that in some cases the *N*-acetylglucosamine is not acetylated at position 6[33-35]. The function of ECA is unknown.

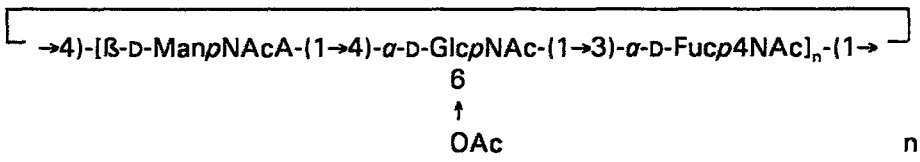


Figure 1.7 The structure of ECA isolated from *Shigella sonnei*[30]

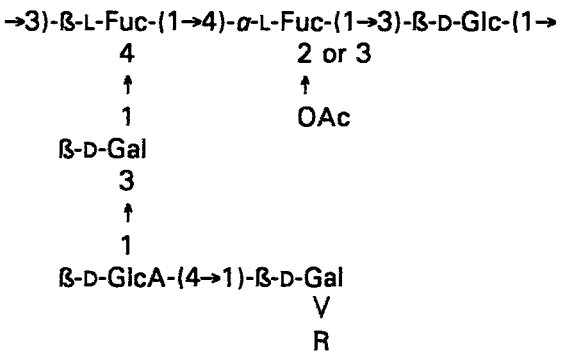
1.1.2.4 Capsular polysaccharides

Most Gram-negative bacteria produce capsular polysaccharides which are usually acidic. Their function is the same as that of Gram-positive capsules, *viz.* they protect the organism against phagocytes by masking the highly immunogenic O-antigens. Capsular polysaccharides are poorly immunogenic and do not elicit an immediate response from the immune system of the host, thus allowing the organism to establish itself firmly in the host. Many capsules suppress the specific immune response to a large degree since they are similar to or identical with host structures[35]. The structures of many Gram-negative capsular polysaccharides have been determined, revealing a high degree of structural diversity. A few homopolysaccharide structures have been studied, including capsules of strains of *Escherichia coli* K1 and *Neisseria meningitidis* B and C, which are sialic acids[1,2,36]. The majority of Gram-negative capsules are heteropolysaccharides made up of

oligosaccharide repeating units comprising between two and seven sugar residues, although a seventeen-residue repeating unit has been proposed recently for the capsular polysaccharide from *Agrobacterium radiobacter*[37]. The capsular polysaccharides of *Escherichia coli*, being particularly relevant to this thesis, are discussed in greater detail in Section 1.2.

1.1.2.5 Other exopolysaccharides

All bacteria belonging to the family *Enterobacteriaceae*, when grown at low temperature and high salt concentration, produce a mucous polysaccharide known as the M-antigen. M-antigens from different bacterial strains have the same carbohydrate backbone, shown in Figure 1.8. The group acetalically linked either at O-3 and O-4 or O-4 and O-6 of the terminal ß-D-galactose residue may be pyruvic acid, acetaldehyde or formaldehyde[38,39].



R = 3,4-*O*-carboxyethylidene, 3,4-*O*-ethylidene, or 4,6-*O*-carboxyethylidene

Figure 1.8 The carbohydrate backbone of M-antigen[38,39]

A strain of *Pseudomonas fragi*, responsible for the rotting of meat, has recently been found to exude an exopolysaccharide in the form of a glycocalyx, enabling it to adhere to the surface of the meat[40]. The structure of this polysaccharide is shown in Figure 1.9.

→4)-3-*O*-[(*R*)-1-carboxyethyl]-α-D-Glc*p*-(1→3)-ß-D-Man*p*NAc-(1→4)-ß-D-Glc*p*-(1→

Figure 1.9 The structure of the glycocalyx from *Pseudomonas fragi* ATCC 4973[40]

Heteropolysaccharides lacking regular repeating units have been obtained from *Pseudomonas aeruginosa*[41] and *Azotobacter vinelandii*[42,43]. These polysaccharides resemble those produced by marine algae. Their production is variable, the quantity depending on the conditions under which the organisms are grown[2].

1.2 THE CHEMISTRY OF THE K-ANTIGENS OF *ESCHERICHIA COLI*

The bacterial species *Escherichia coli* forms part of the natural colonic flora in man and a number of strains are consequently harmless. However, many strains are opportunistic pathogens and are able to cause a wide variety of urinary tract infections and colonic conditions[44]. More virulent strains may cause septicaemia, appendicitic peritonitis, meningitis and pyelonephritis[45]. The majority of *E. coli* strains are encapsulated, and 103 different capsular antigens (K-antigens) were originally identified[46]. These K-antigens were subdivided into categories A, B and L on the basis of their O-antiserum agglutinability and thermolability[44]. Difficulties were encountered in distinguishing between L and B antigens, since the basis for distinction did not reflect any differences in chemical composition. Two of the antigens, *viz.* K88 and K99, were subsequently found to be proteinaceous and were reclassified as F-antigens, while others were not shown to be independent of the O-antigens and were therefore removed from the list[44]. At present, seventy-four polysaccharide K-antigens are recognised and structures have been determined for most of these. Recent reviews by Dutton and Parolis[47] and De Bruin[48] list most of the known structures.

The K-antigens are characteristically acidic heteropolysaccharides comprised of oligosaccharide repeating units of two to six sugar residues, with the exception of K1 and K92, which are homopolymers of Neu5Ac[36,49]. Neutral hexoses are most common, but uronic acids, 6-deoxyhexoses, deoxyaminohexoses and pentoses are also commonly encountered. A number of rare or unusual sugars have been found, eg. D-Man*p*NAc in K84[50] and K50 (this thesis), D-ManNAcA in K7[51], L-Fuc*p*NAc in K87[52], KDO*f* in K95[53], D-Fuc*p*3NAc in K45 (this thesis), 4-deoxy-2-hexulosonic acid in K3[54], and the new diamino sugar L-Rha*p*2,3diNAc in K48[55]. Table 1.1 lists

monosaccharides found in *E. coli* K-antigens.

Table 1.1 Monosaccharides present in the capsular polysaccharides of *E. coli*

MONOSACCHARIDE	CAPSULAR ANTIGEN(S)
D-Glc*p*	Common
D-Gal*p*	Common
D-Gal*f*	K2, K53, K93
D-Man*p*	Common
L-Fuc*p*	K27, K28, K33, K42, K87
L-Rha*p*	Frequent, eg. K3, K103*
D-Fru*f*	K4, K52, K11
D-Rib*f*	Frequent, eg. K6
D-Glc*p*A	Common
D-Gal*p*A	Common
D-Glc*p*NAc	Common
D-Gal*p*NAc	Frequent, eg. K9, K101*
D-Man*p*NAc	K84, K50*
L-Fuc*p*NAc	K87
D-Man*p*NAcA	K7
KDO*p*	Frequent, eg. K6
KDO*f*	K95
Neu*p*5Ac	K1, K9, K92
D-Fuc*p*3NAc	K45*
L-Rha*p*2,3-diNAc	K48
4-deoxy-2-hexulosonic acid	K3

*This thesis

Non-carbohydrate substituents may also be present. Phosphate is found as the acidic component of a large number of K-antigens and *O*-acetyl groups are common. The latter are significant as they affect the antigenicity and immunological specificity of the capsule[56]. Pyruvate, acetalically linked at positions *O*-2 and *O*-3, positions *O*-3 and *O*-4, or, most commonly, positions *O*-4 and *O*-6, has also been found in a number of capsules and is the sole acidic component of K37[57], K47[58], K50

(this thesis) and K103 (this thesis). *O*-propionate is present in non-stoichiometric amounts in the K14[59] and K52[60] antigens. The amino acids L-serine and L-threonine have been found amidically linked to uronic acids in K40[61], K49[62] and K54[63].

Now that more extensive information pertaining to the chemical composition of the polysaccharide antigens is available it is possible to divide them into two groups based on chemical, physical and microbiological criteria[56,64,65].

Group I polysaccharides are co-expressed with O-groups O8, O9, O20, and O101. They have high molecular masses and low charge densities, and may be substituted at the reducing end with core-lipid A. The acidic component in this group of polysaccharides is usually GlcA, GalA or, in a few cases, pyruvate. Group I polysaccharides can be further divided into those containing amino sugars and those without. The latter closely resemble the K-antigens of *Klebsiella* and in fact are related to them at a supra-generic level, eg. *E. coli* K55 has the same structure as *Klebsiella* K5[66].

Group II polysaccharides are not restricted to distinct O groups, generally have low molecular masses and high charge densities, and some are substituted at the reducing end with phosphatidic acid. The group can be further subdivided according to the nature of the acidic component[65]. Many group II polysaccharides contain phosphate. Of these the K18, K22 and K100 antigens are polyribosyl ribitol phosphates[67,68] and are closely related to the capsular polysaccharide of *Haemophilus influenzae b*[65]. Other polysaccharides in this sub-group show structural similarities to capsular polysaccharides of *Neisseria meningitidis b*. A number of the group II polysaccharides contain KDO or a related sugar, 4-deoxy-2-hexulosonic acid, which is present in the K3 antigen[54]. Nine out of thirteen KDO-containing polysaccharides also contain ribose[56]. Only three K-antigens have been reported to contain Neu5Ac, *viz.* K1[36], K9[69] and K92[49], while ManNAcA has been found only in K7[51].

These groupings are summarised in Table 1.2, adapted from Jann *et al.*[56].

Table 1.2 Grouping of capsular polysaccharides of *E. coli*

PROPERTY	CAPSULAR POLYSACCHARIDE GROUP	
	I	II
Acidic component	GlcA, GalA, pyruvate	GlcA, Neu5Ac, KDO, ManNAcA, phosphate
Expressed below 20°C	Yes	No
Coexpression with	O8, O9, O20, O101	Many O antigens
Lipid at reducing end	Core Lipid A*	Phosphatidic acid
Removal of lipid at pH 5-6/100°C	No	Yes
Molecular mass	> 100 kd	< 50 kd
Amino sugars present	Yes	No
Chromosomal determination at (close to)	*rfb(his), rfc(trp)*	*kpsA(serA)*
CMP-KDO synthetase activity elevated	No	Yes
Intergeneric relationship	*Klebsiella*	*H. influenzae, N. meningitidis*

* This substitution has been verified with only a few polysaccharides

There is evidence for the existence of a third group of capsular polysaccharides, provisionally termed Group I/II[70]. Expression of these capsules is not temperature-regulated, as for Group I, but the polysaccharides, like Group II, are substituted with phospholipid at the reducing end[71]. The capsular antigens of *E. coli* K96 and K54, which have identical structures except that the K54 polysaccharide is substituted with threonine and serine, belong to this intermediate group[72].

CHAPTER TWO: ISOLATION AND PURIFICATION OF BACTERIAL POLYSACCHARIDES

2.1 ISOLATION PROCEDURES

It is a prerequisite of all natural products chemistry that the material to be examined be obtained in as high a yield and as pure a form as possible, since structural studies performed on impure material or a mixture of compounds are difficult to interpret. It is also important that the techniques employed for isolation and purification are sufficiently mild to preclude the loss of labile substituents and degradation of the compound itself. To achieve both aims requires careful formulation of methodologies, and it is often necessary to perform the isolation procedure more than once if sensitive linkages are later discovered to be present in the compound.

In the case of bacterial polysaccharides, the isolation technique employed is dependent on whether the bacterial culture is grown on solid medium or in liquid culture, and on the type of polysaccharide to be studied. Westphal and Jann have described the isolation of capsular polysaccharide directly from a liquid culture[73]. Isolation of polysaccharides from bacterial cultures grown on solid medium essentially involves solubilisation of the required polymer with subsequent separation from other cellular components. The isolation of exopolysaccharides such as capsular antigens, gum exudates and slime polysaccharides may proceed under relatively mild conditions since these compounds are not strongly attached to the cell; however, in order to solubilize polysaccharides which are linked to the cell wall or membranes, such as lipopolysaccharides, harsher conditions are often necessary. A wide range of isolation techniques have been established for the different types of bacterial polysaccharides and these have been reviewed[74-7]. The method employed in this laboratory for the isolation of capsular polysaccharides is a modification of that described by Okutani and Dutton[78].

Serologically homogenous encapsulated bacteria are propagated on an appropriate sterile nutrient agar medium at 37°C until actively growing colonies are present. In this laboratory either Mueller-Hinton or Luria-Bertani medium is used for *Escherichia coli* (Appendix I). Single colonies are removed and transferred to sterile tubes containing sterile nutrient broth and, following an

incubation period of approximately 8 hours, are poured onto sterile stainless steel trays containing the appropriate nutrient agar. The trays are then incubated at 37°C, which is considered to be the optimal temperature for capsule production. Ørskov *et al.* have reported that the growth temperature may have a profound effect on the expression of the capsule in *E. coli*[79]. The bacteria are harvested after approximately 18-24 hours by scraping the culture off the agar surface, and are then suspended in a solution of 1% phenol which kills the cells without lysis. Longer periods of incubation result in desiccation of the medium and, in some cases, the production of a non-specific galactan in addition to the usual capsular antigen[86]. The cell suspension is stirred at 4°C to allow dissolution of the capsular polysaccharide, and subsequent ultracentrifugation removes cellular matter. The dissolved polysaccharide is usually recovered at this stage by precipitation into ethanol. However, many strains of *E. coli* produce thin, tight capsules and the concentration of capsular polysaccharide in the supernatant is often very low. In such cases, we have found that the yield of polysaccharide may be increased by as much as five-fold if the supernatant is first dialysed against running tap water to remove the phenol, freeze-dried, and the resulting solid then reconstituted in a minimum quantity of distilled water. The acidic capsular polysaccharide may be isolated at this stage from this more concentrated solution by precipitation with a 5% solution of cetyltrimethylammonium bromide (CTAB), which effectively separates it from other soluble products such as neutral O-antigen. The polysaccharide-CTAB complex is collected by centrifugation, dissociated by dissolution in aqueous 3 M NaCl and precipitated into ethanol. The precipitate is redissolved in water, dialysed exhaustively against tap water to remove salt and residual CTAB, ultracentrifuged to remove residual cellular debris and freeze-dried to yield the K-antigen as the sodium salt.

A problem sometimes encountered with K-antigens of *E. coli* is the substitution of the polymer with lipid. This interferes with the CTAB complexation process and it is often beneficial to treat the crude isolate with 1% acetic acid followed by centrifugation to remove precipitated lipid before the addition of CTAB.

2.2 PURIFICATION PROCEDURES

Fractionation procedures, such as the CTAB-complexation described above, do not always result in pure polysaccharide material and further purification is often necessary. Sometimes the polysaccharide may be sufficiently purified by a repetition of the fractionation process[74]. A number of complexation agents are available besides CTAB, although the latter seems to be the most widely used[74,80]. Neutral polysaccharides may also be purified by fractionation as their borate complexes[81].

Depending on the nutrient medium used for growth of the bacterial culture and on the isolation method used, a polysaccharide sample may be contaminated with nucleic acids and protein. The isolated material may be treated with proteases and nucleases (trypsin, RNAse or DNAse) to effect purification. Chromatographic methods provide a useful alternative to these chemical procedures, with the added advantage that information concerning the nature of the polysaccharide may be obtained at the same time as purification is achieved.

2.2.1 Gel permeation chromatography

Gel permeation chromatography (GPC) is a technique which is routinely used for the purification and analysis of biological compounds and has been reviewed extensively[82,83]. The technique involves a column packed with a stationary phase of swollen, porous gel beads which is eluted using an appropriate solvent. To achieve a reasonable flow rate in a large column the solvent may be pumped through the stationary phase. Separation of compounds is based on the principle of molecular exclusion. The gel beads of the stationary phase have a fixed pore-size into which molecules below this size are able to penetrate, but larger molecules are not. The larger molecules thus pass through the column by moving through the spaces between the gel beads and are eluted sooner than the smaller molecules, which diffuse into the pores and thus spend more time associated with the stationary phase. The shape of the molecule also plays a role in the separation process, since a small but bulky molecule will also not be able to penetrate the gel pores and will be eluted sooner than a less bulky molecule. The type of gel bead used depends on the required

molecular exclusion range, the type of solvent to be used and the required flow rate. For carbohydrates the solvent is usually pure water or aqueous buffer. A range of gels is available for separation of compounds up to a molecular mass of 20 x 10^6. As well as effecting purification of a polysaccharide or oligosaccharide sample, GPC is also useful for molecular mass determinations, since the elution volume (V_e) of a particular compound can be compared to elution volumes of standard polysaccharides (such as dextrans[82]).

The column eluent is usually monitored continuously by being allowed to flow through a detector such as a differential refractometer, and the readings are traced out by a chart recorder to provide an elution trace. A wide range of specific colorimetric detection techniques is also available[84], although these are not now routinely used except in cases where, for example, the use of a buffer gradient precludes the use of differential refractometry.

2.2.2 Ion-exchange chromatography

This technique involves the use of a stationary phase gel or resin in which charged sites have been covalently bonded to the gel structure. Each charged site is associated with a counter ion which can be exchanged for a charged molecule. The charged sites may be anionic or cationic.

Tap water, used in dialysis of polysaccharide preparations, frequently contains ions, particular sodium ions, which associate with acidic groups on the molecule, such as the carboxylic acid functionality of a uronic acid residue. The sodium form of a polysaccharide frequently has different physical properties from its corresponding free acid form, which can be regenerated by passing a solution of the polysaccharide through a cation exchange resin such as Amberlite™ or Dowex™. This process also serves to separate the acidic capsular polysaccharide from neutral contaminants such as O-antigen. Inorganic salts may also be removed in this way.

If charged sites are incorporated into a molecular exclusion gel, improved sample resolution can be achieved. For carbohydrates, anionic exchange gels are more useful. A range of such gels with

diethylaminoethyl (DEAE), aminoethyl (AE) and quaternaryaminoethyl (QAE) groups is available. Separation is achieved by first adsorbing the polysaccharide onto the gel and eluting with buffer to remove uncharged components. The sample is then eluted from the column using gradient elution. In these cases colorimetric analysis of the fractions is essential, the phenol-sulphuric acid method being a convenient one for routine use[85]. Excellent separation of bacterial polysaccharides and their degradation products containing uronic acids is achieved on DEAE-Sepharose™ CL-6B, as shown in Figure 2.1.

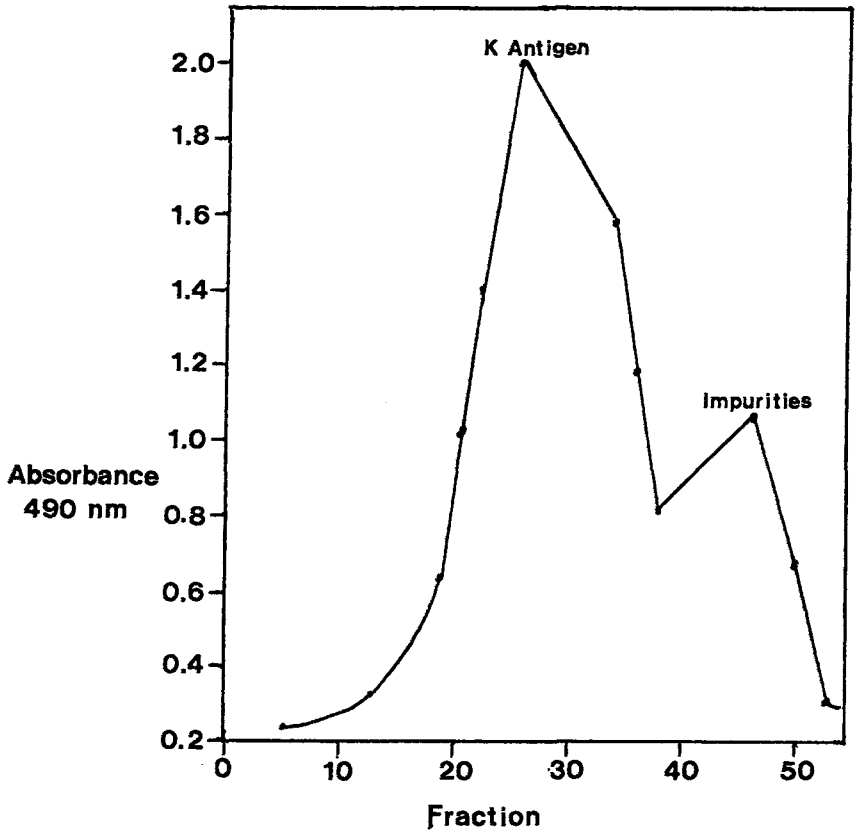


Figure 2.1 Elution diagram of the capsular polysaccharide of *E. coli* K50 on DEAE-Sepharose CL-6B

CHAPTER THREE: STRUCTURAL ANALYSIS OF COMPLEX POLYSACCHARIDES

Once a sufficient quantity of pure polysaccharide is available, structural studies may commence. Structural elucidation techniques have been revolutionised since the development of analytical techniques such as gas-liquid chromatography (GLC), high-performance liquid chromatography (HPLC), mass spectrometry (MS), and nuclear magnetic resonance (NMR) spectroscopy, which require very small amounts of material to obtain definitive results. NMR spectroscopy has the added advantages of being non-destructive and requiring no chemical modification of the material to be studied. The traditional techniques involving chemical degradation are now used to provide complementary evidence for structural features, and in some cases to provide oligosaccharide fragments suitable for further study. Several techniques have been developed[75,86] and it is not possible to discuss all of these in this review. Rather, a few of the more important techniques will be discussed and emphasis placed on those used in our laboratories.

Structural elucidation of bacterial polysaccharides is simplified in so far as most of the specific polysaccharides are composed of regular repeating units, whereas plant and algal polysaccharides are often long chains of monosaccharides with no apparent regularity in structure[88]. Complete structural elucidation of such a polysaccharide thus requires determination of the monosaccharide composition, conformation, and configuration, the linkage position and type, and the presence and position of non-carbohydrate substituents. The chemical repeat unit may not be the same as the biological repeat unit[1], and it has been shown that *Escherichia coli* and *Salmonella enterica* strains have characteristic preferred O-antigen chain lengths[87]. Since one of the primary motivations for studying bacterial polysaccharides is an attempt to correlate structure with activity, complete structural elucidation will also include chain-length determination, which can be achieved by determining the average molecular mass of the polymer. The conformation of the polymer in solution is thus also of interest.

3.1 DETERMINATION OF MONOSACCHARIDE COMPOSITION

3.1.1 Total hydrolysis

The classical approach to compositional analysis of a polysaccharide involves degradation to the constituent monosaccharides which may then be identified by a variety of analytical techniques[75]. The degradation procedure must hydrolyse the polysaccharide completely to its constituent sugars in order to obtain an accurate total sugar ratio (TSR), but must not be so harsh that the monosaccharides obtained are themselves degraded to any significant extent. Various acids have been used to effect total hydrolysis, including hydrochloric acid, sulphuric acid, formic acid, nitric acid, oxalic acid[88], trifluoroacetic acid (TFA)[89], and liquid hydrogen fluoride (HF)[90]. In general, TFA and sulphuric acid are preferred as they have been found to cause least degradation[91], although Manna *et al.* recently found that even TFA hydrolysis produces degradation products such as anhydro sugars[92]. TFA has the added advantage that it can be removed from the reaction mixture under reduced pressure and therefore does not have to be neutralized[93]. Although HF hydrolysis usually proceeds with minimum degradation and excellent release of components, it is not generally used because of its corrosiveness and toxicity. However, Sanger and Lamport have developed an apparatus which allows for comparatively safe handling of the reagent[90]. In our laboratories, we have found that treatment with 4 M TFA at 120°C for 1 h is sufficient to effect total hydrolysis with minimal degradation for most applications[93].

Polysaccharides containing uronic acids or 2-amino-2-deoxyhexoses are resistant to hydrolysis[86]. When uronic acids are present total hydrolysis without accompanying decomposition is rarely possible, and in these cases it is preferable to reduce the carboxyl function prior to hydrolysis so as to obtain the corresponding neutral hexose. This also prevents lactonisation of the uronic acids, which can cause problems in analysis[94]. Reduction may be effected by first treating the polymer with anhydrous 3% methanolic HCl (methanolysis)[95] with subsequent reduction of the carboxylic acid methyl ester with sodium borohydride. Alternatively, the method proposed by Taylor and Conrad[96], which involves activation of the carboxyl groups with aqueous carbodiimide followed by

reduction with sodium borohydride, can be used. Reagents such as lithium aluminium hydride[97] or lithium triethyl borohydride[98] in THF can be used to effect reduction of hydrophobic polymers, eg. methylated polymers.

Acid hydrolysis of polysaccharides containing amino sugars results in *N*-deacetylation, and the protonated free amino group thus formed stabilises the glycosidic linkage. Polystyrene sulphonic acid has been used to effect hydrolysis of polymers containing amino sugars[99]. A number of mixed reagent systems which preserve the *N*-acetyl group have also been developed, including a mixture of acetic acid, water and TFA (5:75:20)[100], hydrogen chloride in ethanethiol (mercaptolysis)[101], and acetic anhydride, acetic acid and sulphuric acid (acetolysis)[102]. Liquid HF is also a useful reagent, since it does not in fact cleave *N*-acetyl groups and achieves quantitative hydrolysis with minimal degradation[103].

Less severe hydrolysis conditions are appropriate for more labile residues, eg. 6-deoxy sugars, 2-deoxy sugars, and furanosidically-linked sugars, and for polymers which decompose easily. Methanolysis[95] and enzymic hydrolysis[104] are effective techniques in such cases.

Once total hydrolysis has been achieved, the monosaccharide constituents must be identified. The commonly-occurring monosaccharides may be identified easily by analytical techniques such as GLC or HPLC, providing standards are available.

3.1.2 Gas-liquid chromatography

Since its introduction in 1952 by James and Martin[105] GLC has become one of the most extensively used analytical techniques in scientific research and industry. Separation of compounds is based on partition between an immobilized liquid phase and a mobile gaseous phase. The original packed GLC columns have been superseded by the more efficient capillary columns for routine analytical work[106,107]. Capillary columns have an internal diameter of 0.2-0.5 mm and are up to 30 m in length. The stationary phase may be coated on the inner wall of the column itself, as in a wall-

coated open tubular (WCOT) column, on a solid support, as in a support-coated open tubular (SCOT)[108] column, or on a crystalline porous layer (PLOT). The column itself is usually composed of fused silica which has greater flexibility and is thus more durable than a glass column. Wide-bore capillary columns, with an internal diameter of 0.53 mm, were developed in 1983 for larger sample sizes and allow interfacing of the gas chromatography with a Fourier transform infra-red spectrometer (GLC-FTIR)[109]. Although the resolution obtained with capillary columns is far superior to that obtained with the traditional packed columns, problems relating to column bleed at higher temperatures are still encountered[110]. Non-volatile components also become adsorbed onto the solid support and stationary phase, and the resolution capabilities of the column gradually deteriorate as a result. Bonded-phase columns have the stationary phase chemically bonded to the fused silica wall. These columns therefore have much greater thermal stability and may be flushed with solvent to remove the build-up of non-volatile compounds[110].

Many stationary phases are available in capillary columns and a variety have been employed in the separation of carbohydrates[111]. In this laboratory bonded-phase columns with stationary phases of intermediate polarity such as DB-1, DB-17 and DB-225 are used for routine monosaccharide analysis, while the polar DB-WAX column is used to separate derivatised amino sugars. The compositions and applications of a few stationary phases are listed in Table 3.1.

One of the advantages of open tubular capillary columns is that their higher specific permeability allows a much lower flow rate, reducing the cost of the carrier gas and allowing a smaller sample volume. A split injection system allows the majority of the sample to be voided, thus preventing column overload. For efficient analysis of the nanogram quantities of material which actually reach the column, extremely sensitive detection systems are needed. Flame ionization, thermal conductivity and electron capture detectors are routinely used[112]. Eluting compounds are recorded as peaks on a chromatogram by a chart recorder or recording integrator and are usually identified by comparison with standard compounds. Quantitative estimations are made by comparing the areas under the appropriate peaks. The response of the detector may vary with different classes

of compounds, and this becomes important when analyzing polysaccharides which do not contain repeating units. However, when dealing with polysaccharides composed of regular repeating units exact quantitation is not essential as the ratio of the sugars to each other provides sufficient information. In these cases the molar response factor, which is estimated by assuming that each type of carbon atom contributes to the flame response to the same extent in all compounds[113], gives satisfactory results. In our laboratory we assume that response factors are equal on an equal weight basis.

Table 3.1 Compositions and applications of some stationary phases[114]

Stationary Phase	Composition	Applications
DB-1	100% Methyl silicone, gum	Acetylated and trimethylsilylated oligosaccharides, disaccharide acetates
DB-17	50% Methyl silicone 50% Phenyl silicone	Partially methylated alditol acetates and acetylated methyl glycosides, peracetylated alditols, pertrimethylsilylated alditols, trimethylsilyl ethers, derivatised amino sugars, pertrimethylsilylated ketoximes of Neu5Ac and NeuGlyc, peracetylated ketoximes
DB-225	25% Cyanopropyl silicone 25% Phenyl silicone 50% Methyl silicone	partially methylated alditol acetates, peracetylated alditols, pertrimethylsilylated alditols, trimethylsilylated *O*-methyloximes
DB-WAX	Polyethylene glycol	Derivatised amino sugars, methyl ethers, acetals, acetates

The monosaccharide products of total hydrolysis are not sufficiently volatile to allow direct analysis by GLC and require derivatisation. In addition, the hydroxyl, carboxyl and amino groups present in the compound may cause adsorption onto the column surface, causing tailing and impairing resolution[115]. A variety of derivatives are available for analyzing carbohydrates[111,115,116]. Such derivatives as acetylated and trifluoroacetylated methyl glycosides[117], methyl ethers[115], and pertrimethylsilyl ethers[118] are conveniently prepared and provide good resolution. However since the anomeric centre and the ring are conserved in their preparation, each compound gives rise to at least four peaks in the chromatogram due to the formation of anomers in the pyranose and

furanose forms. The occurrence of multiple peaks in the chromatogram leads to reduced resolution and complicates the analysis. The sugars may be converted into oximes prior to trimethylsilylation[119] but both *syn* and *anti* derivatives are formed, so that each compound still gives rise to two peaks in the chromatogram.

Acyclic derivatives are preferable since they generally give only one peak per compound. Alditol acetates[120], peracetylated aldononitriles[121] (PAANs), pertrimethylsilylated alditols[122], diethyldithioacetals[123], and trimethylsilylated or acetylated deoxy(methoxyamino) alditols[124] have been used, with alditol acetates and PAANs being the most popular. Alditol acetates are usually prepared by reduction with sodium borohydride followed by acetylation catalysed by pyridine[125] or sodium acetate[89]. These derivatives have been used successfully for the analysis of aldoses and Figure 3.1 shows the separation of a mixture of monosaccharide alditol acetates. However, different sugars (eg. arabinose and lyxose) may give the same alditols when reduced, resulting in incorrect identification of components. This problem may be overcome by marking C-1 with a chromatographically distinct group, such as a nitrile group, prior to acetylation, as in the case of peracetylated aldononitriles (PAANs). Furneaux reported that significant amounts of by-products were formed using the traditional method of preparation[126], but an improved procedure developed by Chen and McGinnis occurs without by-product formation[127]. Ketoses may be analyzed by the preparation of a related derivative type, the peracetylated ketoximes[128].

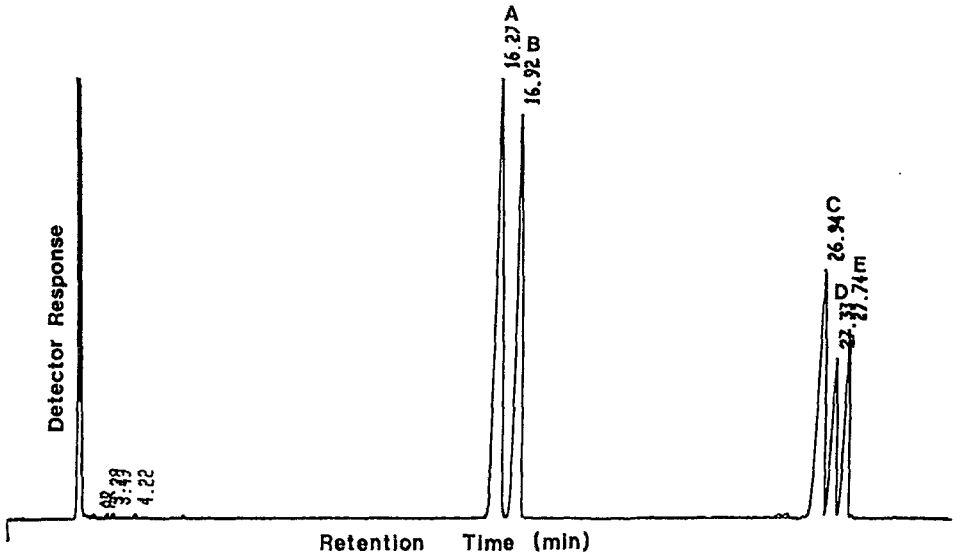


Figure 3.1 The GLC analysis of a mixture of monosaccharide alditol acetates on a DB-17 column. A. Rhamnose; B. Fucose; C. Mannose; D. Glucose; E. Galactose

The analysis of hexosamines can be problematic, since their alditol acetates and aldononitriles have excessively long retention times. A DB-17 column may be used for the analysis of the alditol acetates but the peaks are broadened and sensitivity is lower. *O*-methyl oximes[129] provide a useful alternative to alditol acetates since they have much shorter retention times and give a single peak in the chromatogram. More recently, acetylated deoxy(methoxyamino) alditols have been used[124].

GLC can also be applied to the analysis of methylated sugars and in the determination of monosaccharide configuration. These applications will be discussed in the relevant sections.

3.1.3 High performance liquid chromatography

High performance liquid chromatography (HPLC) is a form of column chromatography in which compounds are partitioned between a stationary phase composed of a monodisperse particulate solid and a mobile liquid phase which is pumped through the column at high pressure[130]. The column is constructed from stainless steel to withstand the high operating pressures of 1 to 30 MPa. The principle of separation may be based on molecular exclusion, ion exchange, ligand exchange, adsorption chromatography, or simply partitioning between the two phases[131]. The high operating pressure allows separation to be achieved in a relatively short time (15-60 minutes) which offers a distinct advantage over gravity-feed and low pressure column chromatography. The technique is versatile since it does not require high operating temperatures, and the wide range of available stationary phases allows analysis of a much wider range of compounds than is possible with GLC. Derivatisation is unnecessary and little or no pretreatment of the sample, beyond degassing and filtration, is thus required[132]. If proteins or high concentrations of salts are present these must be removed as they may cause damage to the column[132]. Many HPLC systems have a short "guard column" preceding the HPLC column to remove any such damaging substances and thus prolong the useful life of the column.

Stationary phases of several types are available. Those most useful for carbohydrate analysis include chemically bonded silica gels (normal and reverse-phase), amine-modified silica gels, and

both cation and anion exchange resins[133]. Reverse phase silica gels such as octyldimethylsilyl and octadecyldimethylsilyl columns are popular due to their stability, and since they can be used with aqueous mobile phases are suitable for routine analysis of carbohydrates. Resin-based columns are also generally used for the analysis of underivatised carbohydrates[132]. HPLC resins may contain cationic or anionic sites and compounds are separated by a combination of molecular size exclusion, ligand exchange, and hydrophilic adsorption. These columns are stable over a wide pH range and do not dissolve in aqueous media. Figure 3.2 shows the separation of a mixture of carbohydrates on the Supelcogel™ C6-11 ion-exchange column.

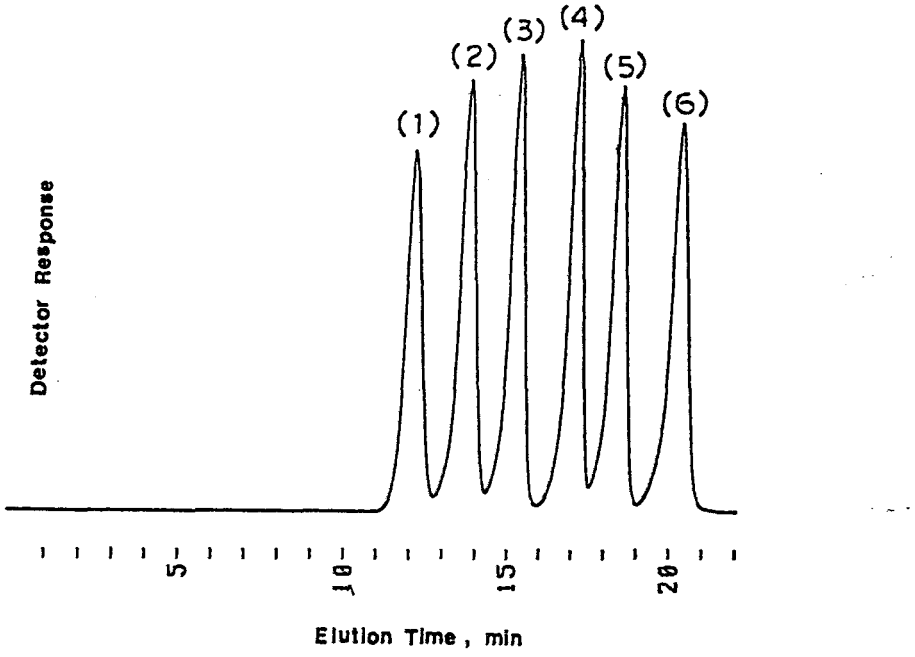


Figure 3.2 **The HPLC separation of a carbohydrate mixture on column Supelcogel C6-11. 1. Raffinose 2. Maltose 3. Glucose 4. Galactose 5. Mannose 6. Fructose Eluent 10^{-4} M aqueous NaOH**

Until recently, HPLC anion-exchange chromatography has been little used, since the silica and polystyrene-based resins available were not suitable for the analysis of underivatised neutral sugars, which could be separated only as their borate complexes with post column derivatisation for the purposes of fluorometric detection[134]. The technique requires long analysis times, is relatively insensitive, and resolution is poor. Recent advances in chromatographic technology have produced new high pH anion-exchange (HPAE) supports providing high-resolution separation of carbohydrate components[135]. Neutral, acidic, and amino sugars may be analyzed since the high pH of the mobile

phase converts neutral sugars into anions which are subsequently eluted using a sodium acetate gradient[136].

Gel filtration columns are used for the analytical or semi-preparative separation of oligosaccharide mixtures. Separation is based on molecular exclusion principles. The gel stationary phases are more fragile and less densely packed than are conventional stationary phases and operating flow rates and pressures are consequently much lower. In this laboratory, Progel™-TSK G3000 PWXL is employed for the analysis of both neutral and ionic samples and Progel™ G-oligo-PW for the analysis of non-ionic oligosaccharides. The mobile phase used for the latter column is usually water or 10^{-4} M NaOH, while the former employs a buffer system such as 0.1 M sodium acetate or 0.1 M sodium sulphate.

HPLC is not as sensitive a technique for carbohydrate analysis as is GLC due to the less sensitive detection systems employed[132] and for this reason is used as a complementary technique rather than a replacement. The most popular detector is the differential refractometer, despite its relative insensitivity. However, it is sensitive to changes in the temperature and composition of the mobile phase, thus necessitating a thermostatic control unit and making gradient elution impossible. The UV detector is insensitive to changes in the solvent composition providing the solvent has no chromophore and is not sensitive to temperature changes; however, its sensitivity towards carbohydrates is low since these compounds have low UV activity. Both detectors are adequate for analytical work; however for quantitative purposes it is preferable to derivatise the sample to produce a chromophore or fluorophore which allows more reliable and sensitive detection by UV or fluorometric techniques[133].

The recent development of pulsed-amperometric detection (PAD) has revolutionized HPLC detection systems. This detection system involves surface-catalysed oxidations of various functional groups (eg. the hydroxyl, carboxyl, and amino groups in carbohydrates) at noble metal electrodes, generating an electrochemical response which can be measured[137]. Other electrochemical

techniques involving constant applied potentials usually result in fouling of the electrode surface, which results in poor sensitivity. This problem is overcome by PAD since the potential is pulsed rather than constant. The technique is extremely sensitive and quantification is better than 100 pmol[136]. PAD coupled with high-pH anion exchange chromatography (HPAE-PAD) has become the method of choice for the determination of carbohydrates in a variety of samples and several studies have recently appeared in the literature[135-139].

3.1.4 Paper Chromatography

Before the advent of GLC and HPLC, paper chromatography was used for the routine qualitative and quantitative analysis of carbohydrate samples. As an analytical technique it has been superseded by GLC and HPLC, but it is still used for the preparative separation of hydrolysed samples in order to isolate new or rare monosaccharides which cannot be identified by the more routine techniques due to the non-availability of standards. Paper chromatography was used successfully to isolate Fuc3NAc from the capsular antigen of *E. coli* K45 (See Section 4.2) in order to determine its absolute configuration.

Compounds are partitioned between the stationary phase, *viz.* the hydrated paper fibres, and the mobile phase which is a water miscible solvent or mixture of solvents[140]. Chromatography may be conducted in the ascending or descending mode and the resulting chromatogram must be chemically developed to visualize the components. An excellent survey of most of the materials, solvents, and detection methods used in paper chromatography has been presented by Churms[84], and Sherma has surveyed more recent developments in and uses of the technique[141]. The technique has the advantages of being suitable for the separation of all classes of monosaccharides (including uronic acids and amino sugars) and small oligosaccharides, derivatisation of the sugars is unnecessary, and excellent sample recovery is achieved. Unfortunately the technique is slow in comparison to GLC and HPLC, the average time required to obtain results being between 16 and 24 hours.

3.1.5 The role of NMR in the determination of monosaccharide composition

Developments in the field of NMR spectroscopy over the last two decades have completely altered the approach towards the structural elucidation of complex polysaccharides. The applications of the technique in this field are numerous, and a more detailed discussion of various experiments and the information they afford appears in section 3.6; however it is appropriate at this point to mention that monosaccharide composition may be determined directly by NMR spectroscopy without recourse to chemical degradation. The basic configuration of each sugar residue may be determined from the characteristic proton coupling pattern it displays. The value of the spin-spin coupling constant, *J*, depends on the size of the dihedral angle between the vicinal protons in the molecule[142]. The coupling constant is large (5-8 Hz) when the protons are transdiaxial, and small (1-3.5 Hz) when they are gauche. Values for the first four vicinal coupling constants (H-1 to H-4) are sufficient to establish the basic configuration of the sugar[143]. Table 3.2 lists these characteristic *J*-coupling patterns. Once the basic configuration of each monosaccharide has been established and the signals in both the ^{1}H and ^{13}C NMR spectra have been assigned to the atoms of the constituent monosaccharides, identification of the residues is possible by comparison of this information with published data[144]. This approach is invaluable in cases where chemical degradation does not achieve complete release of constituent sugars. The capsular polysaccharide produced by *E. coli* K48 contains a novel trideoxy sugar which could not be released by chemical means and was characterised entirely by NMR spectroscopy[55].

Table 3.2 CORRELATION OF ALL VICINAL COUPLING CONSTANT PATTERNS WITH THE STEREOCHEMISTRY OF ALDOPYRANOSYL SUGAR RESIDUES[143]

Vicinal coupling constants[a]				Aldopyranose residue stereochemistry[b]		Examples
$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	Configuration	Conformation	
L	L	L	L	*ß-gluco*	$^4C_1$	ß-D-Glc, ß-D-GlcNAc
L	L	L	S	*α-ido*	$^1C_4$	(*α*-D-Idose)[c]
L	L	S	L	Not possible		
L	L	S	S	*ß-galacto* *α-altro*	$^4C_1$ $^1C_4$	ß-D-Gal, ß-L-Fuc (*α*-D-Altrose)
L	S	L	L	Not possible		
L	S	L	S	Not possible		
L	S	S	L	*ß-allo*	$^4C_1$	(ß-D-Allose)
L	S	S	S	*ß-gulo*	$^4C_1$	(ß-D-Gulose)
S	L	L	L	*α-gluco*	$^4C_1$	*α*-D-Glc, *α*-D-Qui
S	L	L	S	*ß-ido*	$^1C_4$	(ß-D-Idose)
S	L	S	L	Not possible		
S	L	S	S	*α-galacto*	$^4C_1$	*α*-D-Gal, *α*-L-Fuc
S	S	L	L	*α-manno* *ß-manno*	$^4C_1$ $^4C_1$	*α*-D-Man, *α*-L-Rha ß-D-Man, ß-D-ManNAc
S	S	L	S	*α-gulo*	$^1C_4$	(*α*-D-Gulose)
S	S	S	L	*α-* or *ß-altro* *α-allo*	$^4C_1$ $^4C_1$	(*α-* or ß-D-Altrose) (*α*-D-Allose)
S	S	S	S	*α-* or *ß-ido* *α-gulo* *α-* or *ß-talo*	$^4C_1$ $^4C_1$ $^4C_1$	(*α-* or ß-D-Idose) (*α*-D-Gulose) (*α-* or ß-D-Talose)

[a]The symbols L and S represent large (> 5 Hz) and small (< 5 Hz) vicinal proton coupling constants, respectively.

[b]The absolute stereochemistry of a residue cannot usually be established by NMR. Stereochemical designations apply to the D as well as to the L form.

[c]Residues shown in parentheses are theoretically possible, but have not yet been encountered in Nature.

3.2 DETERMINATION OF MONOSACCHARIDE CONFIGURATION

Sugars occur in both the D and L forms in nature and the glycosidic linkage may have either the α or ß configuration. Several methods have been developed for the determination of these configurations, and include both chemical and spectroscopic methods.

3.2.1 Absolute Configuration

The enantiomeric D and L forms of a monosaccharide may be distinguished by measurement of the specific optical rotation[145]. In the case of a polysaccharide, this method necessitates hydrolysis to constituent monosaccharides and isolation of the pure components in sufficient quantity for the analysis. Hudson evaluated the contribution of various portions of an oligosaccharide to the overall optical rotation of the compound in solution. His rules of isorotation[146] may be extended to the determination of monosaccharide configurations in oligosaccharides and polysaccharides, provided that the monosaccharide composition, order and positions of linkage, and ring sizes have already been established[145]. Sugar configuration may also be determined by means of circular dichroism (c.d.) measurements performed on the derived alditol acetates[147]. Each sugar has a characteristic c.d. band at 213 nm and sugars can be identified by comparison of the determined value with published values. This method has the advantage that accurate measurements can be obtained on milligram quantities of material. Although it is not applicable to *meso*-alditols, these can be converted to chiral derivatives.

Enzymatic methods may also be used. Several enzymes react exclusively with recognized molecules. For example, specific D-glucose or D-galactose oxidases oxidize only D-Glc or D-Gal type sugars respectively[148]. The polysaccharide is hydrolysed prior to treatment with the specific enzyme, and the hydrolysate is then analyzed for the absence of the corresponding sugars by chromatographic methods. A major disadvantage of enzymatic methods is that specific enzymes are expensive and not always easily available.

All the above methods require hydrolysis of the polysaccharide and isolation of the constituent monosaccharides to provide definitive results. Although this is necessary when new or unusual sugars are encountered, most commonly occurring monosaccharides can be analyzed by means of GC or HPLC without the need for laborious preparative chromatography. In order to distinguish between the enantiomeric forms of each monosaccharide, either the enantiomers must be converted into diastereomers which are chromatographically distinct, or a chiral stationary phase must be employed. The use of chiral alcohols to form diastereomeric derivatives is well established[149,150] and the method described by Leontein *et al.*[149] is used routinely in our laboratory. Monosaccharide enantiomers are identified by comparison of retention times with those of authentic standards. The particular advantage of this method is that it is possible to obtain the chromatograms expected for both forms of a sugar when only one form is available, since the reaction of the D-form of a monosaccharide with a (-)-alcohol will give the same products as those obtained from the reaction of the L-form with the corresponding (+)-alcohol[149]. This is particularly useful in the case of the amino sugars, since D-forms are readily available but L-forms are not. Unfortunately, each enantiomer will give rise to four peaks, i.e. two furanosides and two pyranosides, resulting in crowded chromatograms particularly if the repeating unit of the polysaccharide is large. This problem can be overcome by the preparation of acyclic diastereomeric derivatives which give only one or two peaks per enantiomer, eg. acetylated or trimethylsilylated diastereomeric dithioacetals[151], (-)-menthyloxime pertrifluroacetates[152], and 1-(*N*-acetyl-*α*-methylbenzylamino)-1-deoxy-alditol acetates[153].

Trifluoroacetylated carbohydrate enantiomers can be analyzed conveniently on the recently developed pentylated cyclomaltohexaose stationary phase[154]. Each monosaccharide still gives rise to more than one peak and the problem of multiple-peak chromatograms still applies. Monosaccharides are again identified by comparison of retention times with those of authentic standards. It is therefore necessary that standards are prepared for all possible enantiomers, thus limiting this method to the analysis of those monosaccharides for which standards are easily available.

It is not possible to distinguish between the two enantiomeric forms of an isolated monosaccharide by means of NMR spectroscopy. However, ^{13}C chemical shifts can be affected not only by inductive effects but also by spatial proton-proton interactions, so that the relative stereochemistries of two linked monosaccharides can be determined from ^{13}C NMR data[155]. Shashkov *et al.* have developed a method whereby the absolute stereochemistry of one sugar constituent of a disaccharide may be assigned on the basis of ^{13}C chemical shift differences, providing that the absolute stereochemistry of the second monosaccharide is known[155]. This is especially useful for the analysis of polysaccharides which are resistant to hydrolysis. For example, the absolute stereochemistry of the new diacetamido sugar present in the capsular polysaccharide of *E. coli* K48 was assigned by NMR spectroscopy since it could not be released by hydrolysis[55]. Other workers have found that the absolute stereochemistry of constituent monosaccharide residues in oligo- or poly- saccharides can be determined by a combination of NMR data and potential energy calculations, providing that the absolute stereochemistry of the adjacent sugar is known[156].

3.2.2 Anomeric Configuration

The anomeric configurations of the monosaccharide constituents in a polysaccharide may be determined chemically or enzymatically, although these methods often produce inconclusive results.

Acetylated oligo- and poly- saccharides may be oxidised with chromium trioxide in acetic acid. Pyranosidic sugar residues in which the aglycone occupies an equatorial position, i.e. ß-residues, are oxidised to 5-hexulosonate residues and surviving sugars can thus be assigned the α-configuration[157]. However, α-D-hexose residues which adopt the $^1C_4$ configuration also have an anomeric bond in the equatorial position and results can thus be inconclusive[158].

Enzymatic determination is based on the cleavage of glycosidic bonds by hydrolases which are specific for a particular anomeric configuration. A wide variety of such enzymes is available and their application to the structural analysis of polysaccharides has been reviewed[159,160]. This

approach is not widely used due to the cost of the specific enzymes required, although it has been used with success in determining the anomeric configurations of various residues in the capsular polysaccharides of *Klebsiella* K29[161] and *Klebsiella* K15[162].

Anomeric configurations of all the constituent monosaccharides can be established readily from a one-dimensional ^{1}H NMR spectrum of the native polysaccharide. The anomeric protons resonate downfield from the remaining ring protons and their respective configurations may be determined from chemical shift, spin-spin coupling data or spin-lattice relaxation time[163]. This will be discussed in greater detail in Section 3.6.

3.3 LINKAGE ANALYSIS

Methylation analysis is the traditional method of determining the linkage positions of the constituent sugars of a polysaccharide. Although this information can be provided non-destructively and less tediously by NMR spectroscopy, methylation analysis is still used to provide confirmatory evidence for the polysaccharide structure and the technique has been reviewed extensively[86,164,165].

The principle of methylation analysis is simple. The free hydroxyl groups in the polymer are etherified (methylated) and the polymer is hydrolysed to its constituent partially methylated monosaccharides which are identified by GLC-MS. Hydroxyl groups which are not etherified mark the positions at which the residues were substituted in the original polysaccharide.

Before the advent of NMR spectroscopy, methylation analysis was considered to be the most important technique available for linkage analysis, and as such has been the subject of considerable research. Many methods are available for effecting methylation of the polysaccharide substrate, most of which are based on ionization of the free hydroxyl groups to form nucleophilic alkoxides which may then attack the electrophilic methylating agent[86]. The success of the reaction is dependent on achieving a sufficient degree of ionization, and on the solubility of the polysaccharide

in a suitable solvent. The most widely used method until recently was that originally developed by Hakomori[166] and modified by Sandford and Conrad[167], which employs the methylsulphinyl anion (dimsyl anion) as the base and methyl iodide as the alkylating agent. The dimsyl anion was originally generated by the reaction of sodium hydride with dimethylsulphoxide (DMSO), which is also used as the solvent for the reaction. This procedure is tedious and the dimsyl anion is susceptible to the presence of moisture which may inactivate the reagent. Commercial sodium hydride also contains impurities which interfere with GLC analysis of the reaction products[168]. Modified procedures which afford improvements with regard to ease of preparation and purity of product include the use of potassium *tert*-butoxide[168], potassium hydride[169], and butyl lithium[170]. Due to its polarity, DMSO is an ideal solvent for most polysaccharides, and complete methylation is usually achieved in a single step. Undermethylation is usually due to the insolubility of a particular polysaccharide in DMSO, which has been attributed to intra- and inter-molecular hydrogen bonding restricting the reactivity of the free hydroxyl groups[171]. This problem can be overcome by the use of 1,1,3,3-tetramethyl urea which relaxes the hydrogen bonding[171]. Alternatively, mixtures of DMSO and *N*-methylmorpholine *N*-oxide (MMNO) may be used to dissolve polysaccharides which are insoluble in DMSO alone[172], although Harris *et al.* observed lower methylation yields with this method[173].

Despite the popularity of the Hakomori procedure, it has certain limitations. Apart from the difficulties encountered with the preparation of the dimsyl anion, the procedure also results in the esterification of the carboxyl groups of uronic acid residues, which may then undergo ß-elimination under the strongly basic reaction conditions[174]. In addition, Zähringer and Rietschel have reported *C*-methylation at *N*-methylacetamido groups, the proportion of these additional products being dependent on the reagent used for the preparation of the dimsyl anion[175]. More recently Ciucanu and Kerek have proposed a new methylation procedure which involves the use of hydrides and hydroxides as bases in preference to the dimsyl anion, claiming higher methylation yields with cleaner resulting gas chromatograms[176]. They claim that undermethylation is due to the reaction of the dimsyl anion with methyl iodide and present an intensive study of alternative basic reagents,

showing that the best yields are obtained in the shortest reaction times using NaOH or KOH in DMSO[176]. Ciucanu and Luca also investigated optimal conditions for keeping ß-elimination with subsequent degradation to a minimum during the methylation of polysaccharides containing uronic acid residues[174]. This method is more convenient due to reagent availability and ease of preparation, however York *et al.* have reported significant oxidative degradation under these conditions[177]. A subsequent study performed by Needs and Selvendran showed that this could be overcome by treating the polysaccharide with sodium hydroxide before the addition of methyl iodide in a sequential manner[178], and it seems that this more convenient, rapid procedure is now the method of choice for polysaccharide methylation.

Neither the Hakomori procedure nor the procedure of Ciucanu and Kerek lends itself to repetitive methylation steps, since exposure of a partially methylated polymer to the strongly basic conditions employed in both methods will result in ß-elimination[86]. Methylation may be completed by treating the partially methylated material with methyl iodide and silver oxide in DMF, according to the method developed by Kuhn *et al.*[179]. The absence of the O-H stretching vibrations in the IR spectrum of the methylated polymer indicates complete methylation.

A particular disadvantage of methylation procedures which employ strongly basic conditions is the loss or migration of base-labile substituents such as *O*-acetyl groups, resulting in loss of structural information. A recent method proposed by Prehm, using methyl triflate in trimethyl phosphate in the presence of the weak base 2,6-di-(*tert*-butyl)pyridine, provides a mild acid-catalyzed methylation procedure which achieves complete methylation in a single step with retention of acyl groups[180]. The presence and location of *O*-acetyl groups can also be readily determined by NMR spectroscopy.

Once the polysaccharide has been exhaustively methylated, it is hydrolysed to its constituent partially methylated monosaccharides which may be derivatized and identified by GLC-MS. Alditol acetates are the preferred derivatives and a large data bank of information pertaining to retention

times and mass spectra is available[181-3]. One of the problems encountered in the analysis of small samples is the loss of material due to repeated transfers between reaction vessels during the various stages of the reaction and due to volatilization of methylated components during concentration of solutions under reduced pressure[173]. In order to minimize such losses, Harris *et al.* developed a one-tube procedure which is especially suited to the analysis of small samples[173]. A recent assessment of this widely-used method by Needs and Selvendran showed that the acid-catalysed acetylation method employed in the Harris procedure results in the underestimation of terminal units in branched polysaccharides due to oxidative degradation[184]. In addition, highly methylated alditols, aldoses and alditol acetates are prone to loss from evaporation and stored GLC samples decompose within two months due to the excess acetic acid present[184]. Despite the advantages of the Harris method, it seems that the more traditional base-catalysed acetylation procedure affords more accurate information.

Traditional methylation analysis has two important drawbacks. Apart from the laborious work-up procedure, the technique is unable to distinguish between permethylated 4-linked aldohexopyranoses and 5-linked aldohexofuranoses, which give the same product (*viz.* a 2,3,6-tri-*O*-methylhexose) after hydrolysis of the glycosidic linkage[185]. The same problem is encountered with 4- and 5-linked aldopentoses and 5- and 6-linked ketohexose residues. Gray and co-workers have developed an alternative to traditional methylation analysis, which involves reductive cleavage of the methylated polysaccharide with triethylsilane in the presence of either boron trifluoride etherate ($BF_3.Et_2O$) or trimethylsilyltrifluoromethanesulphonate (TMSOTf)[186]. This results in regiospecific reductive cleavage at the anomeric carbon atom to afford the corresponding cyclic anhydroalditols (see Figure 3.3), which can be analyzed by GLC-MS as their acetylated derivatives. This method allows determination of linkage positions and ring forms simultaneously. Problems were encountered with the analysis of carbohydrates containing 2-acetamido-2-deoxyhexopyranosyl residues, since these formed water-soluble derivatives which could not be accurately quantitated[187]. D'Ambra and Gray recently modified the reductive cleavage method to allow simultaneous analysis of these residues[187], and used the modified method successfully in the linkage analysis of the

O-antigen of *Pseudomonas aeruginosa* ATCC 33358[188].

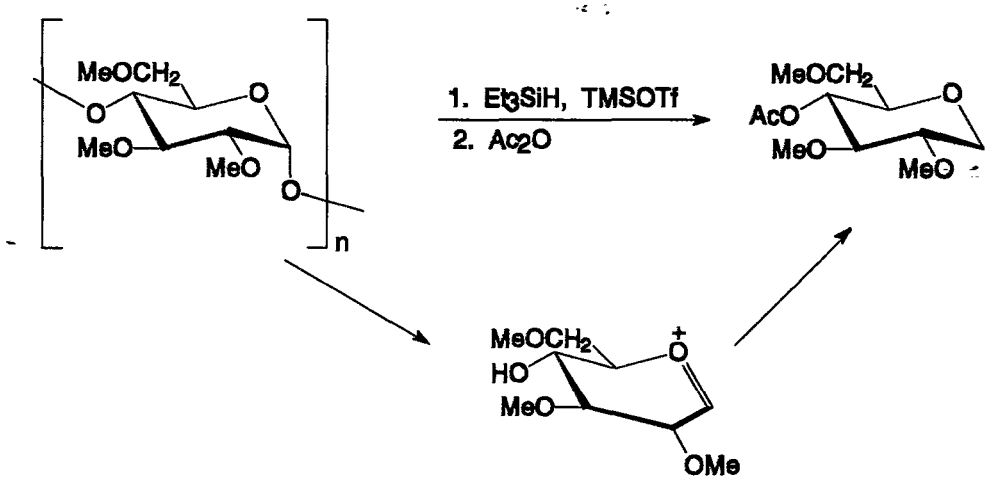


Figure 3.3 **Reaction scheme for reductive cleavage of a glycosidic linkage showing the presumed intermediate cyclic oxocarbenium ion**

3.4 ANALYSIS OF NON-CARBOHYDRATE SUBSTITUENTS

Many bacterial polysaccharides contain non-carbohydrate substituents which significantly influence their physicochemical and biological properties. *O*-acetyl groups affect the immunospecific character of the polysaccharide[56], while pyruvate substituents have been found to act as determinants in the symbiotic association between certain microorganisms and their host plants, besides having a significant effect on the conformation and solubility of the polysaccharide[189]. To obtain a complete understanding of the immunogenic character of a specific polysaccharide it is therefore necessary to identify and accurately locate such substituents.

O-acetyl groups are common and are the most difficult to locate due to their lability and tendency to migrate to other hydroxyl groups in the polysaccharide[190]. Their analysis may also be complicated if they are present in non-stoichiometric amounts. As was mentioned previously, the basic conditions of the Hakomori and the Ciucanu-Kerek methylation procedures cleave *O*-acetyl groups, and the milder Prehm[180] methylation procedure is necessary to locate their position in the polymer. Alternatively, the free hydroxyl groups in the polymer may be protected by conversion

to acetal[191] or phenylcarbamoyl[192] derivatives, after which the sites of the *O*-acetyl groups may be methylated selectively. More recently, Stanikowski and Zeller located the *O*-acetyl group in welan, a commercial polysaccharide produced by *Alcaligenes* ATCC 31555, using a reductive cleavage technique[193].

NMR spectroscopy provides more accurate information in less time and is the method of choice for the analysis of non-carbohydrate substituents. The presence of an acetyl group in the polysaccharide is indicated by a singlet at δ 2.0-2.3 in the ^{1}H NMR spectrum, and a peak at 20-25 ppm in the ^{13}C NMR spectrum. *O*-acetyl groups may be distinguished from *N*-acetyl groups by the characteristic C-N peak at 47-58 ppm in the ^{13}C spectrum. The substituent may be located from chemical shift data, since the methine proton of an acyl substituted carbon will be deshielded and the corresponding signal in the ^{1}H spectrum will occur downfield from other ring protons. Acylation also causes significant deshielding of the α-carbon (0.6-3.5 ppm) and shielding of the adjacent ß-carbons, supplying reliable information as to the position of the substituent[194]. Long range two- and three-bond heteronuclear correlations obtained from a 2D-Heteronuclear Multiple Bond Coherence (HMBC) experiment can locate an *O*-acetyl substituent unambiguously[195].

NMR spectroscopy is also the most widely-used method for the analysis and location of other non-carbohydrate substituents. Table 3.3 lists some commonly encountered substituents and their NMR data.

Pyruvate is usually acetalically linked at positions *O*-2 and *O*-3, positions *O*-3 and *O*-4 or, more commonly, at positions *O*-4 and *O*-6. The acetal linkages are base stable and the positions of substitution can therefore be established by conventional methylation analysis. Reductive cleavage of methylated pyruvate-containing polysaccharides using Me_3SiOSO_2Me-BF_3.Et_2O, BF_3-Et_2O or $Me_3SiOSO_2CF_3$ as catalyst results in the formation of partially methylated anhydroalditol derivatives which can be analyzed by GLC-MS to establish the positions of substitution of the pyruvate group[196,197]. Unfortunately, this technique also results in isomerisation at the acetal carbon and

therefore cannot be used to establish the absolute configuration at this chiral centre. In addition, pyruvate groups occurring as 2,3- or 3,4-linked acetals are more labile than the 4,6-linked form and may not be stable during chemical analysis, particularly if the polysaccharide is in its acidic form. Both the position of the pyruvate substituent and the absolute configuration of its acetalic carbon can be determined using NMR spectroscopy. Garegg *et al.* illustrated that it was possible to determine the absolute configuration of the acetalic carbon of a 4,6-linked pyruvate substituent from ^{1}H and ^{13}C NMR data, due to the difference between the ^{1}H and ^{13}C chemical shift data of axial and equatorial CH_3 groups[198]. Assignment of configuration to pyruvate substituents which are 2,3- or 3,4-linked is more difficult using this method because the chemical shift differences are smaller than in the 6-membered ring. It is also possible to assign absolute stereochemistry from NOE data as the pyruvate methyl protons may show NOE's to specific ring protons, depending on the acetalic configuration, as shown in Figure 3.4[199]. This approach has been used successfully to establish pyruvate stereochemistry in the capsular polysaccharides of *E. coli* K47[58] and K103 (this thesis) and *Klebsiella* K30[200]. Miertus *et al.* have noted that 4,6-linked pyruvate substituents prefer the *S* configuration when linked to D-glucopyranosyl or D-mannopyranosyl residues and the *R* configuration when linked to D-galactopyranosyl residues, and have correlated this preference with conformational stability by means of molecular dynamics studies[189].

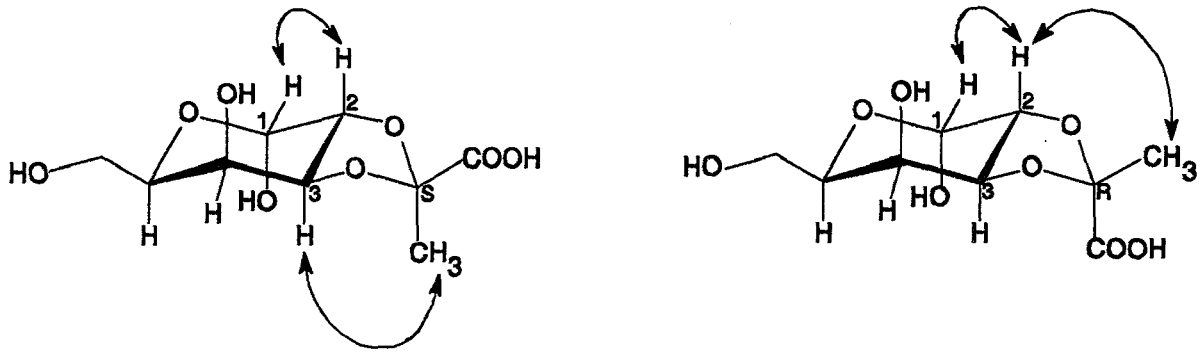
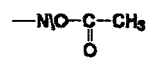
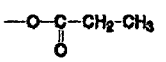
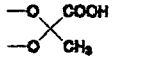
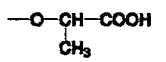
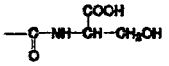
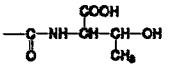


Figure 3.4 NOEs observed between specific ring protons and pyruvate methyl protons may permit assignment of the configuration of the quaternary acetalic carbon[199]

Table 3.3 **Non-carbohydrate substituents present in *E. coli* capsular polysaccharides**

Substituent	Structure		^{1}H and ^{13}C NMR Data (ppm)					
			-C=O	-CH_3	-CH_2-	-CH-	-C-	-COOH
N/O-acetyl		H		2.0-2.3				
		C	170-180	20-25				
O-propionyl		H		1.1-1.3	2.1-2.2			
		C	170-180	10-12	26-28			
Pyruvate		H		1.4-2.1				
		C		17-26			107-109, 100-103[a]	173-175
Lactyl		H		1.3-1.5		4.0-4.5		
		C		19-21		77-79		178-184
Serine		H			2.8-3.2[b]	~5.0		
		C	170-180		59-63	56-59		175-178
Threonine		H		1.0-2.0		4.1-4.5[c]		
		C	170-180	20-22		58-62, 68-71[d]		175-178

[a]For 5- and 6-membered acetal rings respectively; [b]H_a and H_b usually close together
[c]α and ß protons close together or may overlap; [d]H_α and H_β respectively

The amino acids L-serine and L-threonine have been found amidically linked to the carboxyl group of uronic acid residues[61-63]. These substituents can be identified and their linkage positions determined by NMR spectroscopy, since they show characteristic chemical shift and coupling constant patterns[201]. Strongly acidic conditions (eg. 4 M HCl, 100°C, 18 h, under N_2) are required to effect cleavage of the amino acids from the polysaccharide. The absolute stereochemistry may then be determined by GLC analysis of diastereomeric derivatives, or by measuring optical rotations after isolation and purification.

Phosphorus in the form of glycero, ribitol, or sugar phosphates is also a common component of *E. coli* capsular antigens. The phosphate content of the polysaccharide may be determined by means of colorimetric assays[202,203] following cleavage of the substituent from the polysaccharide. Total hydrolysis of phosphate-containing polysaccharides may be effected using 38% aqueous HF at elevated temperatures (65°C, 1 h), followed by extended TFA hydrolysis[135], while cold aqueous 48% HF may be used for dephosphorylation of the polysaccharide[204]. The position of phosphate linkage may be determined by ^{13}C NMR spectroscopy, since phospho-substituted carbon atoms

show characteristic deshielding and their corresponding signals are frequently split due to carbon-phosphorus coupling[205].

3.5 SELECTIVE DEGRADATION OF POLYSACCHARIDES

Once the linkage positions of the constituent monosaccharides have been established, the sequence of the monosaccharides in the repeating unit must be determined. It is possible to determine the monosaccharide sequence unambiguously by means of various NMR experiments, and these will be discussed in Section 3.6. The classical method of sequence determination involves degradation of the polysaccharide by chemical or enzymatic means to produce a specific oligosaccharide or mixture of overlapping oligosaccharides which can be isolated and characterised. NMR spectra of oligosaccharides are better defined and thus easier to interpret than those of polysaccharides, and in cases where the size of the repeat unit results in considerable signal overlap in the spectra of the native polysaccharide, the generation of a suitable oligosaccharide fragment greatly aids structure determination. Degradative procedures require a large amount of material since the yield of the desired product is usually low, and are therefore no longer used as part of routine structural analysis.

A wide variety of degradative techniques which take advantage of specific structural features within the polysaccharide have been established, and several comprehensive reviews have been published[75,86,206]. Only those techniques used in this thesis or in which new developments have been reported will be discussed.

3.5.1 Partial acid hydrolysis

Different glycosidic linkages are hydrolysed at different rates under acidic conditions. For example, furanosides are hydrolysed faster than the corresponding pyranosides by factors of 10-10^3, while glycosiduronic acids and 2-amino-2-deoxyglycosides are extremely resistant to acid hydrolysis[86]. It is therefore often possible to perform graded hydrolysis of the polysaccharide, employing the

same reagents used for total hydrolysis but under milder conditions, and to isolate the resulting oligosaccharide fragments by chromatographic techniques. This approach was used with great success in structural studies of the *E. coli* K37 polysaccharide[57], as shown in Figure 3.5. Partial acid hydrolysis also permits the isolation of labile monosaccharide residues without the degradation which usually accompanies total hydrolysis. This approach was particularly useful in the isolation of Fuc3NAc from the *E. coli* K45 capsular polysaccharide (see Section 4.2).

Non-aqueous acidic conditions such as those employed in acetolysis, trifluoroacetolysis, mercaptolysis and methanolysis may also be useful. In particular, acetolysis is frequently used in conjunction with partial aqueous acid hydrolysis to provide complementary data, since the relative rates of cleavage of the glycosidic linkages are often very different[86].

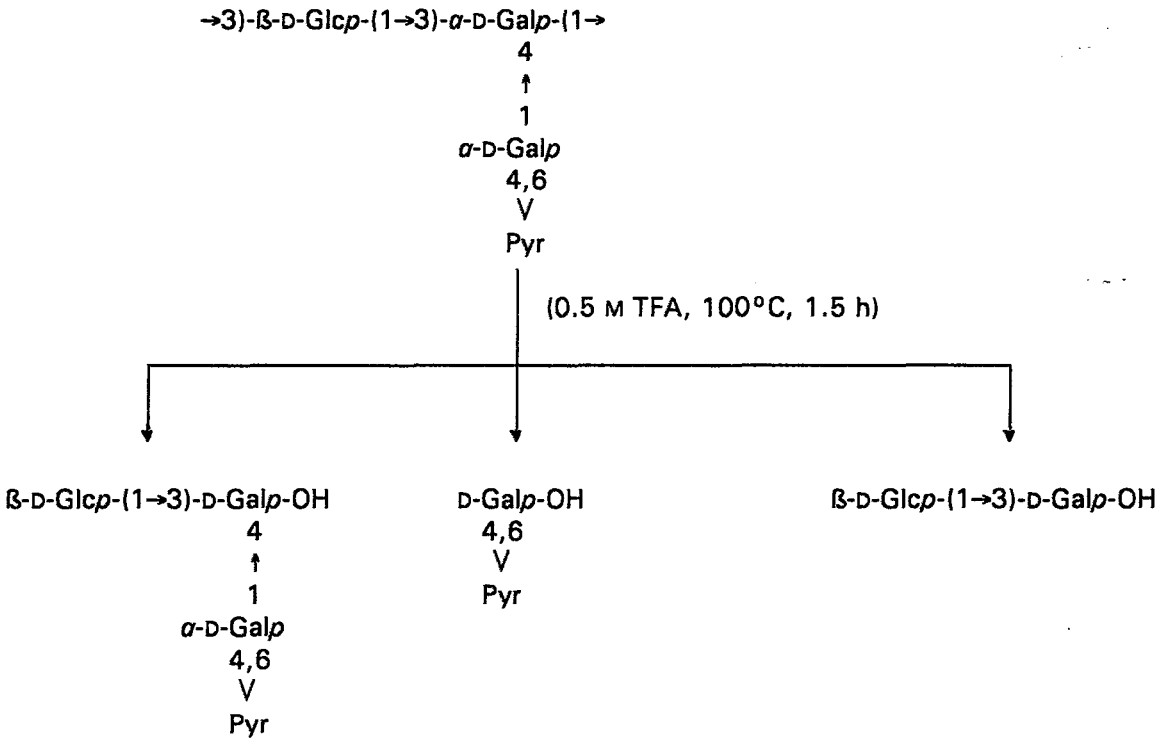


Figure 3.5 Products of partial acid hydrolysis of the *E. coli* K37 polysaccharide

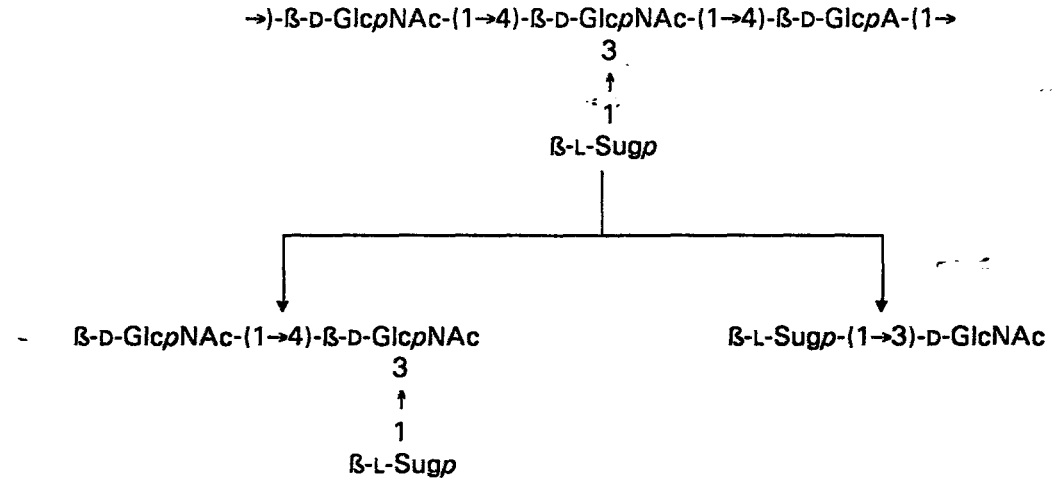
3.5.2 Anhydrous HF solvolysis

The use of liquid anhydrous HF as a medium to effect total hydrolysis of polysaccharides has been mentioned in Section 3.1.1. Despite its ability to cleave linkages normally resistant to acid hydrolysis, anhydrous HF may also be used as a highly selective reagent for partial degradation. Anhydrous HF solvolysis was first used in this capacity in structural studies performed on polysaccharides isolated from *Rhizobium japonicum*[207,208], and has since become a valuable technique for the selective solvolysis of polysaccharides.

A recent comprehensive review by Knirel *et al.* describes the mechanism of the technique and its application in the structural analysis of various polysaccharides[103]. Polysaccharides are dissolved due to the formation of hydrogen bonds between the HF and the polysaccharide, followed by cleavage of the glycosidic linkage to form glycosyl fluoride intermediates with very little accompanying decomposition. *N*-acyl substituents are preserved and *O*-acyl and pyruvate groups may be retained depending on the reaction conditions. Different glycosidic linkages exhibit different reactivities towards the reagent, and the solvolysis procedure may be controlled very selectively by the choice of reaction temperature. Table 3.4 lists the relative labilities of different glycosidic linkages[103].

The high selectivity of this reagent allows it to be used with equal efficacy for the analysis of sensitive polysaccharides containing labile linkages and those containing residues resistant to hydrolysis, such as amino sugars and uronic acids. For example, the technique was applied to the partial degradation of the hydrolysis-resistant capsular polysaccharide from *E. coli* K48 to generate oligosaccharides suitable for analysis by NMR spectroscopy, as shown in Figure 3.6[55].

Unfortunately, not even HF solvolysis could achieve release of the new diamino sugar 2,3-diacetamido-2,3,6-trideoxy-ß-L-mannopyranose to allow its complete characterisation[55].



where Sug is 2,3-diacetamido-2,3,6-trideoxy-ß-mannopyranose

Figure 3.6 **Partial degradation of the *E. coli* K48 capsular polysaccharide with anhydrous HF[55]**

Table 3.4 **Expected labilities of the glycosidic linkages of various sugar residues[103]**

	Temperature of hydrogen fluoride (°C)				
Sugar Residue	< -70	-40	-23 to -20	0	20 to 25
Pentofuranose	+	+	+	+	+
Pentopyranose	±	±	+	+	+
6-Deoxyhexose	-	±	+	+	+
α-Hexose	-	±	+	+	+
ß-Hexose	-	±	±	+	+
2-Amino-2,6-dideoxyhexose	-	-	-	±	+
3-Amino-3,6-dideoxyhexose			+	+	+
4-Amino-4,6-dideoxyhexose			+	+	+
2,4-Diamino-2,4,6-trideoxyhexose					+
2-Amino-2-deoxyhexose	-	-	-	±	±
Uronic acid	-	-	-	±	±
Galactosaminuronic acid	-	-	-	-	±
Mannosaminuronic acid	-	-	-		+
2,3-Diamino-2,3-dideoxyalduronic acids	-	-	-	-	-
5,7-Diamino-3,5,7,9-tetradeoxynonulosonic acid	-	-	-	-	-

Key: (+) cleavage; (-) stable; (±) depends on structural peculiarities.
The labilities indicated are generalizations as limited data are available.

3.5.3 Reductive cleavage

The reductive cleavage method was originally developed by Rolf and Gray as an alternative to traditional methylation analysis[185], as discussed in Section 3.3. Reductive cleavage of the glycosidic bond relies on successful hydride transfer from the silicon hydride catalyst to the activated anomeric carbon centre[209]. A number of catalysts have been used, including boron trifluoride etherate ($BF_3.Et_2O$), a mixture of $BF_3.Et_2O$ and TFA[209], and trimethylsilyltrifluoromethane sulphonate (TMSOTf)[210]. Both $BF_3.Et_2O$ and $BF_3.Et_2O$-TFA have been found to catalyse selective glycosidic cleavage, allowing the generation of oligosaccharides[209,210].

Selective reductive cleavage was used to generate specific oligosaccharides in structural studies performed on branched mannans from strains of *Saccharomyces cerevisiae*[210] and on the glucan pullulan[211]. In the latter study the catalyst $BF_3.Et_2O$ showed selectivity for α-(1→4) linkages while all α-(1→6) linkages remained intact. In 1987 Jun and Gray developed a new catalyst for reductive cleavage, *viz.* a mixture of trimethylsilyl methanesulphonate (Me_3SiOSO_2Me) and $BF_3.Et_2O$, which was found to be highly selective[212]. Fewer side reactions were reported using this catalyst.

Bennek *et al.* investigated the behaviour of amino sugars under reductive cleavage conditions, and found that ß-linked GlcNAc was selectively hydrolysed when using TMSOTf as catalyst[213]. The same study showed that α-linked GlcNAc and GalNAc are completely resistant to reductive cleavage, offering scope for the generation of specific oligosaccharides from polymers containing these residues. Vodonik and Gray found that 4-linked GlcA was cleaved rapidly under reductive cleavage conditions via an acyclic oxonium ion to form an isomeric furanosyl anhydroalditol, although other pyranuronic acids are cleaved at a lower rate than neutral sugars[214]. This selectivity was exploited in structural studies performed on capsular polysaccharides isolated from *E. coli* K57[215] and *E. coli* K38[216].

Zeller and Gray have shown that ester-linked non-carbohydrate substituents and pyruvic acid acetals are stable under reductive cleavage conditions, making it possible to determine their linkage

positions[196]. Using a selective catalyst, it is therefore possible to generate oligosaccharides which contain such substituents, thus aiding structure elucidation.

3.5.4 Degradation by Lithium in Ethylenediamine

Lithium metal dissolved in ethylenediamine was introduced as a reagent for the partial degradation of polysaccharides by Mort and Bauer[207]. A subsequent study by Albersheim *et al.* proved it to be a powerful reducing agent which selectively cleaves polysaccharides at the sites of uronic acid residues, regardless of the substitution position of the uronic acid or the anomeric configuration of the linkage point[217]. The proposed mechanism of this reaction involves the cleavage of the bond between the ring oxygen and C-5 of the uronic acid residue to form an intermediate radical anion with the oxy-anion located at C-1. The oxy-anion then initiates α-elimination with concomitant loss of the aglycone oxygen[217].

This degradation reaction has been used in the structural elucidation of several polysaccharides, including several recent studies on capsular polysaccharides from *Klebsiella* K43[218] and *E. coli* serotypes K26[219], K31[220], K39[221], and K102[222]. Figure 3.7 shows the lithium degradation of the capsular polysaccharide of *Klebsiella* K43[218].

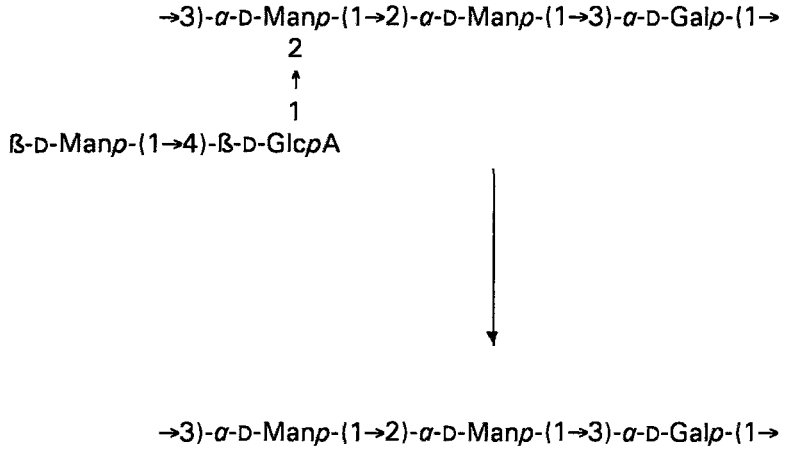


Figure 3.7 **Polymeric product of the lithium/ethylenediamine degradation of the *Klebsiella* K43 polysaccharide[218].**

More recently, Whittaker has explored the application of this degradative technique to polysaccharides containing amino sugars[223]. A modified procedure involving desalting on a Bio-Gel™ P2 column rather than a cation-exchange resin to remove lithium was developed using the capsular polysaccharide of *E. coli* K38, and was then extended to degradation of the *E. coli* K84 capsular polysaccharide. The reaction proceeded successfully but was found to require the addition of a much larger quantity of lithium wire and product yields were low.

3.5.5 Bacteriophage-mediated enzymatic degradation

Bacteriophages are bacteria-infecting viruses. Since their discovery in 1915[224], bacteriophages have been studied extensively and have been found to be highly substrate, and therefore host, specific[225]. Studies performed on the infective processes of bacteriophages revealed that penetration of the host cell was facilitated by degradation of the cell wall. This is achieved by means of enzymes believed to be secreted within the bacteriophage tail spikes[226]. Apart from lysozyme, which degrades the cell wall murein, bacteriophages may produce capsule-degrading enzymes, including deacetylases, lyases, and glycanases[227]. Most of these bacteriophages make use of endoglycanases which catalyse the hydrolysis of the glycosidic bonds in the capsular polysaccharide, thus producing the repeating unit and/or multiples thereof. Several instances of lyase (eliminase) activity have also been reported[228-230], and endo-ß-glucuronidase activity has been reported by Altman *et al.*, this being the only instance in which the resulting oligosaccharide has a terminal reducing uronic acid residue[231]. A study performed by Rieger-Hug and Stirm[227] on the activity of *Klebsiella* bacteriophages established the following:

(i) Most cleavages occur on either side of a negatively charged sugar unit, but reducing uronic acid residues are not usually produced,

(ii) The reducing sugar is often substituted at position 3,

(iii) ß-linkages are the more frequently hydrolysed glycosidic bond, and

(iv) Cleavage at sterically hindered sugar residues is favoured, especially for enzymes having lyase activity.

Bacteriophage-mediated enzymatic degradation thus provides a convenient and selective technique for the generation of oligosaccharides which supply all the necessary structural information about the polymer. The structures of numerous *Klebsiella* and *E. coli* capsular polysaccharides have been elucidated utilising bacteriophage-borne enzymes[227]. A particular advantage of such enzymic hydrolysis is that all labile substituents are preserved and bacteriophage degradation may thus be the only means of obtaining an oligosaccharide with such substituents intact. This approach was used in the structural elucidation of the capsular polysaccharides of *Klebsiella* K32[232] and *E. coli* K103 (this thesis), both of which contain extremely labile pyruvate substituents.

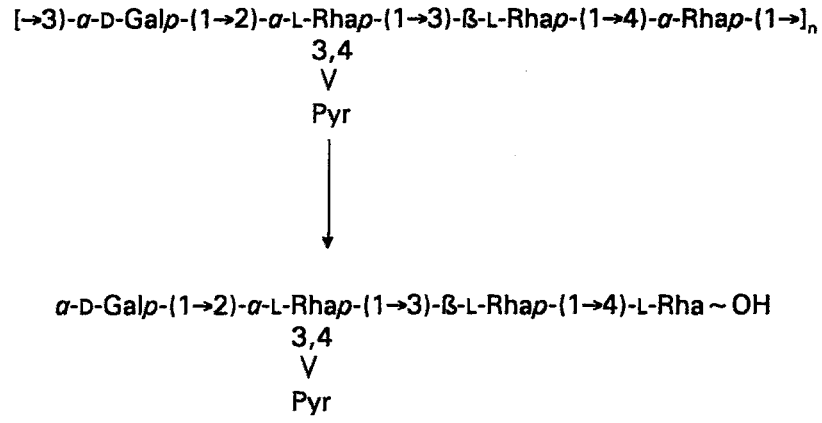


Figure 3.8. Bacteriophage degradation of the *Klebsiella* K32[232] capsular polysaccharide

Until recently the degradation of capsular polysaccharides by virally coded enzymes required rigorous purification of the bacteriophage particles using, for example, polyethylene glycol sedimentation followed by isopycnic centrifugation[233]. However, for the purposes of polysaccharide depolymerisation less rigorous purification is necessary and the simplified procedures proposed by Dutton *et al.* are preferred[234]. These will be described in greater detail in the experimental section (see Section 4.4).

3.6 NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

Developments in nuclear magnetic resonance (NMR) spectroscopy over the last few decades have revolutionized the study of natural products. It is now possible, by the judicious choice of one- and two- dimensional experiments, to determine completely the structures of complex molecules using only NMR techniques. NMR spectroscopy is a non-destructive technique and structural determinations may be carried out on less than 1 mg of sample.

Polysaccharides which are composed of regular repeating units, such as most bacterial polysaccharides, lend themselves readily to analysis by NMR spectroscopy. A spectrum of such a compound is essentially that of a single repeating unit, regardless of the degree of polymerization of the compound, although increasing molecular mass results in line broadening due to increased viscosity of the polysaccharide sample. Sample viscosity can be lowered in order to reduce line broadening by increasing the temperature of the probe; however care must be taken that labile substituents are not lost at elevated temperatures. NMR experiments are usually performed on a solution of the polysaccharide in high grade D_2O, the deuterium of which replaces exchangeable protons in the compound so as to simplify the spectrum. Acetone (chemical shift positions δ 2.23 for ^1H and 31.07 ppm for ^{13}C, relative to TMS) is usually included as an internal reference as tetramethylsilane (TMS) is insoluble in D_2O.

A discussion of the theory behind modern NMR techniques and of the sophisticated instrumentation used is beyond the scope of this thesis. The literature available on the subject is vast, and several excellent reviews have been published[235-239]. Some of the experiments commonly used in the structural elucidation of complex polysaccharides will be discussed.

3.6.1 One-dimensional techniques

Modern pulsed NMR involves the application of a sequence of radio frequency (rf) pulses to the sample in order to effect excitation of the nucleus or nuclei of interest. The resultant free induction decay (FID) signals in the time domain are added by a time-averaging computer and the data Fourier transformed to give the desired frequency domain spectrum[235]. In one-dimensional experiments excitation is restricted to a single type of nucleus, ^{1}H and ^{13}C being those most frequently observed in organic chemistry. ^{31}P NMR spectroscopy has also been developed due to the presence of phosphate in a large number of biological molecules. The positions of the signals in the one-dimensional spectrum are dependent on the chemical environment of the corresponding nucleus, and characteristic chemical shift values allow identification of molecular fragments. Characteristic coupling patterns which arise due to spin-spin coupling between neighbouring nuclei also provide a useful basis of identification.

3.6.1.1 One-dimensional ^{1}H NMR techniques

A typical 1D ^{1}H NMR spectrum of a bacterial polysaccharide is fairly complex, due to the degree of overlap of signals arising from the ring protons. A further complication is the presence of a water-derived HOD signal which may obscure other signals in the spectrum. Careful sample preparation with deuterium exchange is therefore necessary to ensure that the water peak is as small as possible. The position of the HOD signal is temperature dependent and it may be possible to place the signal in a region of the spectrum where no other peaks occur by altering the temperature of the probe[239]. Elevated temperatures are not always desirable since they may result in the loss of labile substituents, and it may be preferable to make use of specific pulse sequences which effect peak suppression[239,240].

Despite the complexity of the ^{1}H NMR spectrum a great deal of preliminary information concerning the polysaccharide can be obtained from it. The spectrum can be divided into three broad regions (see Figure 3.9), *viz.* the anomeric region (δ 4.5-5.5), the ring proton region (δ 3.2-4.5) and the methyl group region (δ 1.2-2.3)[239].

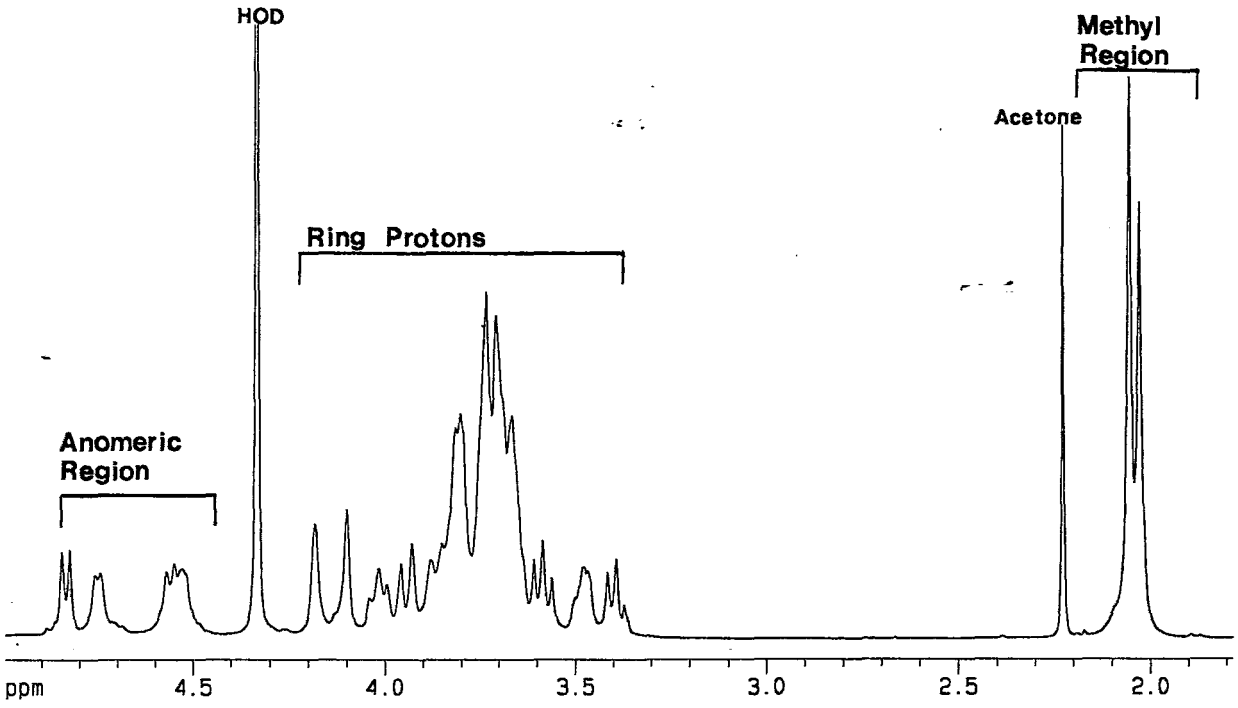


Figure 3.9 **The ^{1}H NMR spectrum (400 MHz) of the *E. coli* K101 capsular polysaccharide, showing the three spectral regions**

The anomeric protons of the constituent monosaccharides are significantly deshielded due to their proximity to the ring oxygen and consequently resonate downfield from the rest of the protons in the compound. The anomeric region yields information concerning the size of the repeat unit, the possible identity of the constituent monosaccharides and their anomeric configurations. The area under each anomeric signal is proportional to the number of protons represented, revealing the size of the repeat unit. The anomeric protons of α-linked *gluco* or *galacto* pyranose sugars resonate in the region δ 5.0-5.5, while ß-anomeric protons resonate at δ 4.5-5.0. The value of the spin-spin coupling constant J depends on the size of the dihedral angle between the vicinal protons[142]. The coupling constant is large (5-8 Hz) when the protons are transdiaxial and small (1-3.5 Hz) when they are gauche. The α-anomeric protons of *gluco* or *galacto* pyranose sugars have relatively small $^3J_{H\text{-}1,H\text{-}2}$ values whereas the ß-anomeric protons have large $^3J_{H\text{-}1,H\text{-}2}$ values. Chemical shift values and coupling constants thus provide information concerning anomeric configuration and permit identification of the basic configuration of the monosaccharide. Residues having the *manno*-configuration pose a problem as the signal for the anomeric proton may resonate close to δ 5.0 and the coupling constants are small for both α- and ß-linked sugars. In such cases the anomeric

configuration can be assigned from $J_{C\text{-}1,H\text{-}1}$ values or from NOE data. The anomeric protons of furanose sugar residues resonate further downfield than their pyranose counterparts, and are therefore easily identified[241].

In certain cases non-anomeric signals may also be found in the region δ 4.5-5.5, leading to misinterpretation of the data. In particular, H-5 of α-GalA, H-2 of ManNAc, and the α-protons of *O*-acylated carbon atoms may resonate in this region. Reducing sugars give rise to two signals in the anomeric region representing the α and ß forms of the sugar and other signals may also be "twinned" as a result of the proximity of their representative protons to the reducing end of the oligosaccharide. In such a case, the areas under the reducing α and ß-anomeric signals represent the ratio of anomers present in solution and should together add up to unity.

The ring proton region of the spectrum (δ 3.2-4.5) yields little direct information due to signal overlap. The methyl proton region (δ 1.2-2.3) reveals the presence of the methyl groups of 6-deoxy sugars, *O*- and *N*-acetyl substituents and pyruvate or lactyl groups. The analysis of non-carbohydrate substituents by NMR was discussed in Section 3.4.

More information can be obtained from a 1D ^{1}H NMR spectrum if selective excitation techniques are employed. A variety of such techniques is available, one of the most useful for carbohydrate analysis being the 1D version of the 2D Homonuclear Hartmann Hahn (HOHAHA) experiment[242]. The 1D HOHAHA experiment involves the use of a 180° shaped pulse to selectively invert a single isolated resonance, with subsequent magnetization transfer within the coupled spin system by the HOHAHA mechanism. 1D subspectra are generated which show the individual spin systems of the protons directly or indirectly scalar coupled to the inverted resonance. For this technique to be effective the resonance selected for inversion must be well separated from other signals in the spectrum, a prerequisite fulfilled by the anomeric protons of the monosaccharide constituents. The subspectra obtained permit identification of the signals belonging to individual monosaccharides as well as the measurement of coupling constants from the ring proton region, which is excessively

overlapped in the normal 1D ^{1}H NMR spectrum. This approach was used to good effect in the study of the bacteriophage degraded product of the capsular polysaccharide from *Klebsiella* K15 (Figure 3.10)[223]. Most of the other commonly used 2D NMR experiments can be converted into 1D subspectra by means of similar shaped pulse sequences. 1D versions of the COSY[243], NOESY[244], and ROESY[245] experiments have been developed.

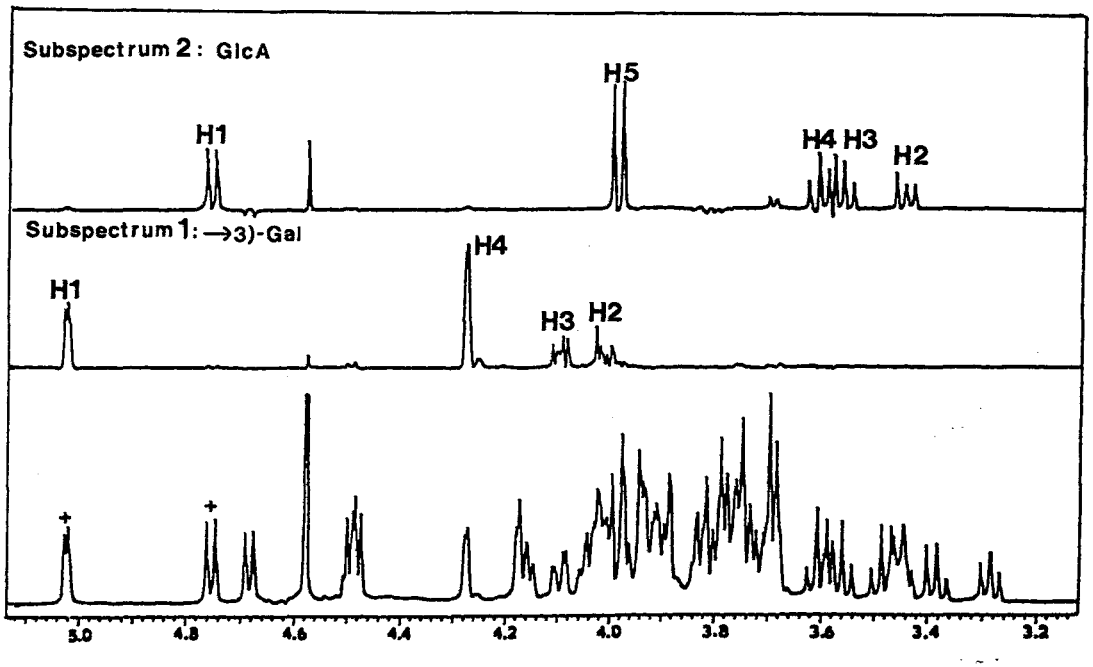


Figure 3.10 **Subspectra of some individual spin systems of the reduced hexasaccharide produced by bacteriophage degradation of the *Klebsiella* K15 capsular polysaccharide[223].**
+Denotes well resolved resonances selected for inversion to produce subspectra.

3.6.1.2 One-dimensional ^{13}C NMR techniques

Due to the low natural abundance of ^{13}C (1.1%), NMR experiments involving excitation of this nucleus require a greater number of scans to obtain a useful spectrum than are necessary in ^{1}H NMR spectroscopy. Although the development of modern Fourier transform techniques permits the routine acquisition of ^{13}C spectra using only the natural ^{13}C abundance of the compound under investigation, several techniques which effect sensitivity enhancement are also available. The removal of ^{13}C-^{1}H coupling simplifies the ^{13}C spectrum and also improves the sensitivity of the

nucleus through Nuclear Overhauser Enhancement (NOE)[235,240]. The intensity of each signal is increased in proportion to the ^{1}H splitting eliminated. ^{1}H-decoupled ^{13}C spectra are usually obtained with broad band or WALTZ ^{1}H-decoupling[246]. Sensitivity enhancement techniques involving polarization transfer such as selective population inversion (SPI)[247] or the non-selective INEPT (Insensitive Nuclei Enhancement by Polarization Transfer)[248] pulse sequence, for example, may also be used. In some instances coupled or partially coupled spectra are useful. These may be acquired by gated decoupling[249], which preserves most of the NOE enhancement, or off-resonance coherent decoupling[250]. Such spectra are useful for the measurement of $J_{C,H}$ values, which permit the anomeric configurations of *manno*-type residues to be established[251].

The information obtainable from the 1D ^{13}C NMR spectrum of a polysaccharide complements that from the ^{1}H NMR spectrum. The ^{13}C spectrum gives better signal separation due to the wider range of chemical shifts involved and the lack of spectral complexity permits ease of interpretation. The ^{13}C NMR spectrum may be divided into a number of regions which yield structural information (Figure 3.11).

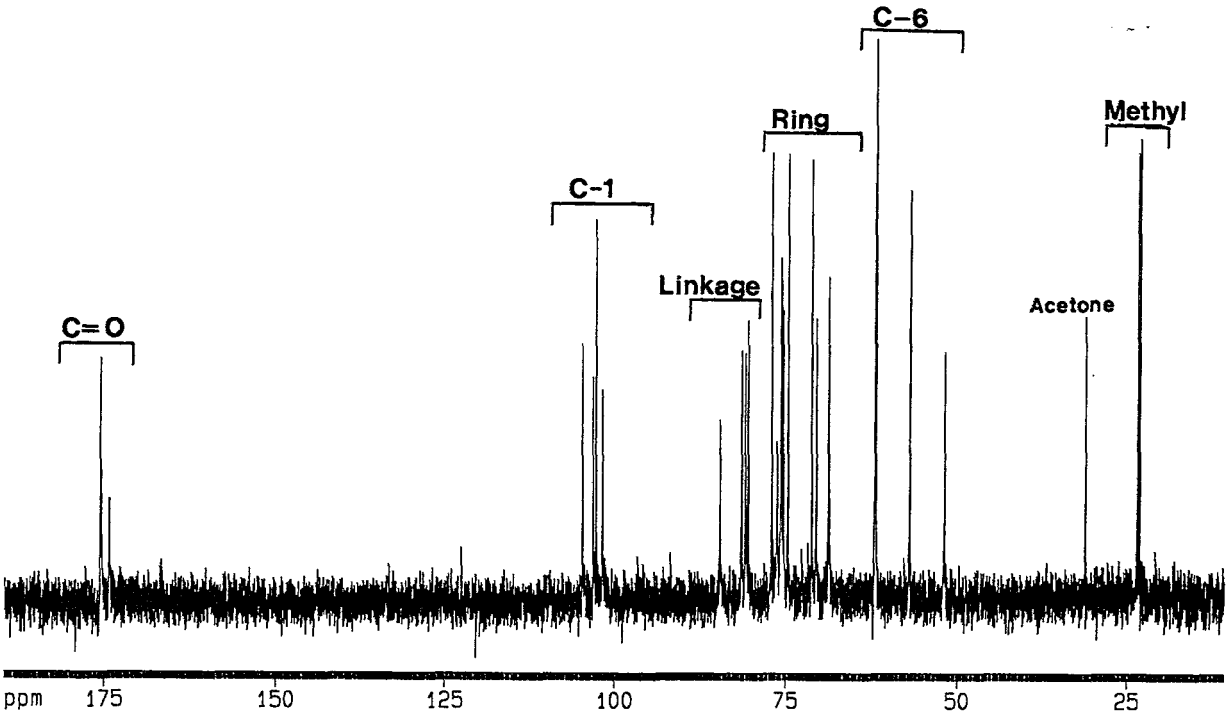


Figure 3.11 One-dimensional ^{13}C NMR spectrum of the *E. coli* K101 capsular polysaccharide

Carbonyl carbons resonate in the region 167-200 ppm and indicate the presence of uronic acids and acetyl or other carbonyl-containing substituents. The anomeric carbons resonate downfield from the rest of the ring carbons due to deshielding effects and can be found in the region 90-110 ppm. Again, the number of signals in this region is an indication of the size of the repeat unit. Carbon atoms which are linked to other sugars or non-carbohydrate substituents resonate in the region 75-85 ppm and this region affords an indication of the degree of branching in the polymer, as well as identifying positions of substitution. The remaining ring carbon resonances occur between 48 and 75 ppm, and include the methine resonances of 2-amino-2-deoxy sugars (48-58 ppm), the hydroxymethylene resonances for 6-linked (66-70 ppm) and unlinked hexoses (60-65 ppm), and ring carbons (67-75 ppm). Signals resonating between 17 and 24 ppm arise from methyl carbons and indicate the presence of deoxy-hexoses and pyruvate and acetate groups. It is possible to identify the monosaccharide constituents of a polysaccharide by comparison of ^{13}C chemical shift data, once assigned to the individual carbon atoms in the molecule, with published values, and a number of studies have been published[144,246,252-254].

The ^{13}C NMR spectrum also provides information pertaining to the anomeric configuration of the glycosidic bonds, since resonances of the anomeric carbons having the aglycon equatorial are generally downfield to those having an axial substituent[252]. Once again, the method cannot be extended to residues with the *manno*-configuration since the values of C-1_α and C-1_β are often indistinguishable. In addition, ambiguities may arise because glycosylation tends to shift the α-carbon downfield and the ß-carbon upfield[239]. Measurement of the $J_{C\text{-}1,H\text{-}1}$ coupling constants, already mentioned, provide a more reliable method of determination.

In addition to conventional one-dimensional ^{13}C NMR spectroscopy, a number of multiple pulse experiments which enable distinction between quaternary, methine, methylene and methyl carbons have been developed. Such techniques facilitate structural elucidation of complex molecules whose spectra contain areas of overlap. Of the many such experiments now available, those most frequently used are the APT (Attached Proton Test)[255], DEPT (Disortionless Enhancement by

Polarisation Transfer)[256] and INADEQUATE (Incredible Natural Abundance Double Quantum Transfer Experiment)[257] pulse sequences. The DEPT sequence relies on one-bond proton-carbon couplings to distinguish between the different types of proton-bearing carbon atoms. Quaternary carbon atoms, which have no bonded protons, are therefore not observed. The INADEQUATE pulse sequence relies on the creation of double quantum coherence between coupled ^{13}C nuclei and results in a spectrum free from uncoupled carbon signals.

3.6.2 Two-dimensional techniques

One-dimensional NMR methods yield limited information for the determination of the structure and stereochemistry of polysaccharides. Although the use of selective excitation to produce subspectra permits the identification of separate spin systems within the molecule, two-dimensional spectra are more efficient for the simultaneous determination of a large number of spin correlations. Developments over the last two decades have been numerous and there is now a wide variety of 2D experiments to choose from.

All 2D experiments involve the use of a multiple pulse sequence containing a variable delay t_1 between pulses in which free induction decays $S(t_2)$ are measured for an evenly spaced series of values of t_1 to build up a matrix of data $S(t_1,t_2)$. This matrix describes the signal as a function of both the time after excitation (t_2) and the delay of the pulse sequence (t_1). Fourier transforming the matrix twice produces a matrix $S(f_1,f_2)$ which is a spectrum in two independent frequency dimensions[258]. There are two basic types of 2D spectra[235]. *J*-resolved spectra have coupling constant information on one frequency axis and chemical shift data on the other, while correlated spectra have chemical shift data on both axes. *J*-resolved spectroscopy is not much used in carbohydrate chemistry because the overlapping of mutually coupled signals causes distortions of the multiplet patterns and interferes with the unambiguous measurement of coupling constants. Furthermore, coupling constant data alone does not lead to unambiguous identification of the monosaccharide constituents. Correlation spectroscopy relies on scalar coupling (homonuclear and heteronuclear), which involves coherent transfer of transverse magnetisation, or on dipolar

coupling, which is based on incoherent transfer of magnetisation. The 2D experiments most frequently used in structural studies on polysaccharides will be discussed briefly.

3.6.2.1 Homonuclear correlated techniques

The most commonly used 2D experiment is the ^{1}H-^{1}H homonuclear shift-correlated (COSY) experiment[259]. The basic pulse sequence involves the application of a 90° pulse to the sample, followed by a delay period (t_1), during which the spin system evolves as it would in normal free induction decay. A second 90° pulse, the mixing pulse, interrupts this evolution, followed by a second time period t_2, which allows evolution of the spin system to give a free induction decay signal which is detected and recorded. Double Fourier transform results in two 1D spectra which can be plotted at right angles to each other to give a COSY contour plot[258]. The diagonal of this plot represents the 1D spectrum. Scalar coupled protons will also show characteristic off-diagonal peaks (cross-peaks), depending on the size of their coupling constants. It is therefore possible to trace a spin system starting from, for example, the anomeric proton of a monosaccharide, by means of these cross-peaks. This is illustrated in Figure 3.12. The pulse angle may be altered from the basic 90° angle and the COSY45 is also frequently used[235].

One disadvantage of the COSY experiment is that cross-peaks can only be observed if the coupling constants between scalar-coupled protons are large enough. In particular, the connectivity pattern for ß-galactose residues can usually only be traced as far as H-4 since $J_{4,5}$ is less than 1 Hz. Similarly, the very small $J_{1,2}$ value for mannose residues results in a very small cross-peak between H-1 and H-2. The COSY experiment may be maximized for the detection of small coupling constants by employing an acquisition time which is much greater than the transverse relaxation time, and introducing fixed delays into the pulse sequence[240]. A further problem is encountered if a number of similar sugar residues are present, in which case there will be considerable overlap of cross-peaks, leading to ambiguous assignment of signals. The homonuclear relayed shift-correlated (RELAY-COSY) experiment[260] involves a second coherence transfer step, achieved by introducing a second mixing pulse into the pulse sequence, and shows correlation between non-

adjacent protons through a mutually coupled neighbouring proton (Figure 3.13). This is particularly useful for obtaining information about the chemical shifts of the non-anomeric protons because it resolves overlapping signals. The experiment can also be extended to enable correlations to indirectly coupled protons further along the chain by adding extra coherence transfer steps, eg. from H-1 to H-4 or H-5 in a two- or three-step RELAY-COSY experiment.

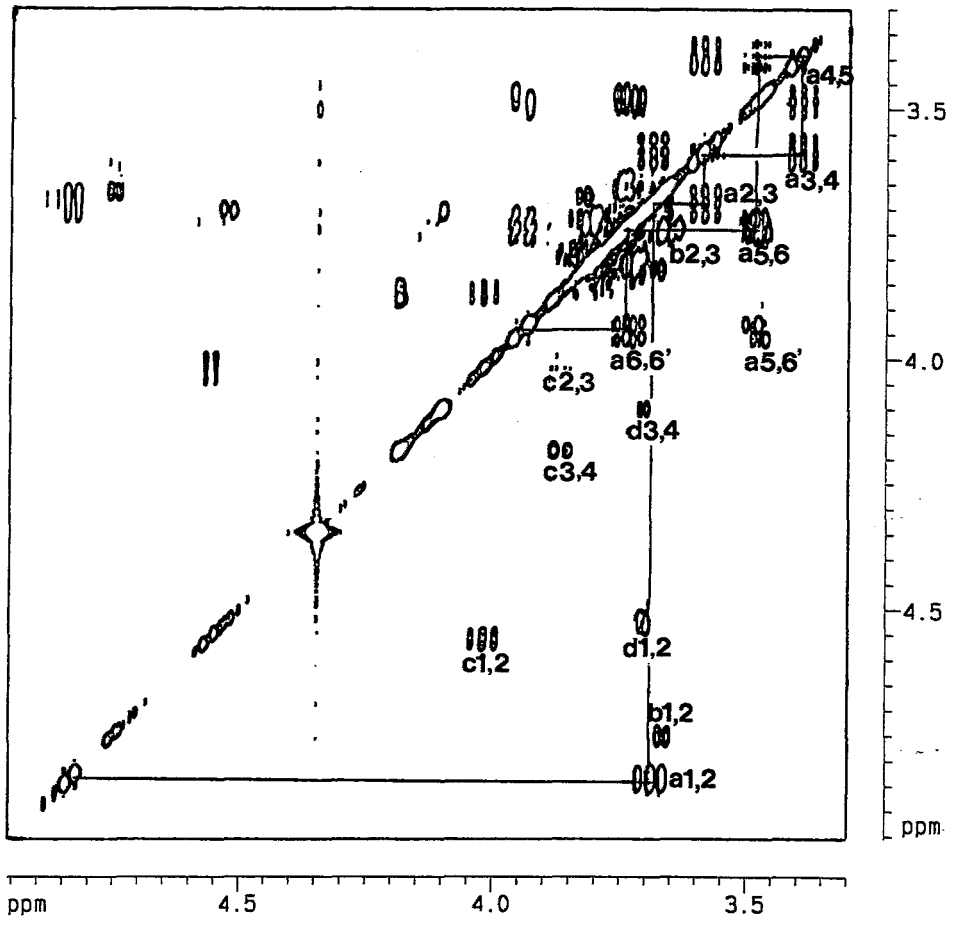


Figure 3.12 **COSY spectrum of the *E. coli* K101 capsular polysaccharide (region δ 3.30 - 4.95). The ^{1}H resonances of the *J*-coupled spin systems are labelled a - d**

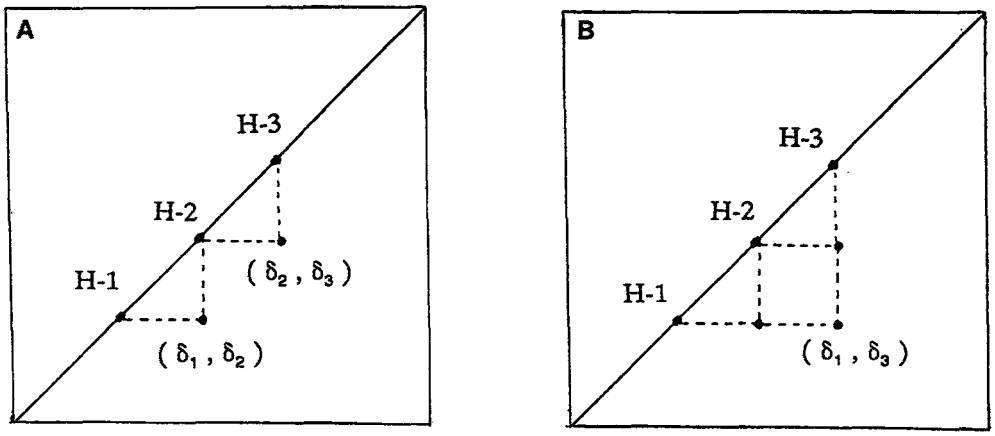


Figure 3.13 **Diagrammatic representation of a COSY (A) and a RELAY-COSY (B) experiment**

A number of phase-sensitive variations of the basic COSY experiment have been developed, and include phase-sensitive correlation spectroscopy (PS-COSY) and quantum filtered correlation spectroscopy (QF-COSY). The latter is particularly useful since it removes resonances that do not participate in coupling, such as residual solvent peaks[237]. The phase-sensitive COSY experiment is also very useful for the accurate measurement of chemical shifts and coupling constants from cross-peaks arising from a heavily overlapped 1D spectrum[258]. Such experiments usually take longer to perform because of the higher degree of resolution required to make the best use of phase information.

The 2D homonuclear Hartmann Hahn (HOHAHA) experiment, developed by Davis and Bax[261], provides a very effective means of separating out individual spin-systems. The experiment involves the transmission of magnetization throughout a scalar-coupled proton spin system by a polarisation transfer mechanism[262]. A complete spin system can thus be identified if there is at least one resonance in the spin system, such as the anomeric proton, which is well isolated and which has a reasonably large coupling to its neighbouring proton. This experiment does not suffer from the signal decay which occurs as a result of successive coherence transfer steps in the RELAY-COSY experiment, and is therefore frequently used to assign the signals in the ^{1}H NMR spectrum. As in the RELAY-COSY, magnetization transfer is dependent on the magnitude of the coupling constants, and if one of the couplings is small it will block the flow of magnetization. Inagaki *et al.* have recently described a relayed version of the HOHAHA experiment which overcomes this limitation to a large degree by including an additional conventional relay step in the pulse sequence[242]. The graphical representation of the 2D HOHAHA experiment is similar to that of a multiple-step RELAY-COSY.

Correlation spectroscopy experiments involving incoherent magnetization transfer by means of dipolar coupling between protons which lie in close proximity to each other are also possible. The 2D nuclear Overhauser (NOESY)[263] experiment makes use of this mechanism to detect NOEs between nearby protons (3 Å or closer) in the molecule. This is particularly useful in the study of

polysaccharides, since intra-residue NOEs provide information pertaining to the configuration of the monosaccharide, while inter-residue NOEs are primarily used for sequencing the residues and determination of linkage positions[238]. Information from the NOESY spectrum thus complements that obtained from other experiments. A NOESY experiment was used to establish the linkage sequence of the residues in the polysaccharide isolated from *E. coli* K43 (see Section 4.5.3). Figure 3.14 shows the expected intra- and inter-residue NOE contacts for a disaccharide.

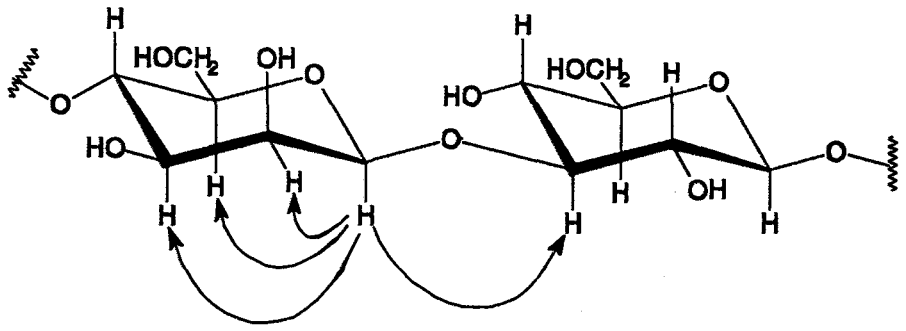


Figure 3.14 **Expected intra- and inter-residue NOE contacts for the disaccharide →4)-ß-Man*p*-(1→3)-ß-Gal*p*-(1→**

The NOESY experiment is not without its problems. The observance of an NOE peak is critically dependent on the mixing time chosen for the experiment. An expected peak may be absent either because the mixing time is too short, and the NOE has not had sufficient time to build up, or because the mixing time is too long, and the NOE has already dissipated[237]. Due care must therefore be taken when choosing the parameters for the experiment, and when interpreting the results. In addition, NOE effects are dependent on molecular tumbling rate and vary with molecular size. NOE peaks are frequently not observed for rapidly tumbling molecules of intermediate size, such as small oligosaccharides, when using the conventional NOESY experiment. This problem can be overcome by performing the NOE experiment in the rotating frame (ROESY), using an appropriate pulse sequence[264]. Both experiments still have the disadvantage that crosspeaks arising from scalar coupling may be present. This may be overcome by recording the spectrum in the phase-sensitive mode[237].

3.6.2.2 Heteronuclear correlated techniques

In the heteronuclear analogue of the COSY experiment, the two frequency axes represent the chemical shifts of different nuclei, typically ^{1}H and ^{13}C. The typical heteronuclear correlated experiment relies on transferring proton spin polarization through one-bond couplings, but uses a variable delay t_1 to make the magnitude and phase of the transferred signal depend on the proton chemical shift[258]. Since the magnetization originates as ^{1}H signals but is detected as ^{13}C signals, there is no diagonal peak equivalent to that observed in the COSY contour plot. This basic ^{13}C-detected heteronuclear correlation experiment, known as the HETCOR experiment, suffers from the disadvantage of low sensitivity due to the low natural abundance of ^{13}C. Extended experiment times are therefore required and sample sizes of more than 15 mg are necessary to obtain a useful spectrum.

The development of inverse-detected 2D NMR techniques have afforded a more sensitive alternative, in which heteronuclear chemical shift correlation is achieved via proton detection of heteronuclear multiple quantum coherence (HMQC)[266]. All variants of this experiment utilize the creation and manipulation of heteronuclear multiple quantum coherence followed by reconversion to detectable single quantum proton coherence. This HMQC pulse sequence requires smaller sample sizes and shorter experiment times and is preferred to its ^{13}C-detected counterpart in most cases. Both HETCOR and HMQC experiments result in a two-dimensional data matrix, the contour plot of which shows correlations between carbon and protons which are linked, providing a means of identifying proton resonances from their linked carbons and *vice versa*. Once all proton and carbon resonances have been assigned, it is possible to identify the sugar residues and to determine their linkage positions by comparing the chemical shift data obtained with published data[144].

Several relayed heteronuclear correlation experiments are available, using both ^{13}C- and ^{1}H-detected techniques. The latter are preferred because of their enhanced sensitivity. One of the most useful experiments of this type is the 2D HMQC-TOCSY (Figure 3.15) which utilizes an isotropic mixing interval after restoration of single quantum coherence to permit propagation of magnetization further through the system[267]. This experiment is valuable in instances where

chemical shifts cannot be assigned unambiguously from an HMQC spectrum. The ^{1}H-detected HMBC (heteronuclear multiple bond correlation) experiment[195] makes use of extensive phase-cycling to allow evolution of the magnetization components that lead to long-range multiple quantum coherence[266]. The contour plot of the 2D data matrix obtained shows two- and three- bond correlations between protons and carbons, and is an invaluable aid to determining the linkage positions of the monosaccharide constituents since magnetization is relayed across hetero-atoms such as oxygen and thus across the glycosidic linkage.

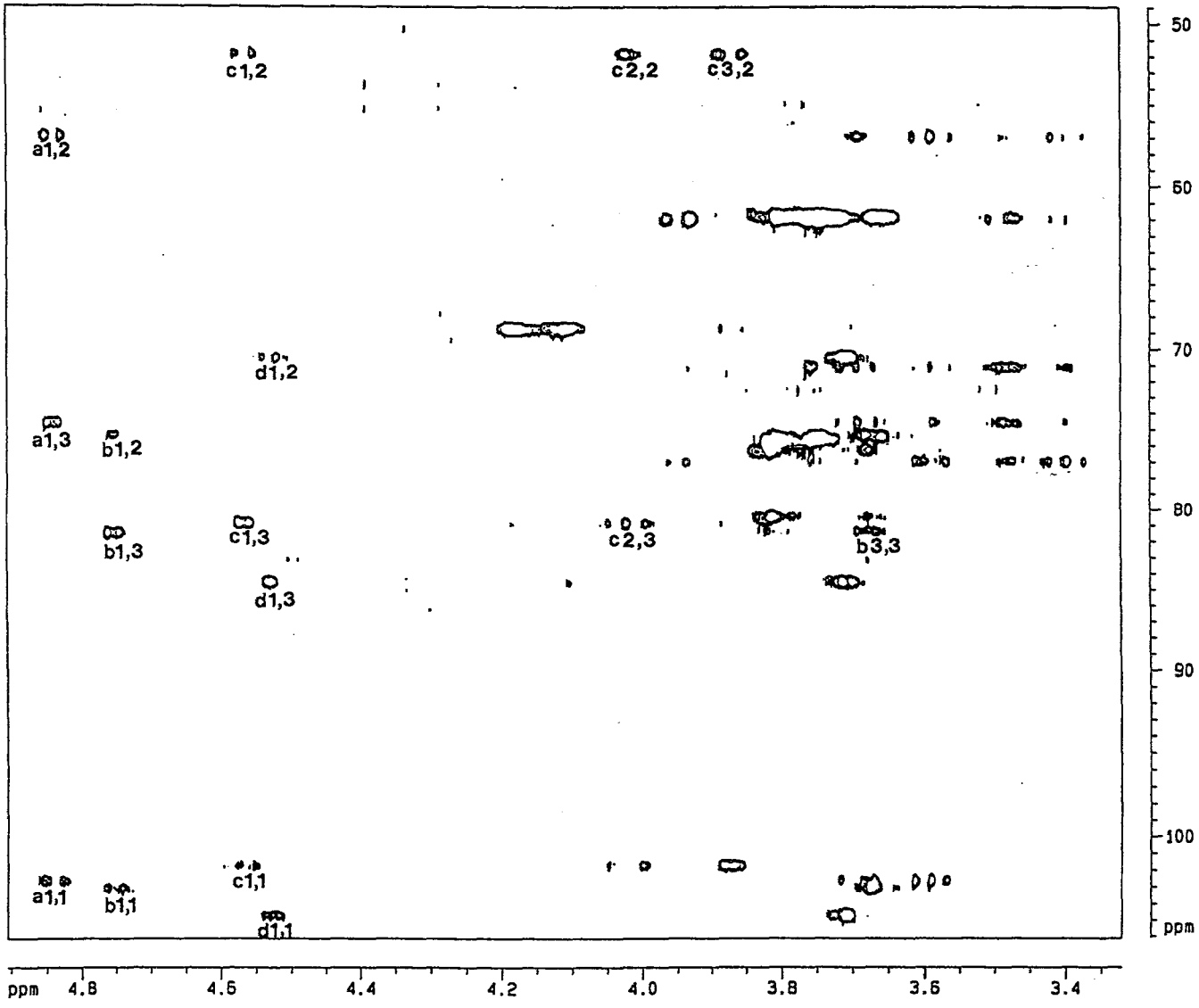


Figure 3.15 HMQC-TOCSY ^{1}H-^{13}C contour plot of the spectral region f_1 50 - 106 ppm (^{13}C) and f_2 δ 3.30 - 4.90 (^{1}H) of the *E. coli* K101 capsular polysaccharide. Correlated resonances are labelled a - d

3.7 MASS SPECTROMETRY

Mass spectrometry (MS) is a versatile analytical technique which is used extensively in the structural elucidation of organic compounds. Although not as definitive a technique as NMR spectroscopy, it has the advantage that microgram quantities of material can be analyzed, and the information yielded complements structural data obtained by other means. In addition the mass spectrometer may be interfaced with the gas chromatograph or the HPLC system[109], permitting the analysis of mixtures of compounds. The sample is volatilized in the inlet of the mass spectrometer (150-180°C), after which the vapours are ionized by a stream of high velocity particles. This results in the formation of a molecular ion, which is usually unstable and fragments to give smaller charged species. These fragments are separated via an electric or magnetic field based on their mass to charge ratios[268]. Most organic compounds exhibit predictable fragmentation patterns, and the fragment ions produced therefore provide information on the basic structural skeleton of the parent molecule. Unfortunately it is not possible to distinguish between stereoisomers since they fragment to give similar mass spectra.

Since polysaccharides are not volatile, it is necessary that they be degraded to their constituent monosaccharides, or at least to smaller oligosaccharides, for analysis by mass spectrometry[268]. Even oligo- and mono-saccharides are thermally unstable and require derivatisation to improve their volatility. The fragmentation patterns of various derivatives have been studied and the use of MS in carbohydrate structural analysis has been reviewed[268-271].

3.7.1 Electron-impact mass spectrometry

The most common method of sample ionization is by impact with high-velocity electrons. This usually results in the elimination of an electron from the sample molecule and the unstable positively-charged molecular ion thus produced fragments to form smaller charged ions. A molecular ion peak is seldom observed but the molecular mass may be determined by the addition of pairs of fragments. EIMS, in tandem with GLC, is a useful technique for the analysis of

monosaccharides. Distinctive fragmentation patterns permit distinction between pyranose and furanose sugar forms, between aldoses and ketoses, and between hexoses and pentoses, as well as enabling the positions of various substituents to be determined. The fragmentation pathways of a number of different derivatives have been studied[268-270], but the mass spectra of alditol acetates and partially methylated alditol acetates (PMAAs) are of particular interest because these derivatives are used most frequently in GLC analysis of carbohydrates.

Primary fragmentation of the PMAA occurs via fission of the alditol carbon chain. Cleavage occurs most readily between adjacent methoxyl functions, followed by cleavage between acetylated and methoxylated carbons, with cleavage between two acetylated carbon atoms being the least favoured[181,272]. Secondary fragments are formed by single or consecutive losses of acetic acid (60, m.u.), methanol (32 m.u.), ketene (42 m.u.) or formaldehyde (30 m.u.). Typical fragmentations are shown in Figure 3.16.

Recognition of functionalities such as *N*-acetyl groups of amino sugars, carboxyl groups of uronic acids, and deoxy sugars is possible due to the presence of characteristic fragments in their spectra, and the position of such substituents can also be determined[181,182]. A wide range of monosaccharides can thus be identified on the basis of their mass spectra, providing that these can be compared with the mass spectra of authentic samples. Collections of reference spectra of PMAAs have been published by Jansson *et al.*[181] and by Stellner *et al.*[182].

EIMS is not used routinely for the study of oligosaccharides as these compounds usually fragment to produce extremely complex spectra. EIMS of permethylated oligosaccharides can provide sequence information[86,269] but the results are generally unreliable because of the low relative abundance of structurally informative high-mass fragments and molecular ions. EIMS spectra of oligosaccharides containing more than four sugar residues are generally too complex to be useful.

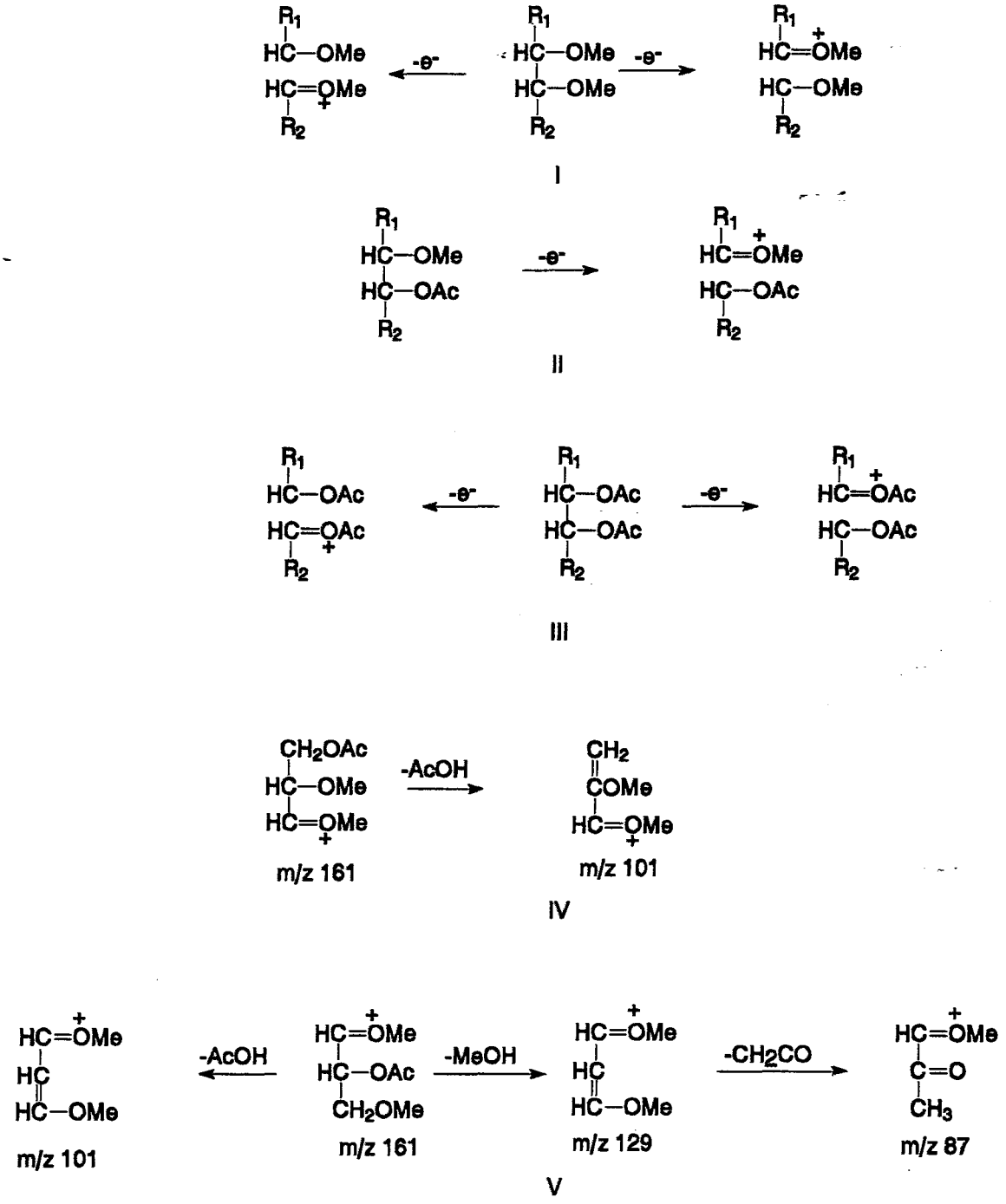


Figure 3.16 Primary fragmentations of a typical permethylated alditol acetate (pathways I, II and III), followed by typical secondary fragmentations (IV and V)

3.7.2 Chemical ionization mass spectrometry

The high energies used for ionization in EIMS frequently cause excessive fragmentation of the sample, resulting in complex spectra which are dominated by small fragment ions. EIMS spectra give little indication of molecular mass and structural detail at high mass is lost, complicating compound identification. Chemical ionization mass spectrometry (CIMS) is a "soft" ionization technique which results in the formation of larger ion fragments with consequently less complex spectra[273]. The essential feature of chemical ionization is the electron ionization of a reagent gas to generate an ion plasma, which is then able to ionize the sample[274]. The reagent gases employed most frequently are methane, isobutane, or ammonia which produce the corresponding reacting species CH_5^+, $C_4H_9^+$, and NH_4^+. Ionization of the sample occurs via charge or proton transfer, or the formation of collision-stabilized adduct ions to produce positively charged, high mass ions[274]. Characteristic fragmentation of these ions produces the CI mass spectrum, which is far simpler than the EI mass spectrum, and in most cases an intact molecular ion enables determination of the molecular mass. The information obtained from CIMS is thus complementary to that obtained from EIMS and the two techniques are often used together in structural elucidation.

The prevalence of high mass ions and the simplicity of the spectra obtained make CIMS a useful technique for the investigation of higher molecular mass compounds such as oligosaccharides. For example, CIMS was used to determine the linkage sequence of residues in the capsular polysaccharide isolated from *E. coli* K38[218] (Figure 3.17). As oligosaccharides often undergo undesirable pyrolysis in conventional CIMS, it may be advantageous to use a modification of the basic CI technique. Direct chemical ionization (DCI) involves the placement of the sample directly within the ionization chamber where heat induced desorption occurs from an extended probe directly into the ion plasma[274]. Samples analyzed by this direct method frequently provide mass spectra with enhanced high mass fragments. DCI-MS of oligosaccharides, using ammonia as the reagent gas, has been shown to preferentially cleave glycosidic linkages, resulting in the production of high molecular mass fragments from which the monosaccharide sequence can be determined[274]. This approach was used successfully in studies on a tetrasaccharide isolated by means of bacteriophage degradation of the *E. coli* K44 polysaccharide (Figure 3.18)[275].

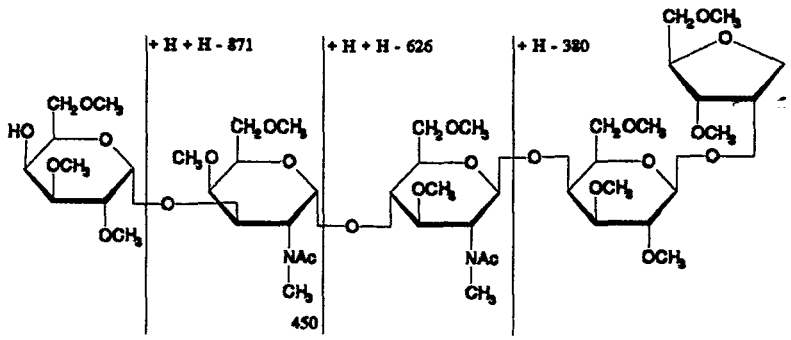
ß-D-Gal*p*-(1→3)-ß-D-Gal*p*NAc-(1→4)-*α*-D-Glc*p*NAc-(1→4)-*α*-D-Gal*p*A-(1→2)-anhydroribitol [PM]

4

↑

OH

M+1 = m/z 1075



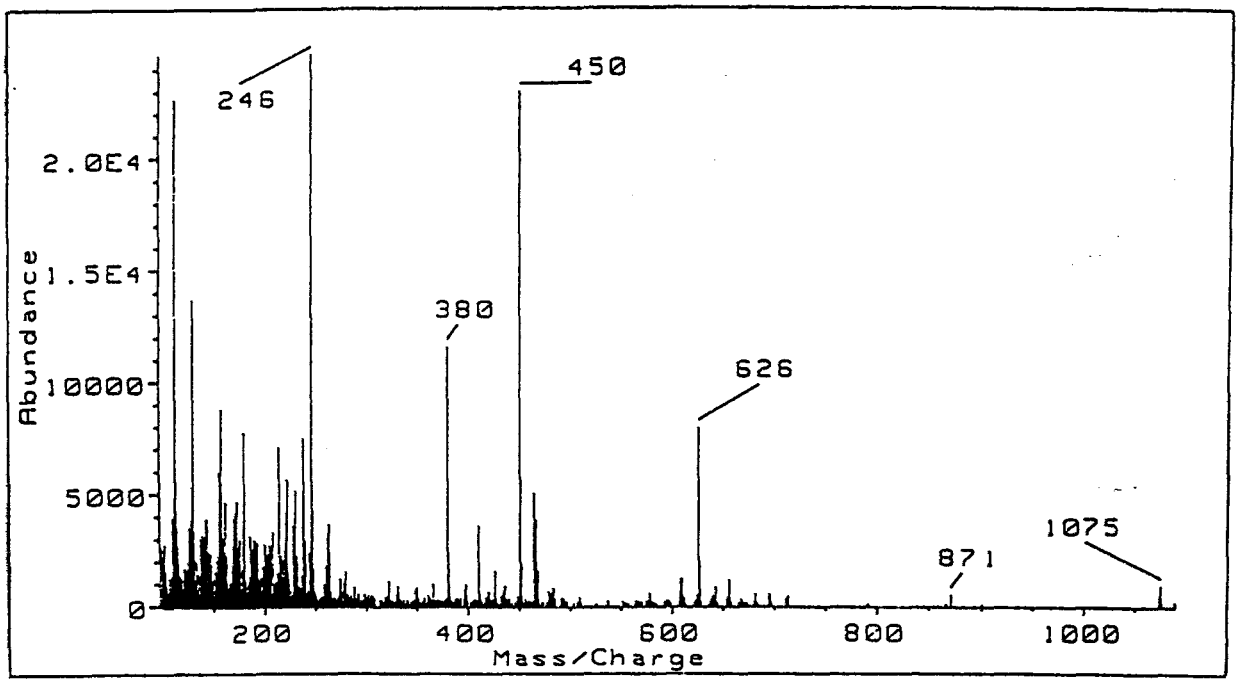


Figure 3.17 CIMS spectrum of the anhydroribitol terminated pentasaccharide anhydroalditol product obtained following specific reductive cleavage of the *E. coli* K38 polysaccharide[218].

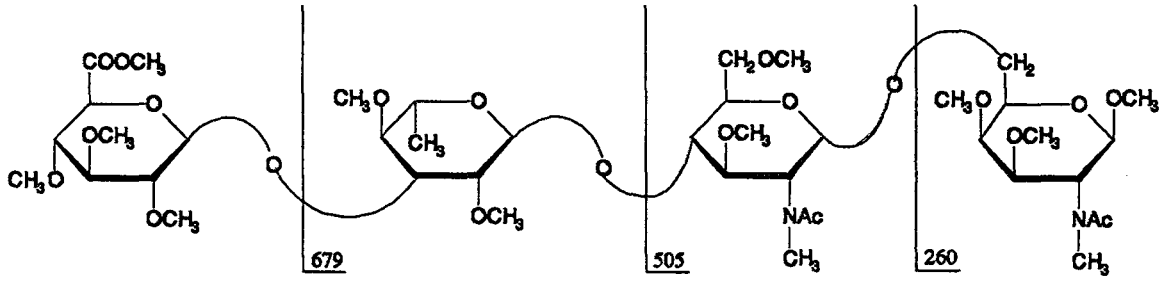


Figure 3.18 Methylated tetrasaccharide P1 derived by bacteriophage degradation of the *E. coli* K44 polysaccharide showing its fragmentation on DCI-MS[275].

3.7.3 Fast atom bombardment mass spectrometry

Fast atom bombardment mass spectrometry (FAB-MS) is another "soft" ionization technique which provides valuable information concerning the molecular masses of large oligosaccharides and small polysaccharides[271]. The sample is ionized directly from solution in a matrix (usually glycerol) by bombardment with accelerated atoms or ions, such as xenon or argon. Since gas-phase ions are generated without prior volatilization of the sample, the technique results in less fragmentation than other MS techniques, and is thus ideal for the investigation of polar, non-volatile, higher molecular mass compounds.

Bombardment of the sample with accelerated atoms results in the formation of both positive and negative ions, and these spectra are recorded separately. The addition of certain salts, eg. NaCl, to the sample stabilizes the molecular species formed as the cationic capture ion, resulting in the appearance of a pseudo-molecular ion peak in the positive-mode spectrum. A peak representing the molecular anion is also readily visible in the negative-mode spectrum.

As in the case of DCI-MS, oligosaccharide fragmentation under FAB-MS conditions usually involves cleavage at the glycosidic linkages, providing a means of determining monosaccharide sequence as well as permitting identification of monosaccharide constituents. As with other MS techniques, the method does not distinguish between stereoisomers and the anomeric configuration of monosaccharides cannot be determined.

FAB-MS can also be applied to the structural study of both native and derivatized carbohydrates[271] and has been used to determine the composition and sequences of bacterial polysaccharides[215,216]. Pappas *et al.* have shown that it is possible to analyze mixtures of related oligosaccharides by FAB-MS, since each structure gives rise to a definitive molecular ion (Figure 3.19)[276]. This is particularly useful in partial degradation studies where the derived oligosaccharides are difficult to separate.

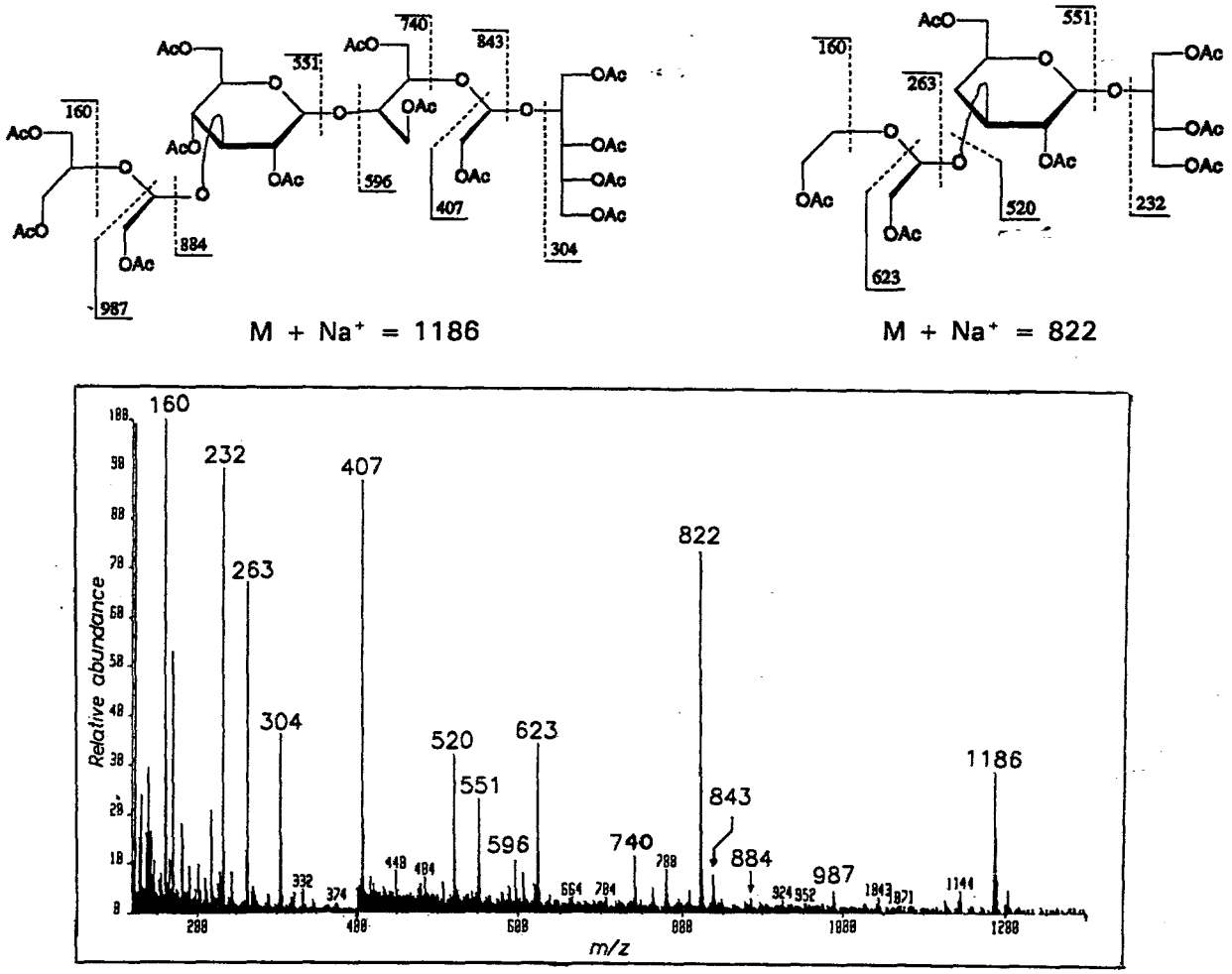


Figure 3.19 **FAB-MS spectrum of derivatives of nigeran tetrasaccharide and a related trisaccharide resulting from Smith degradation[278].**

FAB-MS is also an ideal technique for analyzing polar substances containing labile functional groups because derivatisation and/or volatilization of the sample, which may result in the loss of such substituents, are not required. The analysis of polysaccharides containing *O*-acetyl groups may be problematic due to the tendency of these substituents to migrate to other hydroxyl groups in the polysaccharide[180]. Although this problem can be overcome by the use of NMR spectroscopy, the NMR spectrum is complicated if the *O*-acetyl groups are present in non-stoichiometric amounts, as is often the case. The data obtained in a FAB-MS experiment provides precise molecular mass information, so that it is possible to define which substituent groups are present as well as their exact number[277]. FAB-MS has been used to locate many different types of labile substituents, including pyruvate, formyl, propionyl, and acetyl groups[277,278].

The coupling of the HPLC system to the FAB mass spectrometer is particular advantageous for the analysis of biological compounds which are sensitive to elevated temperatures and to pH. It is frequently not possible to subject such compounds to derivatisation procedures or rigorous purification, particularly if labile substituents are present in the molecule. Separation by HPLC followed by analysis by FAB-MS is the analytical method of choice in such cases. Santikarn *et al.* made use of tandem HPLC-FAB-MS in structural studies on oligosaccharides and reported enhanced sensitivity using a moving belt interface[279].

More recently, Lam *et al.* have reported the use of laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry[280] for the analysis of bacterial oligosaccharides[281,282]. This hybrid technique provides sequencing information and permits identification of linkage positions. Spectra obtained in the negative-ion mode can be used to distinguish the anomeric configuration of the constituent monosaccharides and the presence and position of labile substituents can also be confirmed[282].

CHAPTER FOUR: STRUCTURAL STUDIES ON SOME *Escherichia coli* CAPSULAR ANTIGENS

4.1 STRUCTURAL ELUCIDATION OF THE CAPSULAR POLYSACCHARIDE OF *Escherichia coli* SEROTYPE O20:K101:H⁻

4.1.1 Abstract

The primary structure of the acidic capsular antigen of *Escherichia coli* O20:K101:H⁻ was shown by glycose analysis, methylation analysis, and one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy to be composed of repeating branched tetrasaccharide units having the structure:

→3)-ß-D-Gal*p*NAc-(1→4)-ß-D-Glc*p*A-(1→3)-ß-D-Gal*p*-(1→
3
↑
1
ß-D-Glc*p*NAc

4.1.2 Introduction

The investigation of the structure of the *Escherichia coli* K101 capsular polysaccharide forms part of a continuing program of structural studies on *E. coli* exopolysaccharides. These studies have been undertaken in an attempt to relate chemical structure to antigenic character. Structures have been reported for sixty of the seventy-four known *E. coli* capsular antigens.

The great majority of the capsular (K) antigens of *E. coli* are acidic polysaccharides and have been subdivided into two groups on the basis of their physical, chemical and microbiological characteristics[64]. Group I polysaccharides are usually co-expressed with O groups 8,9,20, and 101. They are high molecular mass, acidic polysaccharides, are heat stable at pH 5-6, and may contain amino sugars. The capsular polysaccharide of *E. coli* O20:K101:H⁻ belongs to the subgroup of I whose capsules do contain amino sugars.

The bacterial strain O20:K101:H⁻ (culture number 1473) has been found to cause diarrhoeal disease in new-born, colostrum-deprived piglets and may cause enterotoxigenic diarrhoea in humans[44].

4.1.3 Results and Discussion

Isolation, composition and 1D NMR spectra of the polysaccharide.— *E. coli* K101 bacteria were grown on Mueller-Hinton agar, and the acidic capsular polysaccharide (**PS**) was isolated and purified by precipitation with cetyltrimethylammonium bromide (CTAB), followed by delipidation with aqueous 1% acetic acid. The released lipid was removed by ultracentrifugation. A small sample was further purified for NMR spectroscopy by gel-permeation chromatography (GPC) on Sephacryl™ S500. The **PS** showed a broad distribution of molecular masses with an average M_r at 1.5 x 10^5. Hydrolysis of the **PS** followed by GLC-MS analysis of the derived alditol acetates showed that Gal, GalN, GlcN, and Glc were present. Methanolysis and reduction of the **PS** with $NaBH_4$, followed by hydrolysis, reduction, and acetylation, revealed no additional monosaccharides, but the proportion of Glc present increased, indicating the presence of GlcA. A further hydrolysis of the polysaccharide was performed, followed by reduction with $NaBD_4$, and the derived alditol acetates were examined by GLC-MS. The Glc component exhibited mass fragments one and two mass units higher than the corresponding Gal component, indicating that the Glc was in fact all derived from the reduction of GlcA and that there was no unsubstituted Glc present in the polysaccharide. Partial reduction of GlcA to Glc occurs via the formation of the glucurono-3,6-lactone during hydrolysis[94]. The constituent sugars were shown to have the D configuration by GLC analysis of the derived acetylated (-)-2-octyl glycosides[148].

The ^{1}H NMR spectrum of the **PS** (Figure 4.1) in D_2O contained H-1 signals at δ 4.85 ($^3J_{1,2}$ 8.2 Hz), 4.75 ($^3J_{1,2}$ 6.1 Hz), 4.57 ($^3J_{1,2}$ 8.0 Hz), and 4.52 (broad doublet), and two signals for the methyl protons of two NAc groups at δ 2.03 and 2.06 (3H each). The chemical shifts of the H-1 signals indicated that all linkages were ß. The ^{13}C NMR data complemented the ^{1}H NMR results and confirmed the tetrasaccharide repeating unit for the **PS**, with signals for C-1 at 104.77, 103.17, 102.69, and 102.16 ppm, and signals for carbonyl carbons at 175.53, 175.25, and 172.07 ppm. Signals at 56.92 and 51.86 ppm indicated the presence of two C-N bonds, and confirmed the presence of two NAc groups. In addition, C-6 resonances were observed at 62.01, 61.92, and 61.88 ppm, indicating that none of the residues were 6-linked.

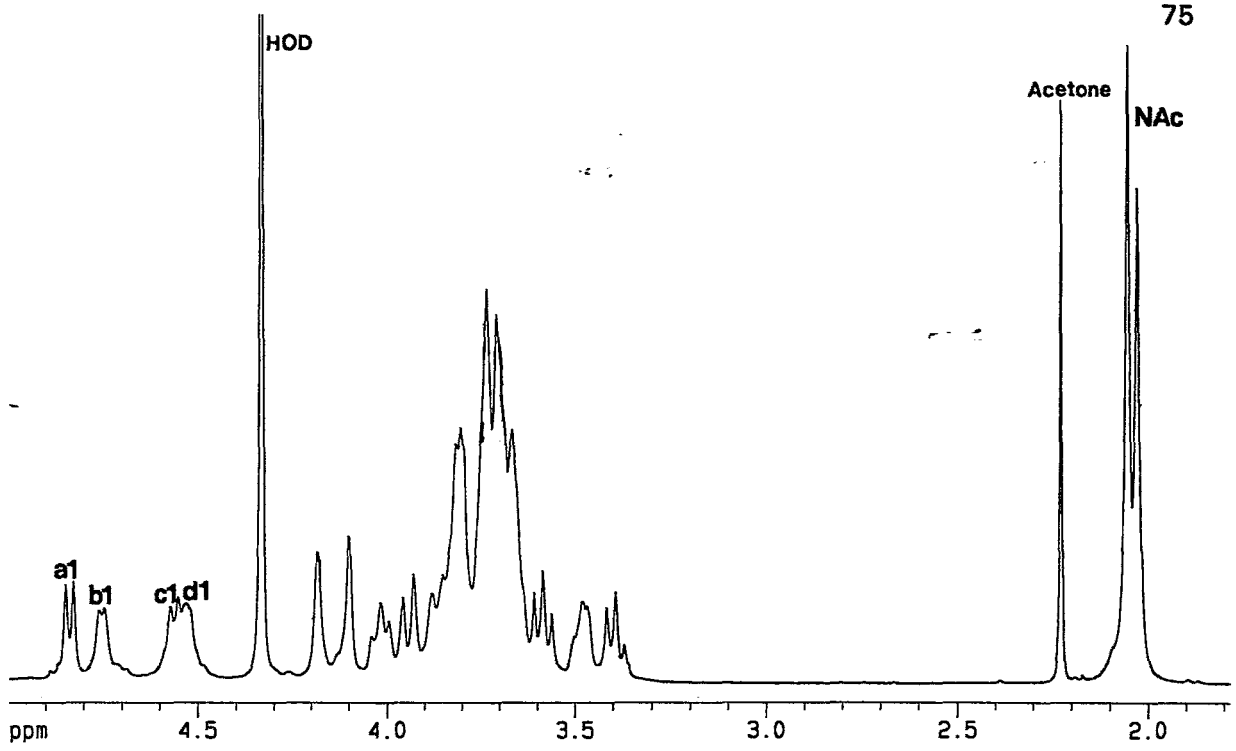


Figure 4.1 **^{1}H NMR spectrum of the K101 PS (400 MHz, 340K)**

Methylation analysis.— The **PS** was methylated by the modified Hakomori procedure of Phillips and Fraser[189] using potassium dimsyl and methyl iodide. GLC and GLC-MS analysis of the derived partially methylated alditol acetates (with carboxyl reduction) showed the presence of 2,4,6-tri-*O*-methylgalactose, 2-deoxy-2-methylacetamido-4,6-di-*O*-methylglucose, 2-deoxy-2-methylacetamido-3,4,6-tri-*O*-methylglucose, and 2,6-di-*O*-methylglucose. These results accord with a branched tetrasaccharide repeating unit with GlcA as the branch point and GlcNAc as the side chain.

2D NMR studies of the E. coli *K101 polysaccharide.*— The sequence of the residues in the repeating unit was established by 2D NMR experiments, which also confirmed the glycosylation sites in the polysaccharide. The residues in the repeating unit were labelled **a** - **d** in order of decreasing chemical shift of their anomeric protons (Figure 4.1). Most of the proton resonances were established from COSY[258], one- and two-step relay COSY[260], and 2D Homonuclear Hartmann Hahn (HOHAHA)[261] experiments. The two-step relay COSY and HOHAHA contour plots are shown in Figures 4.2 and 4.3 respectively, and the ^{1}H assignments are presented in Table 4.1.

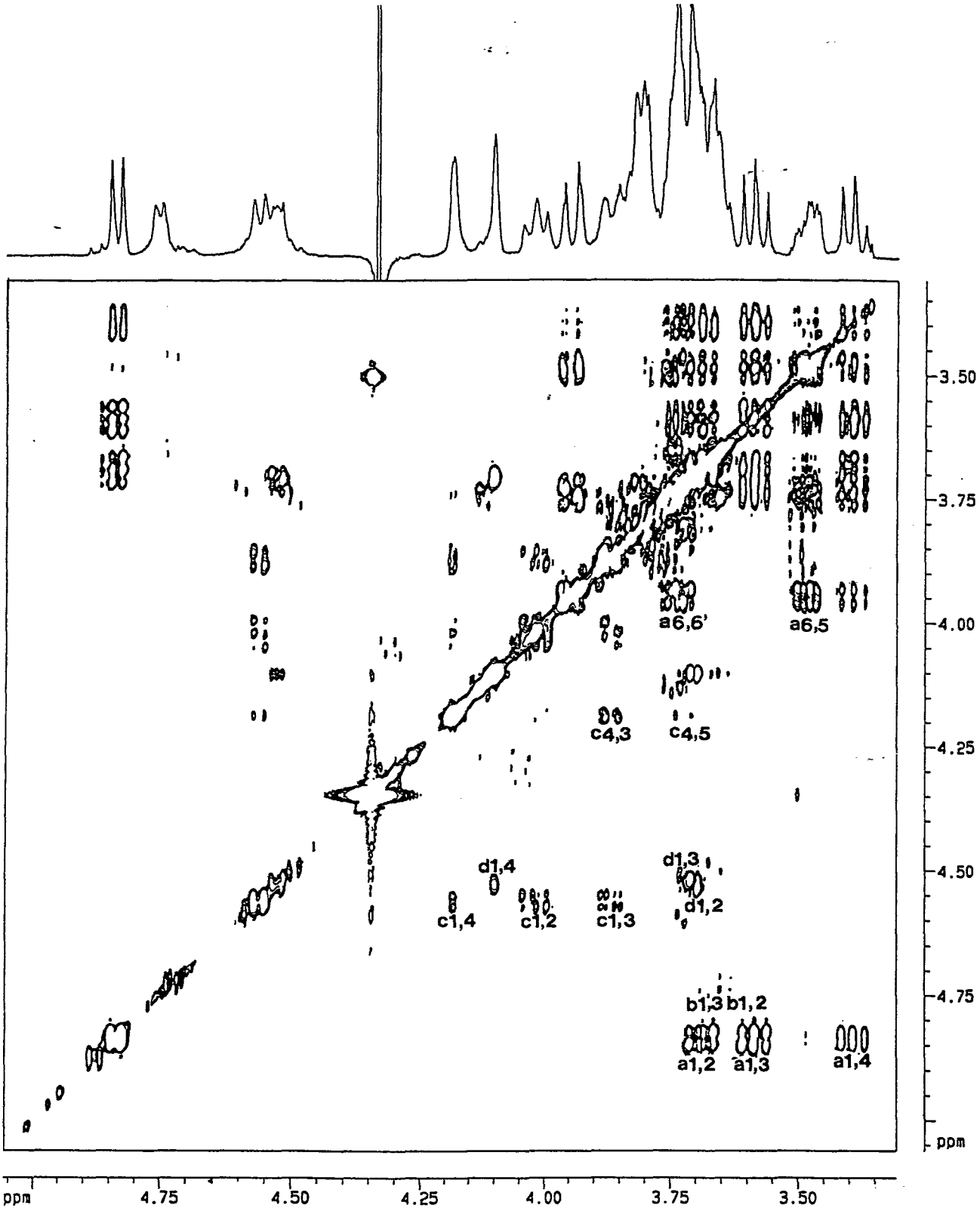


Figure 4.2 Two-step RELAY COSY contour plot of the PS for the region δ 3.30 - 5.05: a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc. The ^{1}H NMR spectrum is projected along the f_2 axis

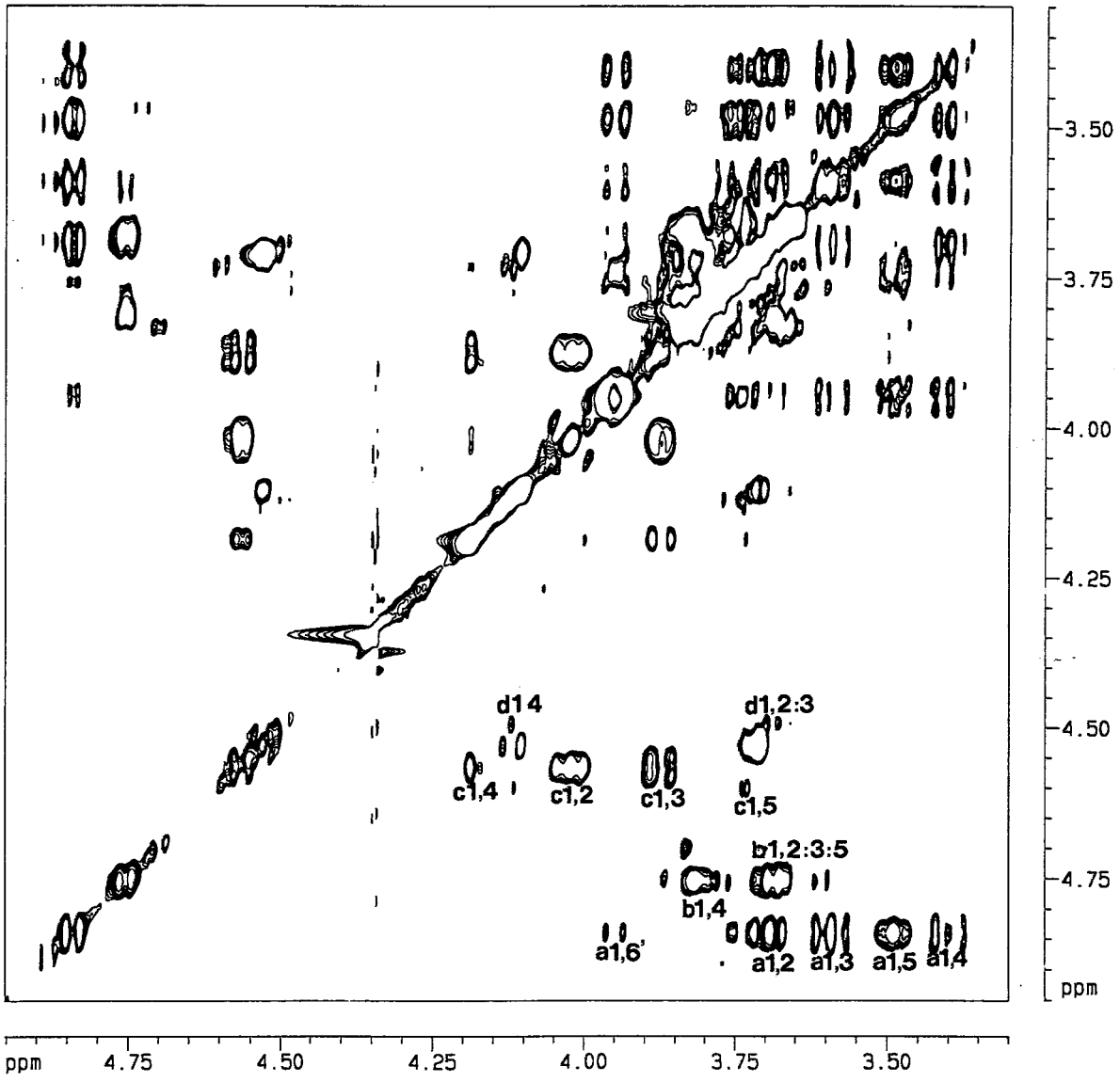


Figure 4.3 HOHAHA contour plot of the PS for the region δ 3.00-5.05: a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc.

Residue a [ß-D-Glc*p*NAc]: The ^{1}H resonances for residue a were traced readily via their cross-peaks in the COSY and RELAY COSY spectra. Magnetism was relayed very efficiently through the spin system due to the large vicinal coupling constants and all cross-peaks were clearly visible. The ^{13}C resonances for residue a were assigned by comparing the ^{1}H assignments with the ^{1}H-^{13}C correlation data obtained from an HMQC[283] experiment (Figure 4.4 and Table 4.1).

Residue b [→3,4)-ß-D-Glc*p*A]: The assignment of the ^{1}H resonances for this residue was complicated by the amount of signal overlap in the H-2 region of the COSY spectrum. The HOHAHA spectrum (Figure 4.3) showed that the signals for H-2, H-3, and H-5 were largely overlapped. As a result, it was not possible to distinguish between the C-2 and C-3 resonances in the HMQC spectrum (Figure 4.4). Assignment of these carbon resonances was only possible using an HMQC-TOCSY[267] experiment, which clearly showed correlations between the H-1 and C-1, H-1 and C-2, and H-1 and C-3 resonances of this residue (Figure 4.5).

Residue c [→3)-ß-D-Gal*p*NAc]: The ^{1}H resonances for H-1 to H-5 of residue c were assigned via their cross-peaks in the COSY and RELAY-COSY spectra. The ^{13}C resonances for the corresponding C atoms were assigned by correlation from the HMQC spectrum. The C-6 signal was identified from the H-5/C-6 cross-peak in the HMQC-TOCSY spectrum (Figure 4.5) and the H-6 signal was then confirmed from the ^{1}H-^{13}C correlation data.

Residue d [→3)-ß-D-Gal*p*]: The ^{1}H resonances for H-1 to H-4 were assigned from the COSY and RELAY-COSY spectra. No further information could be obtained from the HOHAHA spectrum due to signal overlap and poor relay of magnetism in this spin system due to the small $^{3}J_{4,5}$ value[143]. The corresponding ^{13}C signals were assigned by comparison with the correlation data from the HMQC spectrum. The two remaining unassigned pairs of ^{1}H-^{13}C resonances from the ^{1}H-^{13}C correlation data could be assigned, by inspection, to H-5/C-5 and H-6/C-6 of residue d. The connectivity between these two pairs of resonances was confirmed from the HMQC-TOCSY spectrum.

Comparison of the chemical shift data for residues **a** - **d** with those reported for methyl glycosides[144,284,285] identified the residues as indicated in Table 4.1. The significant deshielding displayed by C-3 and C-4 of residue **b**, C-3 of residue **c**, and C-3 of residue **d** identified the linkage positions. The shift positions of the C-2 signals of residues **a** and **c** confirmed the position of the *N*-acetyl groups in these residues. These data are in agreement with the results of the methylation analysis for the **PS**.

Table 4.1 NMR data[a] for *E. coli* K101 polysaccharide

Residue		Proton or carbon						
		1	2	3	4	5	6	6′
ß-D-Glc*p*NAc (a)	H	4.85	3.68	3.59	3.39	3.48	3.74	3.94
	C	102.69	56.92	74.62	71.21	77.06	62.01	
→3,4)-ß-D-Glc*p*A (b)	H	4.75	3.66	3.67	3.83	3.72		
	C	103.17	75.36	81.35	80.47	75.58	172.07	
→3)-ß-D-Gal*p*NAc (c)	H	4.57	4.02	3.88	4.18	3.73	3.82	3.82
	C	102.16	51.86	80.85	68.69	75.74	61.92	
→3)-ß-D-Gal*p* (d)	H	4.52	3.69	3.71	4.11	3.64	3.73	3.73
	C	104.77	70.49	84.59	68.69	75.58	61.88	

[a]Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 ppm for ^{1}H and ^{13}C, respectively

Sequencing of the residues.— The sequence of the residues **a** - **d** in the repeating unit was established by a heteronuclear multiple bond correlation (HMBC) experiment[195]. Correlations between H-1 of GlcNAc and C-3 of GlcA, between H-1 of GlcA and C-3 of Gal, between H-1 of GalNAc and C-4 of GlcA, and between H-1 of Gal and C-3 of GalNAc were clearly visible, thus establishing the sequence of the residues in the repeating unit. Correlations between the carbonyl resonances at 175.53 and 175.25 ppm and the H-2 signals of GalNAc and GlcNAc respectively permitted assignment of these carbons to the two amino sugars. The remaining carbonyl resonance at 172.07 ppm could then be assigned to GlcA. Other intra-residue connectivities served to confirm the assignments made for the residues as listed in Table 4.2.

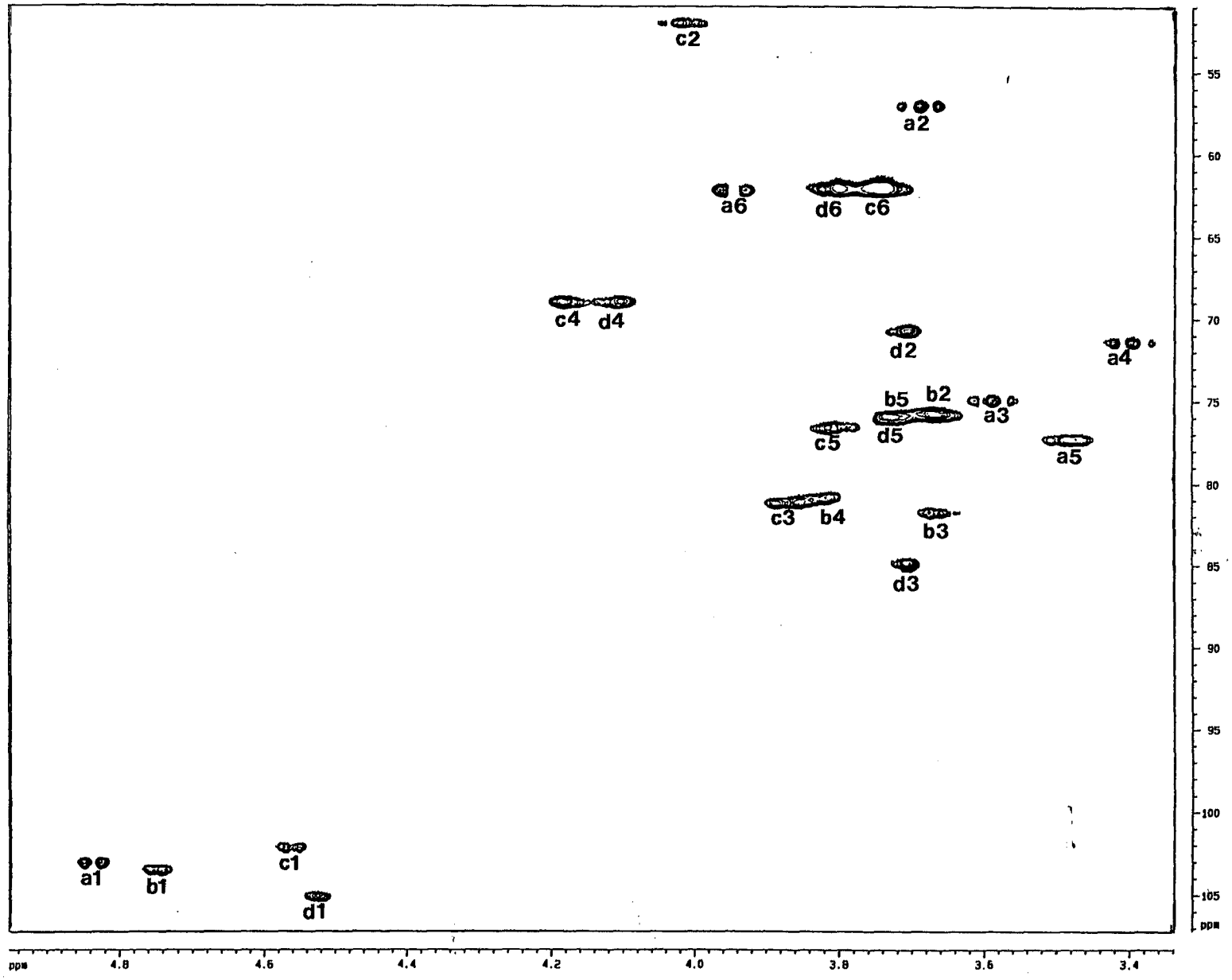


Figure 4.4 HMQC ^{1}H-^{13}C shift correlation map of the spectral region f_1 50 - 106 ppm (^{13}C) and f_2 δ 3.34 - 4.94 of the PS

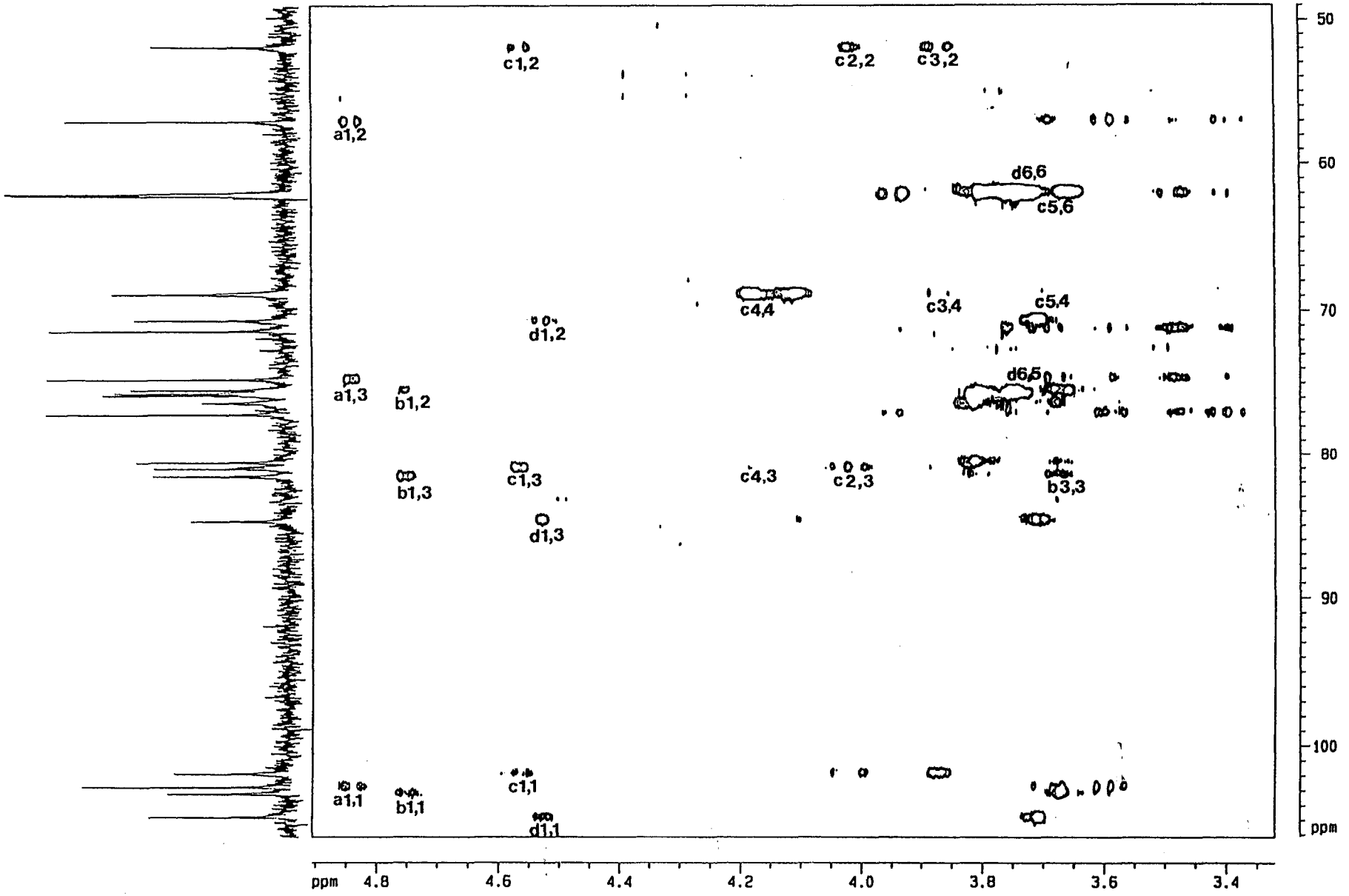


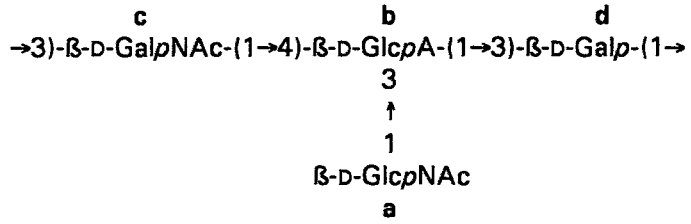
Figure 4.5 HMQC-TOCSY ^{1}H-^{13}C contour plot of the spectral region f_1 50 - 106 ppm (^{13}C) and f_2 δ 3.35 - 4.90 (^{1}H) of the PS. The ^{13}C NMR spectrum is projected along the f_1 axis.

Table 4.2 Two- and three-bond ^{1}H-^{13}C correlations for *E. coli* K101 PS

Residue	Proton	Correlation to
ß-D-Glc*p*NAc (a)	H-1	56.92 (a, C-2), 81.35 (b, C-3)
	H-2	102.69 (a, C-1), 74.62 (a, C-3), 175.25 (a, C=O)
	H-3	56.92 (a, C-2), 71.21 (a, C-4)
	H-4	74.62 (a, C-3), 77.06 (a, C-5), 61.88 (a, C-6)
	H-6	71.21 (a, C-4), 77.06 (a, C-5)
→3,4)-ß-D-Glc*p*A (b)	H-1	84.59 (d, C-3)
	H-2	103.17 (b, C-1), 81.35 (b, C-3), 75.58 (b, C-5)
	H-4	81.35 (b, C-3), 75.58 (b, C-5)
→3)-ß-D-Gal*p*NAc (c)	H-1	80.47 (b, C-4)
	H-2	102.16 (c, C-1), 80.85 (c, C-3), 175.53 (b, C=O)
	H-3	51.86 (c, C-2)
	H-4	51.86 (c, C-2), 80.85 (c, C-3)
	H-5	68.69 (c, C-4), 61.88 (c, C-6)
→3)-ß-D-Gal*p* (d)	H-1	80.85 (c, C-3)
	H-2	104.77 (d, C-1), 84.59 (d, C-3)
	H-4	70.49 (d, C-2), 84.59 (c, C-3)
	H-5	68.69 (d, C-4), 61.88 (d, C-6)

4.1.4 Conclusion

The combined chemical and NMR data permit the structure of the tetrasaccharide repeating unit of the K101 capsular polysaccharide to be written as:



The K101 capsular polysaccharide is the first of the *E. coli* series found to contain a terminal D-GlcNAc residue and joins a group of thirteen other K-antigens whose constituent monosaccharides are all ß-linked. The "3 + 1" structure of the branched tetrasaccharide repeating unit is shared by five other capsular antigens of the *E. coli* series, *viz. E. coli* K27, K28, K33, K36, and K48[48,55]. The occurrence of lipid on this and other capsular antigens of the *E. coli* series is

noteworthy, and the nature of the lipid at the reducing end of the polysaccharide chain is one of the criteria for sub-group classification[58]. Although the lipid present in the K101 antigen was not examined, it is probably core lipid A, in accordance with a few other Group I polysaccharides[58].

4.1.5 Experimental

General methods.— Unless otherwise stated, solutions were concentrated under reduced pressure at temperatures not exceeding 40°C (water bath). Analytical GLC was performed on a Hewlett-Packard 5890A gas chromatograph, fitted with flame-ionization detectors and a 3392A recording integrator, with helium as carrier gas. A J&W Scientific fused-silica DB-17 bonded-phase capillary column (30 m x 0.25 mm; film thickness 0.25 μm) was used for separating partially methylated alditol acetates (programme I), alditol acetates (programme II), and acetylated octyl glycosides (programme II). A J&W Scientific fused-silica DB-Wax bonded-phase capillary column (30 m x 0.25 mm; film thickness 0.15 μm) was used for separating alditol acetates of ManNAc, GlcNAc, and GalNAc (240°C isothermal). A J&W Scientific DB-225 bonded-phase capillary column (30 m x 0.25 mm; film thickness 0.25 μm) was also used for separating acetylated octyl glycosides (240°C isothermal). The temperature programmes used were: I, 180°C for 2 min, then 3°Cmin^{-1} to 240°C; II, 180°C for 2 min, then 2°Cmin^{-1} to 240°C. The identities of all derivatives were determined by comparison with authentic standards and confirmed by GLC-MS on a Hewlett-Packard 5988A instrument, using the appropriate column. Mass spectra were recorded at 70 eV and an ion-source temperature of 200°C.

Isolation and purification of the K101 polysaccharide.— An authentic stab culture of *E. coli* K101, obtained from Dr I Ørskov (Copenhagen), was plated on Mueller-Hinton agar and incubated at 37°C overnight. A single actively growing colony was selected and replated to ensure strain purity. Single colonies were then transferred to each of 9 sterile culture tubes containing 10 mL of sterile nutrient broth and shaken at 37°C for 8 h. The contents of the tubes were used to inoculate 9 sterile stainless steel trays (60 cm x 40 cm) containing Mueller-Hinton agar (1.2 L) and the glass-covered trays were incubated for 18 h at 37°C. The bacterial slime was harvested,

diluted with an equal volume of 2% phenol solution, and gently stirred at 4°C overnight to dissolve the capsular material. Cell debris and other undissolved material was removed by ultracentrifugation (35 000 rpm on a Beckman L8-M ultracentrifuge, rotor type 70 Ti, 3 h) and the supernatant was precipitated into ethanol (5 vols). The precipitate was isolated using low-speed centrifugation (3000 rpm, 10 min) and was redissolved in a minimum quantity of water (*ca.* 200 mL). Cetyltrimethylammonium bromide (CTAB, 5% w/v in water; *ca.* 10 mL) was added dropwise and the resulting precipitate was isolated by low-speed centrifugation as above. The acidic polysaccharide-CTAB complex was dissolved in 3 M NaCl (150 mL) and was precipitated into ethanol (5 vols). The precipitate was collected by centrifugation, redissolved in water (150 mL) and dialysed exhaustively (12-14 000 mw cut-off) against tap-water. The solution was freeze-dried to yield 1.28 g of the acidic polysaccharide. This crude isolate gave an opaque solution when dissolved in water and was thus treated to remove lipid. Crude capsular polysaccharide (600 mg) was dissolved in 1% acetic acid (100 mL) and heated at 100°C for 1 h. The solution was cooled, ultracentrifuged (25 000 rpm on a Beckman L8-M ultracentrifuge, rotor type SW 28; 1 h), dialysed overnight and freeze-dried to yield lipid-free K-antigen (385 mg), which was used for chemical analysis without further purification. A small sample (30 mg) was purified by GPC on a dextran-calibrated column (1.6 m x 65 cm) of Sephacryl™ S500, using 0.1 M sodium acetate buffer (pH 5.00) as eluent. Material was detected by refractive index.

Glycose analysis.— A small portion (1 mg) of the **PS** was hydrolysed with 4 M trifluoroacetic acid (1 mL, 125°C, 1 h). The reaction product was concentrated under reduced pressure after which it was co-concentrated several times with aliquot portions of water to ensure that all the acid had evaporated. Alditol acetates were prepared by reduction of the liberated sugars in aqueous solution with $NaBH_4$ for 1 h followed by acetylation with 2:1 Ac_2O-pyridine (1.5 mL) for 1 h at 100°C. The resulting alditol acetates were recovered by extraction into $CHCl_3$ and analyzed by GLC. The identity of the uronic acid was revealed by methanolysis of the **PS** (5 mg) using 3% methanolic HCl (80°C, 16 h). The acid was neutralized with Ag_2CO_3, followed by treatment with $NaBH_4$ in anhydrous CH_3OH to effect carboxyl reduction. Glycose analysis of the

modified **PS** was carried out as detailed above.

A second sample of native **PS** was hydrolysed, followed by reduction with $NaBD_4$ and subsequent preparation of alditol acetates to investigate the formation of 3,6-glucuronolactone.

Absolute configuration of the monosaccharides.— The absolute configuration of the constituent monosaccharides were determined by GLC analysis of their acetylated (-)-2-octylglycosides[149]. The native polysaccharide (10 mg) was methanolysed (3% methanolic HCl, 80°C, 16 h), reduced ($NaBH_4$ in dry CH_3OH, overnight), and hydrolysed with 4M TFA (125°C, 1 h). The hydrolysate was dissolved in 5 mL water containing 0.5 mL methanol and was re-N-acetylated by treatment with Ac_2O (2 mL) for 2h at room temperature[286]. After concentration of the solution, the residue in water (2 mL) was treated with aqueous 25% NH_4OH (0.5 mL) and then freeze-dried in an ampoule. To the dried material was added 1 drop of concentrated TFA, 0.5 mL (-)-2-octanol and a small magnetic stirrer bar. The ampoule was sealed and ultrasonicated to solubilize the material, after which it was heated, with stirring, in an oilbath at 130°C for 16 h. The reaction mixture was then concentrated to dryness (55°C, vacuum pump attached to the rotary evaporator) and acetylated with 2:1 Ac_2O-pyridine (1.5 mL) at 100°C for 1 h. The derived acetylated octyl glycosides were recovered by extraction into $CHCl_3$ and analyzed by GLC.

Methylation of the PS.— Permethylation of the **PS** was achieved by a modified Hakomori procedure[169]. A dried sample of acidic **PS** (15 mg) was dissolved in dry DMSO (2 mL), methylsulphinyl (dimsyl) anion (1 mL) added under N_2, and the reaction mixture stirred at room temperature for 1 h. The solution was cooled in an ice-water bath and MeI (0.5 mL) was added dropwise with stirring. The reaction mixture was allowed to warm to room temperature and was stirred for 4 h, after which it was dialysed overnight against running water (12-14 000 mw cut-off) and freeze-dried. A sample of the permethylated product (5 mg) was hydrolysed (4M TFA, 1 mL, 125°C, 1 h), reduced ($NaBH_4$ in aqueous solution) and the resulting partially methylated alditols were acetylated with 2:1 Ac_2O-pyridine (1.5 mL, 100°C, 1 h). The partially methylated alditol acetates were analyzed by GLC-MS.

A second sample of dried permethylated **PS** (10 mg) was methanolysed (3% methanolic HCl, 80°C, 16 h), reduced ($NaBH_4$ in dry CH_3OH, overnight), hydrolysed (4 M TFA, 125°C, 1 h) and the partially methylated monosaccharides converted into their alditol acetates as before.

NMR spectroscopy.— Samples were deuterium-exchanged with D_2O by freeze-drying several times from D_2O and then examined as solutions in 99.99% D_2O containing a trace of acetone as internal standard (δ 2.23 for ^1H and 31.07 ppm for ^{13}C). Spectra were recorded at 67°C on a Bruker AMX-400 spectrometer equipped with an X32 computer. The parameters used for 2D experiments were as follows: COSY, two-step RELAY COSY, and HOHAHA [256 x 2048 data matrix, zero-filled to 1024 data points in t_1; 96 or 128 scans per t_1 value; spectral width 1362.4 Hz; recycle delay 1.0 s; fixed delay of 0.036 s (RELAY COSY); unshifted sine-bell filtering in t_1 and t_2 (COSY and RELAY COSY); and shifted sine-squared filtering in t_1 and t_2 (HOHAHA)]. HMQC and HMQC-TOCSY [512 x 4096 data matrix, zero-filled to 1024 data points in t_1; spectral width 14 022 Hz in t_1 and 1362.4 Hz in t_2; 52 and 48 scans, respectively, per t_1 value; recycle delay 1.0 s; fixed delay 3.45 ms; shifted sine-squared filter]. HMBC [256 x 2048 data matrix, zero-filled to 1024 data points in t_1; spectral width 20 733 Hz in t_1 and 1362.4 Hz in t_2; 96 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.45 ms; shifted sine-squared filter].

4.2 *Escherichia coli* SEROTYPE K45 CAPSULAR ANTIGEN: A GLYCAN CONTAINING 3-ACETAMIDO-3,6-DIDEOXYGALACTOPYRANOSE.

4.2.1 Abstract

The primary structure of the acidic capsular antigen of *Escherichia coli* O8:K45:H9 was shown by glycose analysis, methylation analysis and one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy to be composed of repeating branched tetrasaccharide units having the structure:

→6)-*α*-D-Glc*p*NAc-(1→4)-ß-D-Gal*p*A-(1→3)-ß-D-Glc*p*NAc-(1→
3
↑
1
ß-D-Fuc*p*3NAc

The polysaccharide contains the uncommon amino sugar 3-acetamido-3,6-dideoxy-D-galactopyranose, which has not been previously identified in a capsular antigen.

4.2.2 Introduction

The capsular (K) antigen of *E. coli* O8:K45:H9 belongs to the group I polysaccharides of this species[84] and as such is a high-molecular mass, acidic polysaccharide which contains amino sugars. The polysaccharide is unusual in that it contains the amino sugar 3-acetamido-3,6-dideoxy-D-galactopyranose (D-Fuc3NAc) which has previously been shown to be present in lipopolysaccharides and cell-wall antigens of *E. coli* [287], *Pseudomonas maltophilla*[288], *Eubacterium saburreum*[289], *Proteus penneri* [290] and *Xanthomonas campestris*[291]. This is the first reported occurrence of this amino sugar in a capsular polysaccharide. The capsular antigen of *E. coli* K45 has been reported to exhibit serological cross-reactivity with the *E. coli* O74 O-antigen, suggesting that these compounds may contain similar antigenic determinants[44,46]. The structure of the O74 antigen has, to our knowledge, not yet been published.

4.2.3 Results and Discussion

Isolation, composition and 1D NMR spectra of the capsular antigen.— *E. coli* K45 bacteria were grown on Mueller-Hinton agar and the acidic capsular polysaccharide (**PS**) was isolated and purified by precipitation with cetyltrimethylammonium bromide, followed by delipidation with aqueous 1% acetic acid. The average molecular mass of **PS** was found to be M_r 1.5×10^6 on a dextran-calibrated column of Sephacryl™ S500. Hydrolysis of **PS** followed by GLC-MS examination of the derived alditol acetates showed the presence of GlcN and a 3-amino-3,6-dideoxyhexose in the ratio 2:1. Methanolysis of the **PS** followed by reduction of the methoxycarbonyl groups, hydrolysis, and GLC analysis of the derived alditol acetates, revealed the additional presence of Gal. These results indicate that the **PS** comprises a tetrasaccharide repeating unit consisting of GlcNAc, GalA and 3-amino-3,6-dideoxyhexose in the approximate molar ratio 2:1:1.

GlcNAc was shown to have the D configuration by GLC analysis of the derived acetylated (-)-2-octyl glycoside[149]. A sample of the **PS** was partially hydrolysed with 0.5 M trifluoroacetic acid and, after re-N-acetylation of the sample, both GalA and the 3-acetamido-3,6-dideoxyhexose were isolated from the hydrolysate by paper chromatography. The 3-acetamido-3,6-dideoxyhexose was further purified by semi-preparative HPLC and identified as 3-acetamido-3,6-dideoxygalactose on the basis of its ^{1}H NMR spectrum. The $[\alpha]_D$ was determined to be $+89°$ (water), indicating that the sugar had the D configuration[292]. The separated GalA fraction was also shown to have the D configuration by GLC analysis of the acetylated (-)-2-octyl glycoside[149] of the derived Gal.

The ^{1}H NMR spectrum of the **PS** in D_2O (Figure 4.6) contained signals for H-1 of an α-linked sugar at δ 4.94 ($^3J_{1,2}$ 3.6 Hz) and for H-1 of ß-linked sugars at 4.62 ($^3J_{1,2}$ 7.4 Hz), 4.55 ($^3J_{1,2}$ 7.9 Hz) and 4.48 ($^3J_{1,2}$ 7.8 Hz). Two signals for the methyl protons of NAc groups at δ 2.06 (6H) and 2.02 (3H) respectively were also present. A doublet centred at δ 1.26 ($^3J_{5,6}$ 6.3 Hz; 3H) indicated the presence of a 6-deoxy sugar. The ^{1}H-decoupled ^{13}C NMR spectrum (Figure 4.7) confirmed a tetrasaccharide repeating unit with signals for C-1 at 104.55, 104.34, 102.28 and 99.40 ppm and

signals for carbonyl carbons at 175.15, 175.06 and 171.23 ppm. Three signals, at 55.63, 55.43 and 53.07 ppm, indicated the presence of three C-N bonds, and confirmed the presence of three NAc groups. A signal at 16.22 ppm confirmed the presence of a 6-deoxy sugar.

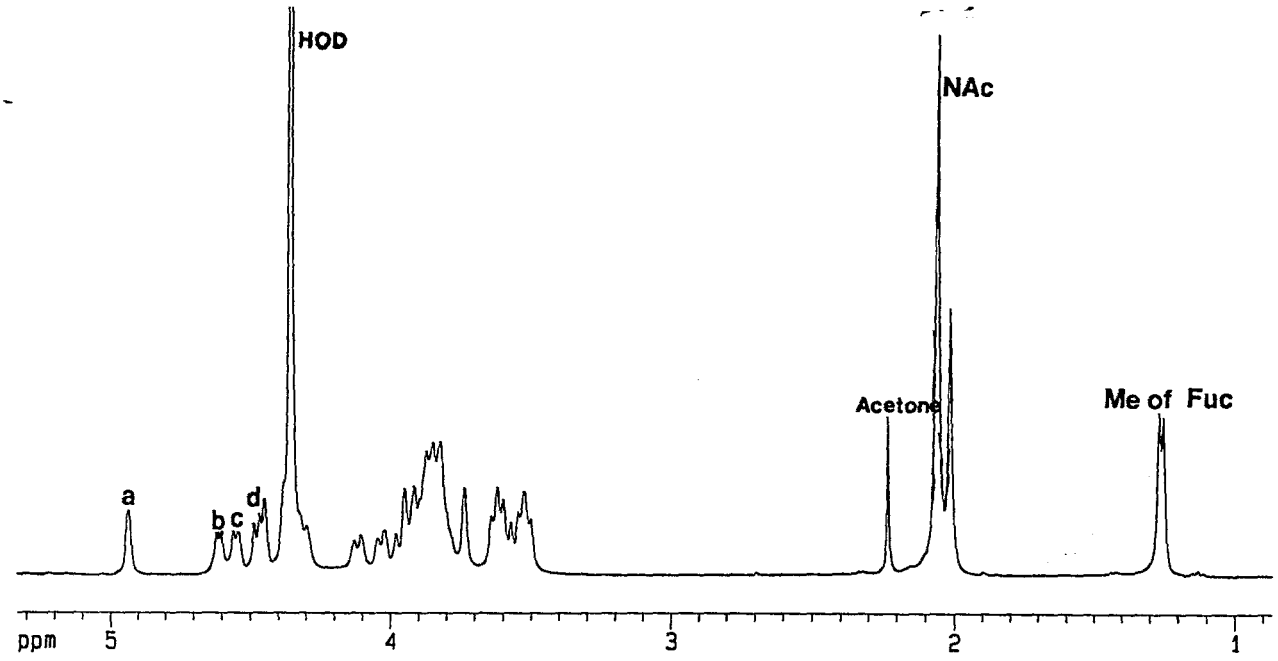


Figure 4.6 ^{1}H NMR spectrum of the *E. coli* capsular polysaccharide (400 MHz; 300K)

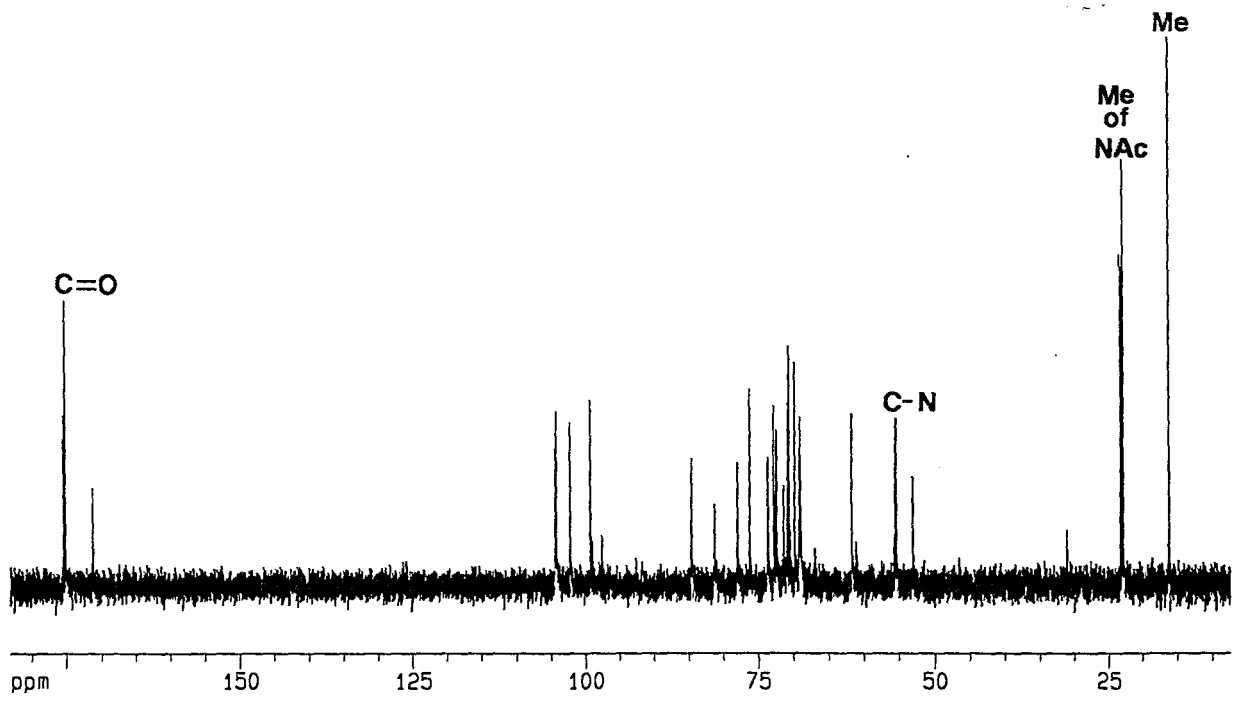


Figure 4.7 ^{13}C NMR spectrum of the *E. coli* K45 capsular antigen (338K)

Methylation analysis.— Methylation of the polysaccharide followed by GLC and GLC-MS analysis of the permethylated alditol acetates derived from the products of an acid hydrolysate gave 3,6-dideoxy-3-methylacetamido-2,4-di-*O*-methylgalactose, 2-deoxy-2-methylacetamido-4-*O*-methylglucose, 2-deoxy-2-methylacetamido-4,6-di-*O*-methylglucose and 2,3-di-*O*-methylgalactose (after carboxyl reduction). These results indicate the presence of 3,6-linked GlcNAc, 4-linked GalA, 3-linked GlcNAc, and terminal Fuc3NAc in a branched tetrasaccharide repeating unit with GlcNAc as the branch point.

2D NMR studies of the E. coli *K45 capsular polysaccharide.*— The identities and sequence of the residues in the repeating unit were established by 2D NMR experiments, which also confirmed the glycosylation sites in the polysaccharide. The residues in the repeating unit were labelled a - d in order of decreasing chemical shift of their anomeric protons (Figure 4.6). The chemical shifts of the proton resonances of residue a - d were established from COSY[259] (Figure 4.8) and 2D Homonuclear Hartmann-Hahn (HOHAHA)[261] (Figure 4.9) experiments. NMR data are given in Table 4.3.

Residue a [→3,6)-*α*-D-Glc*p*NAc]: The ^{1}H resonances for residue a were readily traced via their crosspeaks in the COSY spectrum as far as H-3. The H-3/4 cross-peak overlapped with several other signals, but its position could easily be identified with the help of the HOHAHA spectrum. The chemical shifts of the remaining ^{1}H resonances were then assigned using both the HOHAHA and COSY spectra. As expected for a *gluco*-type residue, magnetism relayed well through the spin system. The ^{13}C resonances for residue a were assigned by comparing the ^{1}H assignments with the ^{1}H-^{13}C correlation data from an HMQC[283] experiment (Figure 4.10).

Residue b [→3)-ß-D-Glc*p*NAc]: The ^{1}H resonances for residue b were traced as far as H-3 via the crosspeaks in the COSY spectrum. The H-3/4 crosspeak overlapped with that for residue a and the H-4 resonance had therefore to be assigned from the HOHAHA spectrum. Magnetism also relayed well through this spin system and the remaining ^{1}H resonances for this residue were readily

assigned from the HOHAHA spectrum and confirmed from the COSY spectrum. The ^{13}C resonances were then assigned from the HMQC spectrum as before.

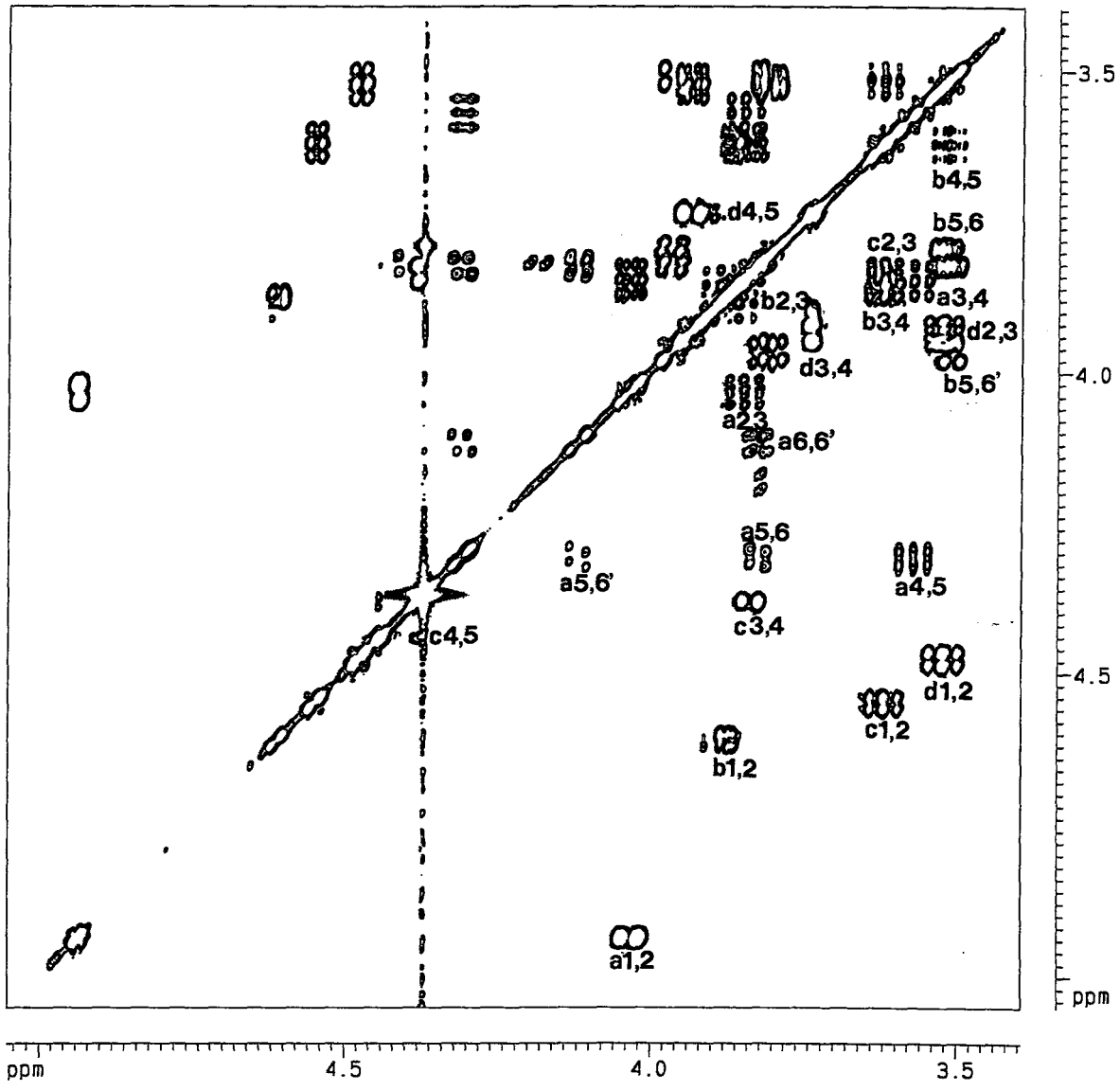


Figure 4.8 COSY contour plot of the region δ 3.42 - 5.02 of the *E. coli* K45 capsular antigen. The proton resonances of the sugar residues are labelled a - d. a1 connotes H-1 of residue a and a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc.

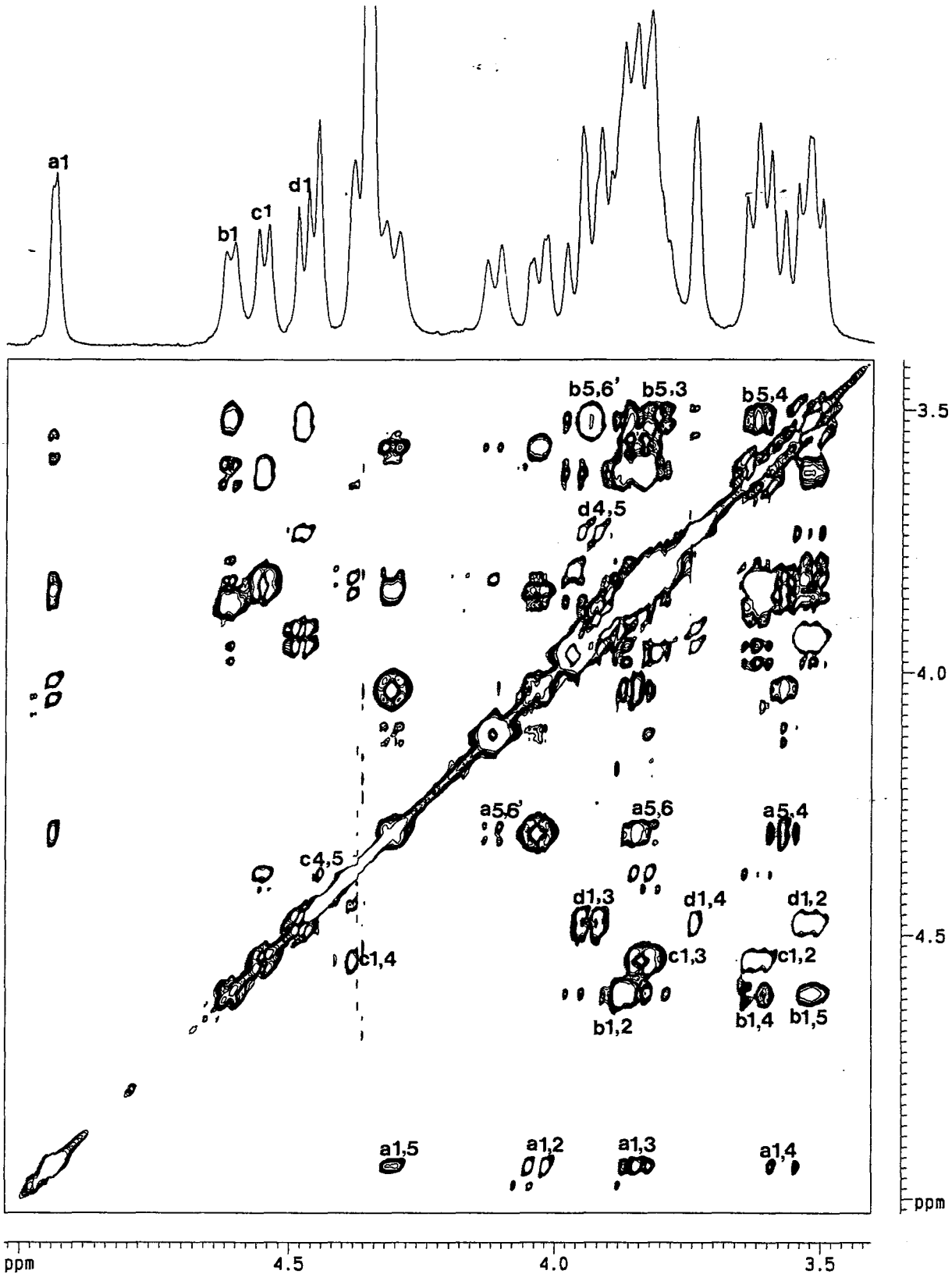


Figure 4.9 HOHAHA contour plot of the region δ 3.42 - 5.02 of the PS. a1 connotes H-1 of residue a, and a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc. The ^{1}H NMR spectrum is projected along the f_2 axis.

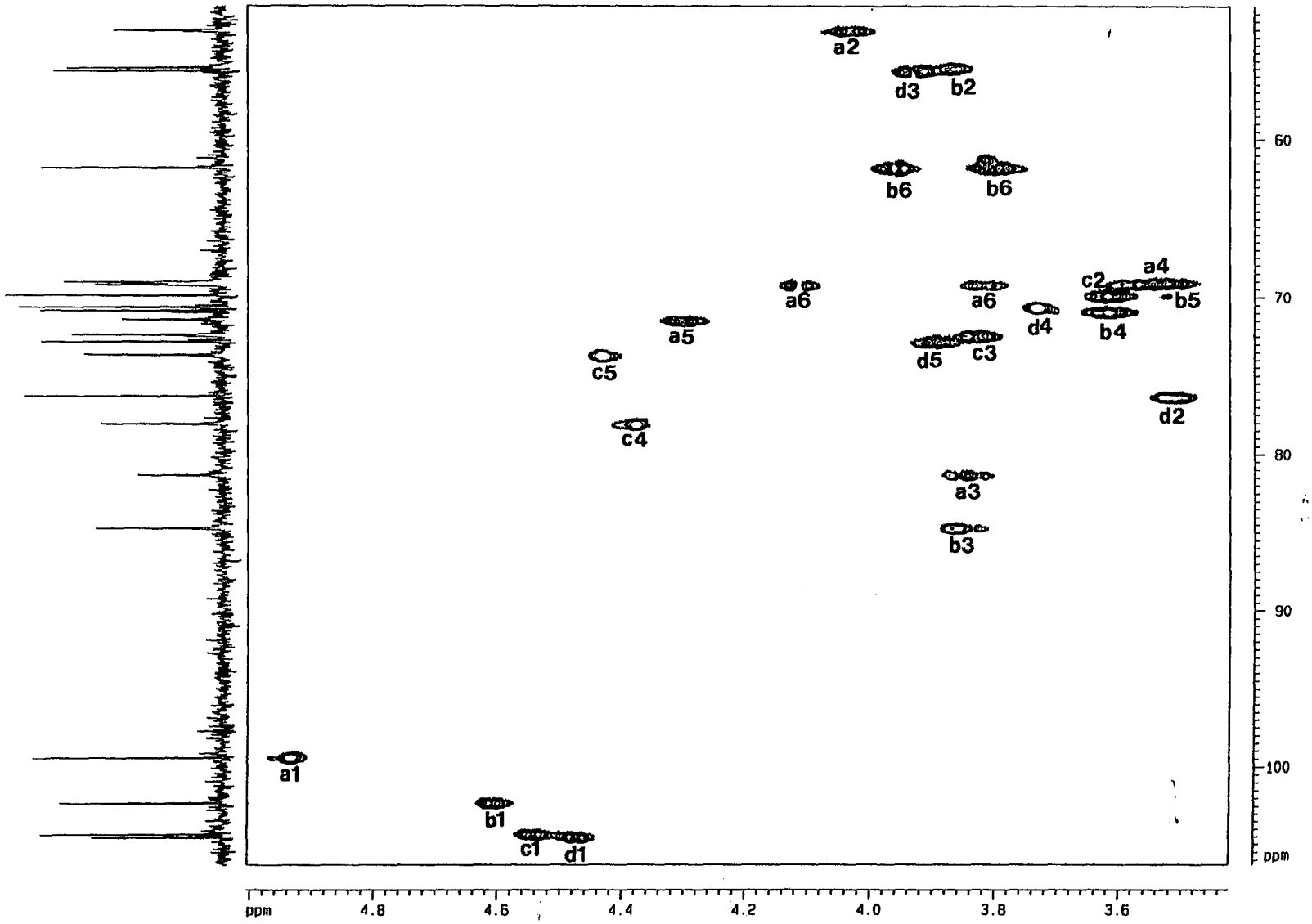


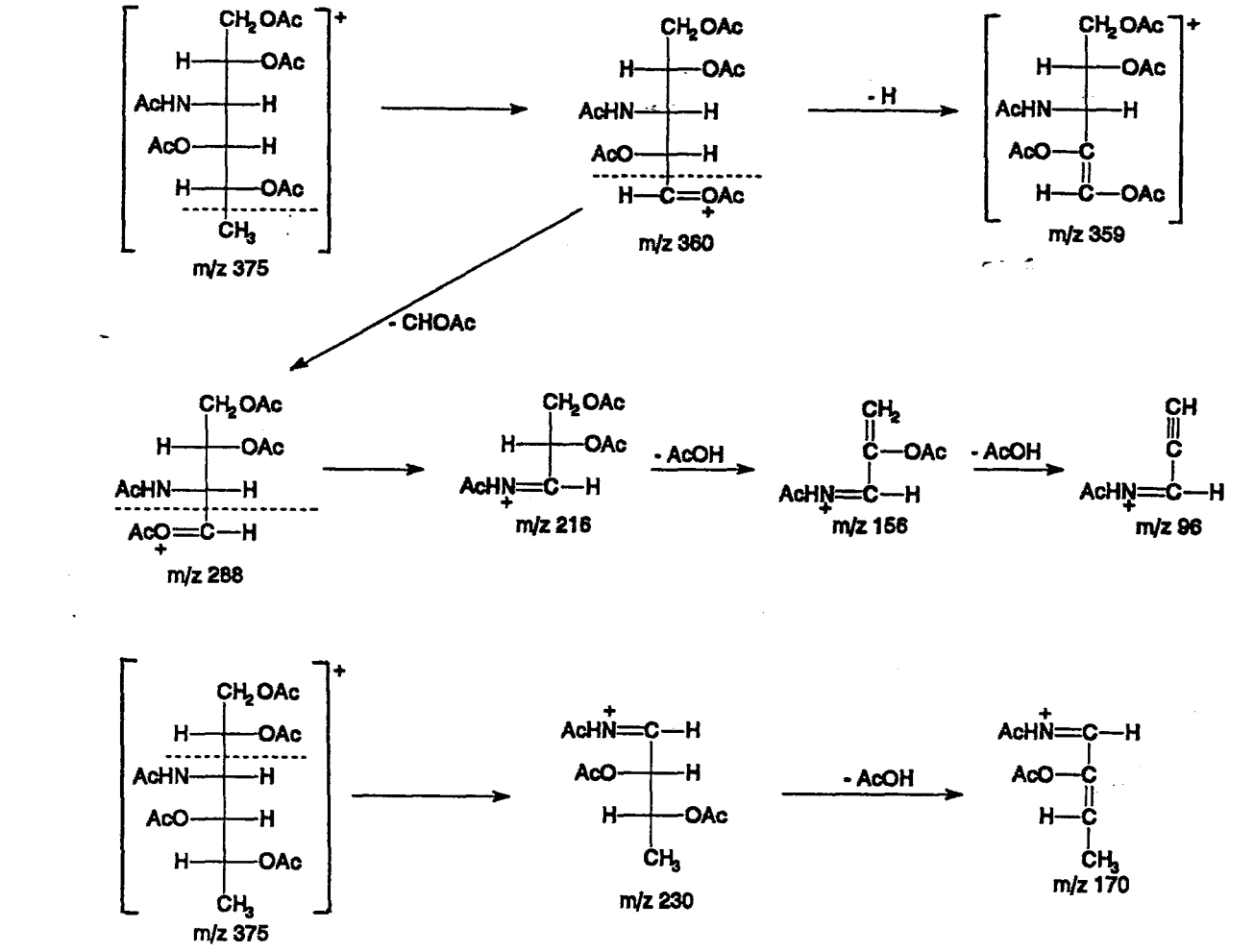
Figure 4.10 HMQC ^{1}H-^{13}C shift correlation map of the spectral region f_1 50 - 106 ppm (^{13}C) and f_2 δ 3.44 - 5.00 (^{1}H) of the PS. The ^{13}C NMR spectrum is projected along the f_1 axis.

Residue **c** [→4)-ß-D-Gal*p*A]: In the COSY spectrum the H-2/3 crosspeak for residue **c** overlapped with the H-3/4 crosspeaks for residues **a** and **b**, however the H-3 resonance could be assigned from the HOHAHA spectrum. The magnetism in this spin system did not relay beyond H-4, as expected for a *galacto*-type residue and the chemical shift of the H-5 resonance was assigned from the small H-4/5 crosspeak in the COSY spectrum. Carbon resonances were assigned from the HMQC spectrum, except for C-6, which was assigned with the help of the HMBC spectrum (See Table 4.4).

Residue **d** [ß-D-Fuc*p*3NAc]: Beginning with the H-1/2 crosspeak in the COSY spectrum, the ^{1}H resonances for residue **d** were traced as far as H-4. The magnetism in this spin system did not appear to relay beyond H-4 so that no further information could be obtained from the HOHAHA spectrum. The remaining ^{1}H resonances were assigned by working back from the H-6 signal at δ 1.26 in the COSY spectrum, which allowed the resonance for H-5 to be assigned and showed that the H-3 and H-5 resonances in fact overlapped, with resulting overlap of the H-3/4 and H-4/5 crosspeaks. The ^{13}C resonances were assigned by comparing the ^{1}H assignments with the ^{1}H-^{13}C correlation data. The carbon resonance at 55.63 ppm could have been assigned to C-3 or C-5, due to the overlap of these two proton signals. However, GLC-MS analysis of the hydrolysed polysaccharide and identification of this sugar component as described earlier pinpointed the position of the acetamido function at C-3, allowing chemical shifts to be assigned for C-3 and C-5 respectively. The fragmentation pathway and mass spectrum of the Fuc3NAc component are shown in Figure 4.11.

Comparison of the chemical shift data for residues **a** - **d** with those reported for methyl glycosides[144,254,284] permitted identification of residue **a** as 3,6-linked α-Glc*p*NAc, residue **b** as 3-linked ß-Glc*p*NAc, residue **c** as 4-linked ß-Gal*p*A and residue **d** as terminal ß-Fuc*p*3NAc.

Sequencing of the residues.— The sequence of the residues **a** - **d** in the repeating unit was established by a Heteronuclear Multiple Bond Correlation (HMBC) experiment[195].



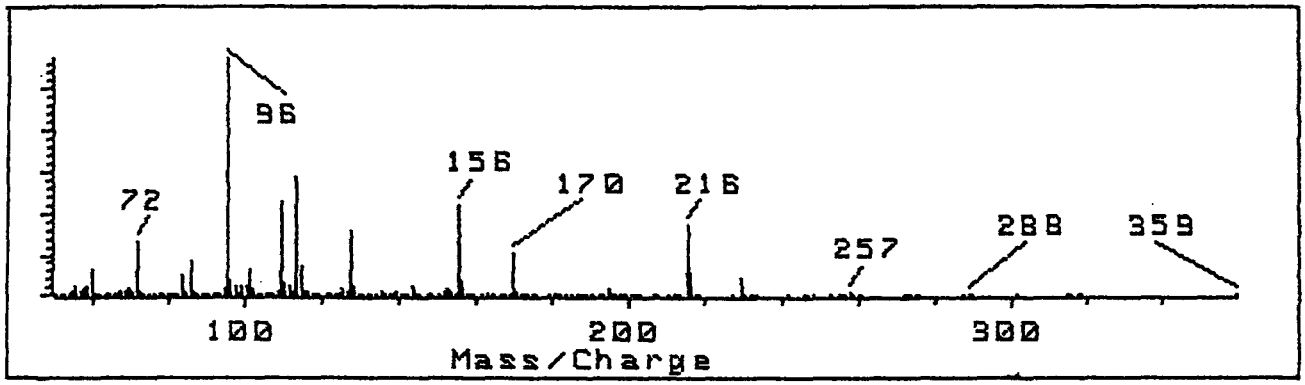


Figure 4.11 Fragmentation pathway and mass spectrum of the derived alditol acetate of Fuc3NAc

Correlations between H-1 of α-GlcNAc and C-4 of ß-GalA, between H-1 of ß-GalA and C-3 of ß-GlcNAc, between H-1 of ß-GlcNAc and C-6 of α-GlcNAc, and between H-1 of ß-Fuc3NAc and C-3 of α-GlcNAc were clearly observed (Table 4.4 and Figure 4.12).

Table 4.3 **NMR data[a] for the *E. coli* K45 polysaccharide**

Residue		Proton or carbon						
		1	2	3	4	5	6	6'
→3,6)-α-D-Glc*p*NAc (a)	H	4.94	4.02	3.85	3.57	4.32	3.82	4.12
	C	99.40	53.07	81.32	69.14	71.43	69.23	
→3)-ß-D-Glc*p*NAc (b)	H	4.62	3.88	3.85	3.62	3.52	3.80	3.96
	C	102.28	55.43	84.68	69.88	76.32	61.77	
→4)-ß-D-Gal*p*A (c)	H	4.55	3.62	3.83	4.38	4.44		
	C	104.34	70.87	72.39	78.07	73.69	171.23	
ß-D-Fuc*p*3NAc (d)	H	4.48	3.52	3.93	3.74	3.91	1.25	
	C	104.55	69.06	55.63	70.63	72.85	16.22	

[a]Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 ppm for ^{1}H and ^{13}C, respectively

Table 4.4 **Two- and three bond ^{1}H-^{13}C correlations for PS**

Residue	Proton	Correlation to
→3,6)-α-D-GlcNAc (a)	H-1	81.32 (a, C-3), 71.43 (a, C-5), 78.07 (c, C-4)
	H-3	104.55 (d, C-1)
	H-6	102.28 (b, C-1),
→3)-ß-D-GlcNAc (b)	H-1	69.23 (a, C-6)
	H-2	102.28 (b, C-1), 55.43 (b, C-2)
	H-3	55.43 (b, C-2), 104.34 (c, C-1)
	H-4	84.68 (b, C-3), 61.77 (b, C-6)
→4)-ß-D-GalA (c)	H-1	84.68 (b, C-3), 70.87 (c, C-2)
	H-2	104.34 (c, C-1)
	H-3	104.34 (c, C-1), 78.07 (c, C-4)
	H-4	73.69 (c, C-5), 99.40 (a, C-1), 171.23 (c, C-6)
ß-D-Fuc3NAc (d)	H-1	81.32 (a, C-3)
	H-2	104.55 (d, C-1)
	H-3	69.06 (d, C-2)
	H-5	55.63 (d, C-3), 70.63 (d, C-4)

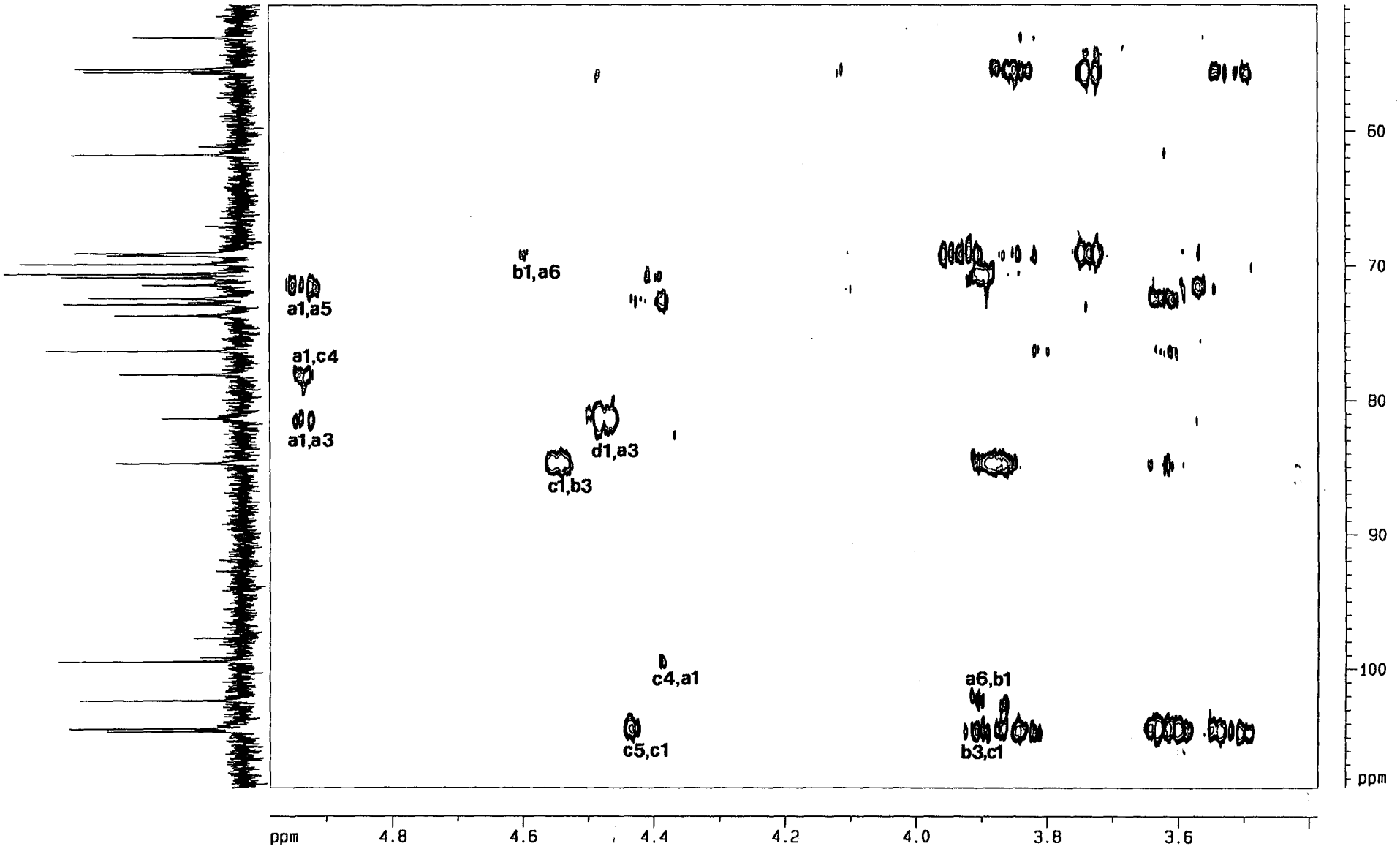
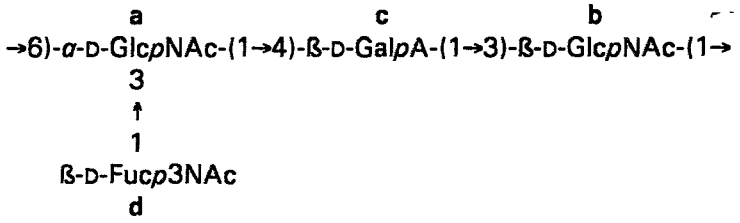


Figure 4.12 HMBC correlation map of the spectral region f_1 50 - 108 ppm (^{13}C) and f_2 δ 3.40 - 5.00 (^{1}H) of the PS. a1,c3 connotes the cross-peak between H-1 of residue a and C-3 of residue c, etc. The ^{13}C NMR spectrum is projected along the f_1 axis

4.2.4 Conclusion

The combined chemical and NMR data permit the structure of the tetrasaccharide repeating unit of the *E. coli* K45 capsular polysaccharide to be written as:



The capsular polysaccharide of *E. coli* K45 joins a group of six other capsular antigens among the *E. coli* series whose repeating units have the "3 + 1" structure, and which include K101 (see Section 4.1.4), K27, K28, K33, K36, and K48[55]. Although 6-linked sugar residues are not uncommon among the capsular antigens of *E. coli*, this is only the second instance of a 6-linked GlcNAc, the other being found in the capsular antigen of *E. coli* K40[61]. In common with many other Group I polysaccharides, the K45 polysaccharide has GalA as its acidic component and contains GlcNAc[46,56]. The occurrence of Fuc3NAc is, however, unusual. Although Fuc3NAc has been isolated from lipopolysaccharides and cell-wall antigens of *E. coli* [287] and is therefore not new to the genus, this is the first occurrence of this uncommon sugar in a capsular antigen. It is also worth noting that in previously reported occurrences of this sugar, it has been α-linked in all cases but one[287-291], and that the ß-linked Fuc3N reported in *Proteus penneri* contained an N-butyramido group rather than a NAc group[290]. This is therefore, to the best of our knowledge, the first ß-linked Fuc3NAc to be reported.

4.2.5 Experimental

General methods.— Instrumentation used was as described for *E. coli* K101. A J & W Scientific fused-silica DB-17 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was used for separating partially methylated alditol acetates (programme I), alditol acetates (programme II) and acetylated octyl glycosides (240° isothermal, pressure 140 kPa). The temperature programmes used were: I, 180° for 2 min, then 3°min^{-1} to 240°, 100 kPa; II, 180° for 2 min, then 2°min^{-1} to 240°, 100 kPa. The identities of all derivatives except Fuc3NAc were determined by comparison with authentic standards and confirmed by GLC-MS. GPC of the K45 polysaccharide was performed on a dextran-calibrated column (1.6 x 65 cm) of Sephacryl™ S500 using 0.1 M sodium acetate buffer (pH 5.00) as eluent. HPLC was carried out on a Supelcogel™ C-611 column (30 cm x 7.8 mm) using 10^{-4} M NaOH as eluent. Material was detected by refractive index. Optical rotation measurements were performed on a Perkin-Elmer model 141 polarimeter, using a 1 cm cell at 23°C (± 2°C).

Isolation and purification of K45 polysaccharide.— An authentic stab culture of *E. coli* O8:K45:H9 was obtained from Dr. I. Ørskov (Copenhagen) and propagated on Mueller-Hinton agar. The capsular polysaccharide was isolated as described for the K101 polysaccharide, and the crude isolate (450 mg) was treated with 1% aqueous acetic acid (100 mL, 100°C, 1 h) to remove lipid. The solution was ultracentrifuged (25 000 rpm, 1 h) and freeze-dried to yield 300 mg polysaccharide.

Glycose analysis.— A small sample of dried acidic polysaccharide (1 mg) was stirred in liquid anhydrous HF (introduced under vacuum using the apparatus described by Sanger and Lamport[80]) for 3 h at 25°C. Excess HF was removed under a stream of N_2 after which 50% aqueous AcOH (2 mL) was added and the mixture stirred for 1 h in order to hydrolyse the resultant glycosyl fluorides. The acetic acid was removed under reduced pressure and the residue reconstituted in water (0.5 mL) and alditol acetates were prepared as described for the K101 polysaccharide. Methanolysis was carried out as described for the K101 polysaccharide, followed

by hydrolysis in HF and preparation of alditol acetates as before. Alditol acetates were analyzed by GLC.

Partial hydrolysis.— K45 polysaccharide (200 mg) was hydrolysed with 0.5 M trifluoroacetic acid (100 mL, 100°C, 1 h) and the hydrolysate in water (5 mL) containing methanol (0.5 mL) was re-N-acetylated by treatment with acetic anhydride (2 mL) for 2 h at room temperature[288]. After concentration of the solution, the residue in water (2 mL) was treated with aqueous 25% NH_4OH (0.5 mL) before being subjected to descending mode paper chromatography on Whatman No 1 paper using ethyl acetate : acetic acid : pyridine : water (5:1:5:3) as eluent. The components were visualised with basic silver nitrate[293] and ninhydrin[294] and extracted from the paper with distilled water. The fraction containing Fuc3NAc (26.7 mg) was further purified by semi-preparative HPLC (yield 3.20 mg). The identity of the compound was confirmed by NMR spectroscopy and had $[\alpha]_D$ +89 (H_2O; *c* 0.32), indicating that its absolute configuration was D[292]. The ^1H NMR data (400 MHz, 303 K) for Fuc3NAc were as follows: α-anomer; H-1 5.19 ($^3J_{1,2}$ 3.3 Hz), H-2 3.80, H-3 3.98, H-4 3.82, H-5 4.12, H-6 1.33 ($^3J_{5,6}$ 6.88); ß-anomer; H-1 4.47 ($^3J_{1,2}$ 7.80 Hz), H-2 3.52, H-3 3.84, H-4 3.72, H-5 3.90, H-6 1.25 ($^3J_{5,6}$ 6.32 Hz).

Absolute configuration of the monosaccharides.— The absolute configurations of the GlcNAc and GalA residues in the repeating unit were established by GLC analysis of their derived acetylated (-)-2-octyl glycosides. A sample of polysaccharide (10 mg) was hydrolysed with 4 M TFA (125°C, 1 h). The reaction mixture was concentrated under vacuum, reconstituted in water (1 mL) containing methanol (0.3 mL), and re-N-acetylated as described above. The solution was evaporated to dryness, dissolved in water (1 mL) containing a few drops of aqueous 25% NH_4OH and freeze-dried in an ampoule. Acetylated (-)-2-octyl glycosides were prepared as described for the K101 polysaccharide and analyzed by GLC. The GalA isolated by paper chromatography of the partially hydrolysed polysaccharide (15 mg) was methanolysed (3% methanolic HCl, 80°C, 24 h), reduced ($NaBH_4$ in dry CH_3OH) and the product converted to its acetylated (-)-2-octyl glycoside as for the K101 polysaccharide.

Methylation of the PS.— Permethylation of the **PS** was achieved by successive treatment according to a modified Hakomori procedure[171] followed by Kuhn methylation[178]. A dried sample of acidic polysaccharide (20 mg) was dissolved in a mixture of dry DMSO (1 mL) and 1,1,1,1-tetramethylurea (1 mL) and methylsulphinyl (dimsyl) anion (2 mL) was added under N_2. The mixture was stirred at room temperature for 2 h. The reaction mixture was cooled in an ice-water bath and MeI (1 mL) was added dropwise with stirring. After 4 h this mixture was dialysed overnight against running water (12-14 000 mw cut-off) and freeze-dried. To the dried product was added DMF (2 mL), MeI (0.5 mL) and a small portion of Ag_2O. The reaction flask was covered with aluminium foil to exclude light and stirred at room temperature for 72 h, with aliquot portions of MeI and Ag_2O being added every 6 hours. The solution was centrifuged (2000 rpm, 10 min), the supernatant decanted and evaporated to dryness to produce a yellow oil. A portion of this permethylated polysaccharide (5 mg) was hydrolysed with liquid anhydrous HF, reduced and acetylated as described above. A second portion of permethylated polysaccharide (10 mg) was methanolysed (3% methanolic HCl, 80°C, 16 h), reduced ($NaBH_4$ in dry CH_3OH), hydrolysed, reduced and acetylated as described above. The resulting partially methylated alditol acetates were analyzed by GLC-MS.

NMR spectroscopy.— Samples were prepared as previously described and spectra were recorded at 65°C on a Bruker AMX-400 spectrometer equipped with an X32 computer. The parameters used for 2D experiments were as follows: COSY, HOHAHA [512 x 2048 data matrix, zero-filled to 1024 data points in t_1; 64 or 72 scans per t_1 value; spectral width 1805 Hz; recycle delay 1.0 s; unshifted sine-bell filtering in t_1 and t_2 (COSY) and shifted sine-squared filtering in t_1 and t_2 (HOHAHA)]. HMQC [512 x 4096 data matrix, zero-filled to 1024 data points in t_1; spectral width 11 067 Hz in t_1 and 1805 Hz in t_2; 64 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.46 ms; shifted sine-squared filter]. HMBC [256 x 4096 data matrix, zero-filled to 1024 in t_1; spectral width 20 826 Hz in t_1 and 1805 Hz in t_2; 72 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.46 ms; shifted sine-squared filter].

4.3 STRUCTURE OF THE CAPSULAR ANTIGEN OF *Escherichia coli* O8:K50:H⁻

4.3.1 Abstract

The primary structure of the acidic capsular antigen of *Escherichia coli* O8:K50:H⁻ was shown by glycose analysis, methylation analysis and one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy to be composed of repeating linear tetrasaccharide units having the structure:

→3)-ß-L-Rha*p*-(1→4)-*α*-D-Glc*p*NAc-(1→2)-ß-D-Man*p*-(1→3)-*α*-D-Man*p*NAc-(1→
4,6
V
Pyr (*S*)

4.3.2 Introduction

The *E. coli* K50 capsular antigen is co-expressed with O group 8[44] and contains amino sugars, thereby placing it in the Group I polysaccharides[64]. *E. coli* K50 has been implicated in appendicital peritonitis[44]. It is interesting to note that the polysaccharide has no uronic acid, its only acidic function being pyruvate. To date only two other polysaccharides of this type have been isolated, from *E. coli* K37[57] and K47[58].

The pyruvic acid ketal is often part of the immunological determinant, and pyruvate substituents have been found to act as determinants in the symbiotic association between certain microorganisms and their host plants[189]. The immunological specificity is affected by the configuration of the monosaccharide to which the pyruvate is linked, the positions of the two hydroxyl groups that are bridged by the ketal, and the configuration of the methyl and carboxyl groups at the asymmetric ketal carbon atom[295].

4.3.3 Results and Discussion

Isolation, composition, and 1D NMR spectra of the capsular antigen.— *E. coli* O8:K50:H⁻ bacteria were grown on Mueller-Hinton agar and the acidic capsular polysaccharide (**PS**) was isolated and purified by precipitation with cetyltrimethylammonium bromide. The

polysaccharide was further purified by ion-exchange chromatography on DEAE-Sepharose™ CL-6B. The average molecular mass of the **PS** was found to be M_r 2.63 x 10^6 on a dextran-calibrated column of Sephacryl™ S-500. Hydrolysis of the **PS** followed by GLC-MS examination of the derived alditol acetates showed that Rha, Man, GlcN, and ManN were present in equimolar quantities. Methanolysis of the **PS**, treatment of the products with $NaBH_4$ to effect carboxyl reduction, followed by hydrolysis, and GLC-MS examination of the derived alditol acetates gave the same result as before, indicating the absence of a uronic acid in the polymer. GLC analysis of the derived acetylated (-)-2-octyl glycosides[149] of the **PS** showed the configuration of Rha to be L and that of the other constituent sugars to be D.

The ^{1}H NMR spectrum of the **PS** (Figure 4.13) in D_2O contained H-1 signals at δ 5.12 ($^3J_{1,2}$ 3.7 Hz), 5.05, and 4.89 (2H), signals for the methyl protons for two NAc groups at δ 2.08 and 2.07, a signal for H-6 of a deoxy sugar at δ 1.33 ($^3J_{5,6}$ 5.4 Hz), and a signal for the methyl protons of pyruvate at δ 1.56. The ^{13}C NMR data complemented the ^{1}H NMR results and confirmed a pyruvylated tetrasaccharide repeating unit for the **PS**, with signals at 101.57, 100.76, 99.43, 97.86, and 96.18 ppm in the anomeric region (95 - 105 ppm), one being that of the acetal carbon of the pyruvate. Signals for carbonyl carbons occurred at 175.46, 174.89, and 174.07 ppm, while signals at 54.96 and 49.79 ppm indicated the presence of two C-N bonds. The origin of the additional minor signals in the ^{13}C spectrum, indicating heterogeneity, were not investigated further. The ^{1}H and ^{13}C NMR spectra are shown in Figures 4.13 and 4.14 respectively and NMR data for the polysaccharide are listed in Table 4.5.

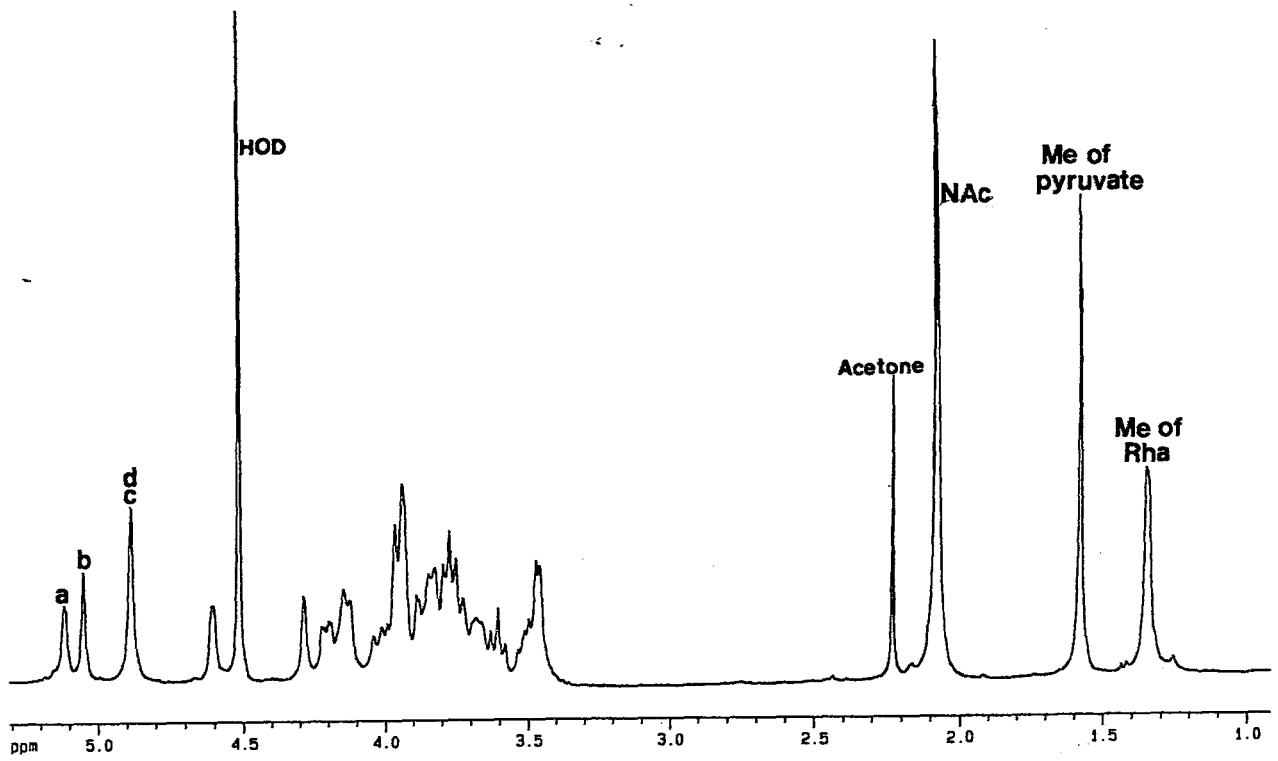


Figure 4.13 ^{1}H NMR spectrum of the *E. coli* K50 capsular polysaccharide (400 MHz, 323K)

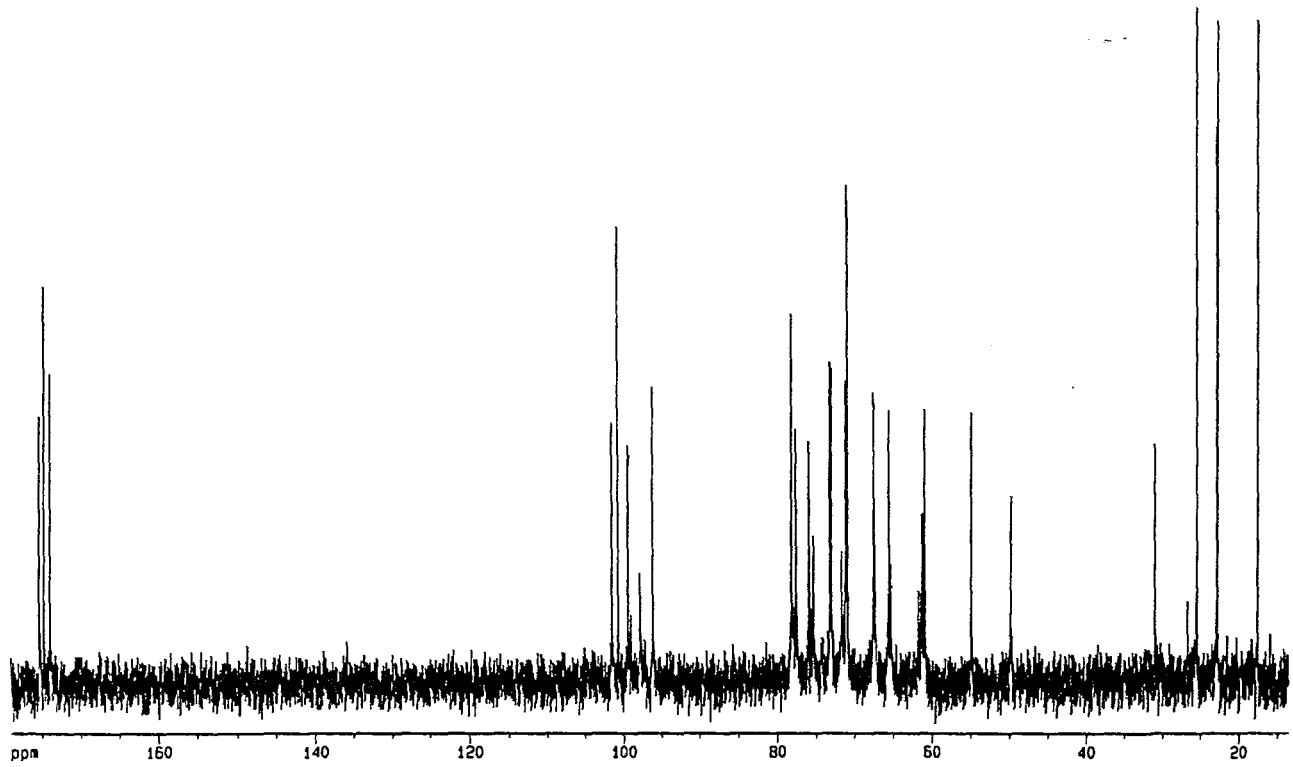


Figure 4.14 ^{13}C NMR spectrum of the *E. coli* K50 capsular polysaccharide

Methylation analysis.— Methylation of the polysaccharide followed by GLC and GLC-MS analysis of the permethylated alditol acetates derived from the products of an acid hydrolysate showed the presence of 2-deoxy-2-methylacetamido-3,6-di-*O*-methylglucose, 2-deoxy-2-methylacetamido-4,6-di-*O*-methylmannose, 2,4-di-*O*-methylrhamnose, and 3-*O*-methylmannose. These results indicated the presence of 4-substituted GlcNAc, 3-substituted ManNAc, 3-substituted Rha, and 3,4,6-substituted Man in a linear tetrasaccharide repeating unit with the pyruvate substituent attached to the Man residue.

2D NMR studies of the E. coli *K50 polysaccharide.*— The sequence of the residues in the repeating unit of the **PS** was established by 2D NMR experiments, which also confirmed the glycosylation sites in the polysaccharide. Carbon and proton resonances were established from COSY[259], 2D Homonuclear Hartmann-Hahn (HOHAHA)[281], HMQC[283], and HMQC-TOCSY[267] experiments. The residues in the repeating unit were labelled **a** - **d** in order of decreasing chemical shift of their anomeric protons (Figure 4.15).

Residue **a** [→4)-D-Glc*p*NAc]: The H-1/H-2 cross-peak was the only one clearly identifiable for this residue in the COSY spectrum (Figure 4.15). The resonances for C-1 and C-2 of this residue could be established from the H-1/C-1 and H-1/C-2 crosspeaks respectively in the HMQC-TOCSY experiment (Figure 4.17). The signal for H-2 could then be identified by correlation with the C-2 resonance from the HMQC spectrum (Figure 4.16), allowing assignment of the H-2/H-3 cross-peak in the COSY spectrum. The remaining proton resonances could then be assigned from the COSY and HOHAHA spectra. The remaining carbon resonances were assigned by comparing the ^{1}H assignments with the ^{1}H-^{13}C correlation data from the HMQC experiment.

Residue **b** [→3)-D-Man*p*NAc]: The ^{1}H resonances for residue **b** were traced readily via their cross-peaks in the COSY and HOHAHA spectra. Magnetism relayed well through this spin system in the HOHAHA experiment despite the small $^{3}J_{1,2}$ value, and all cross-peaks were clearly visible. The ^{13}C resonances were assigned from the HMQC spectrum as for residue **a**.

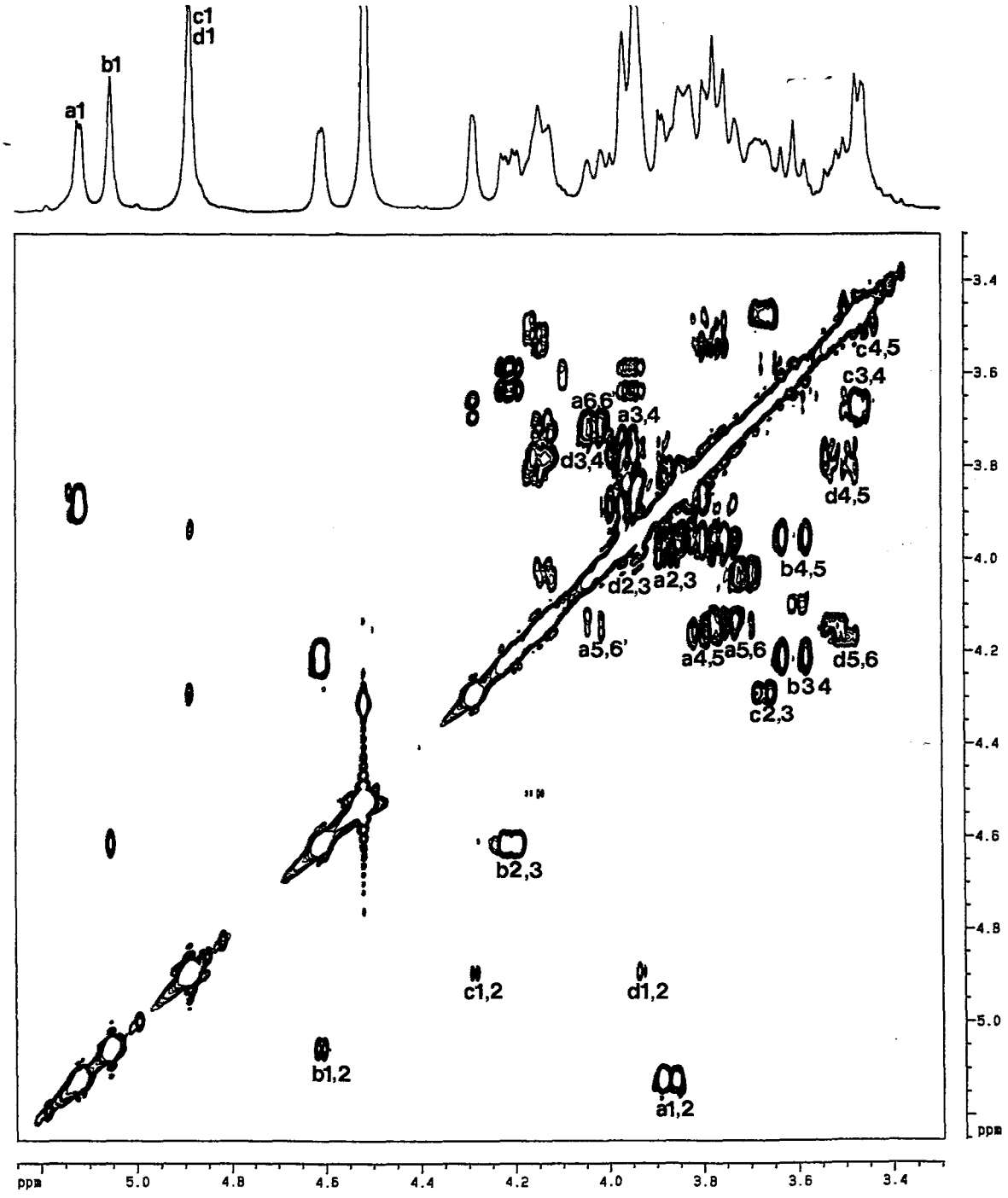


Figure 4.15 COSY contour plot of the region δ 3.30 - 5.25 of the K50 capsular polysaccharide. The 1D spectrum is projected along the f_2 axis. a1 connotes H-1 of residue a and a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc.

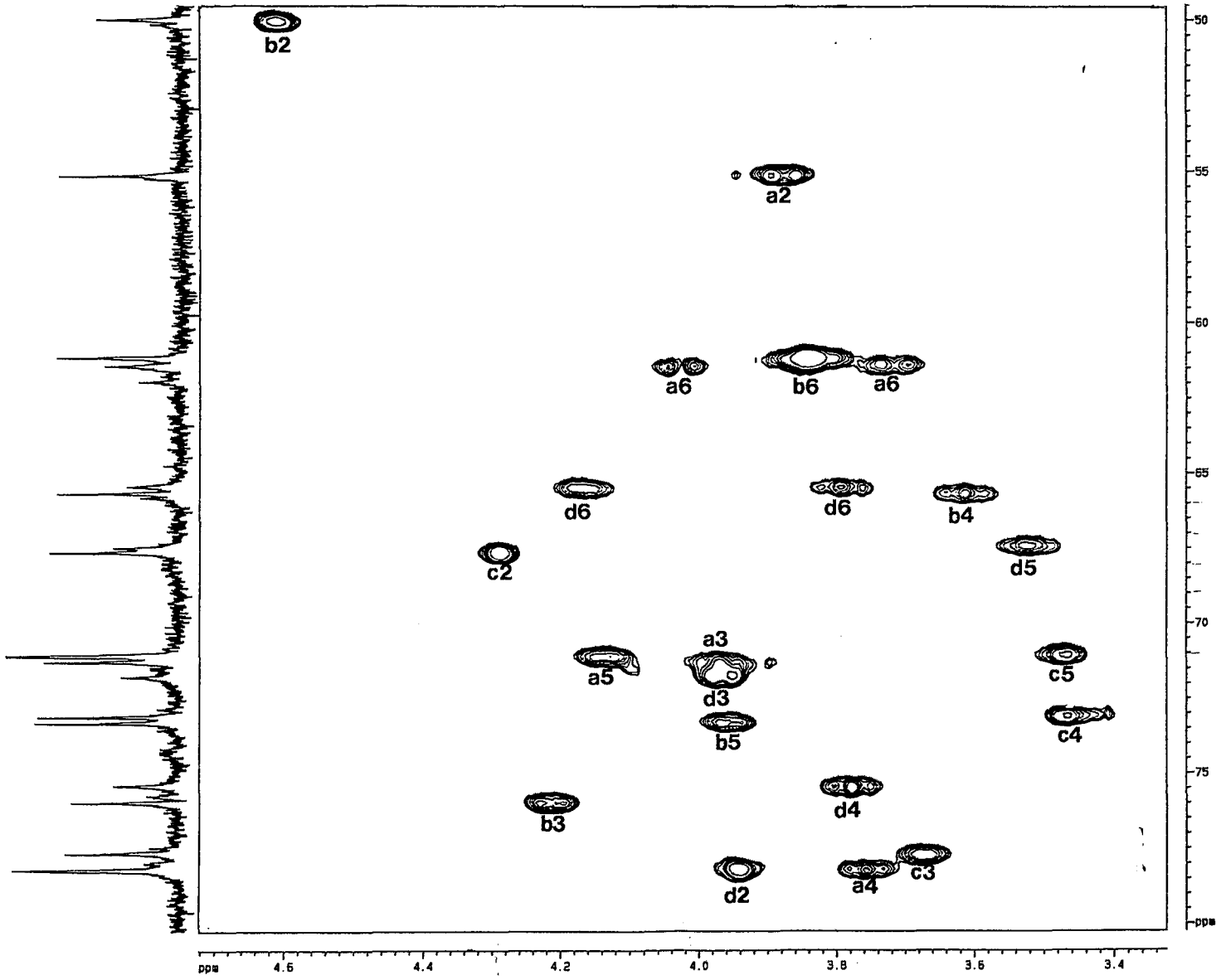


Figure 4.16 HMQC ^{1}H-^{13}C shift correlation map of the spectral region f_1 49 - 80 ppm (^{13}C) and f_2 3.34 - 4.27 (^{1}H) of the PS. The ^{13}C NMR spectrum is projected along the f_1 axis

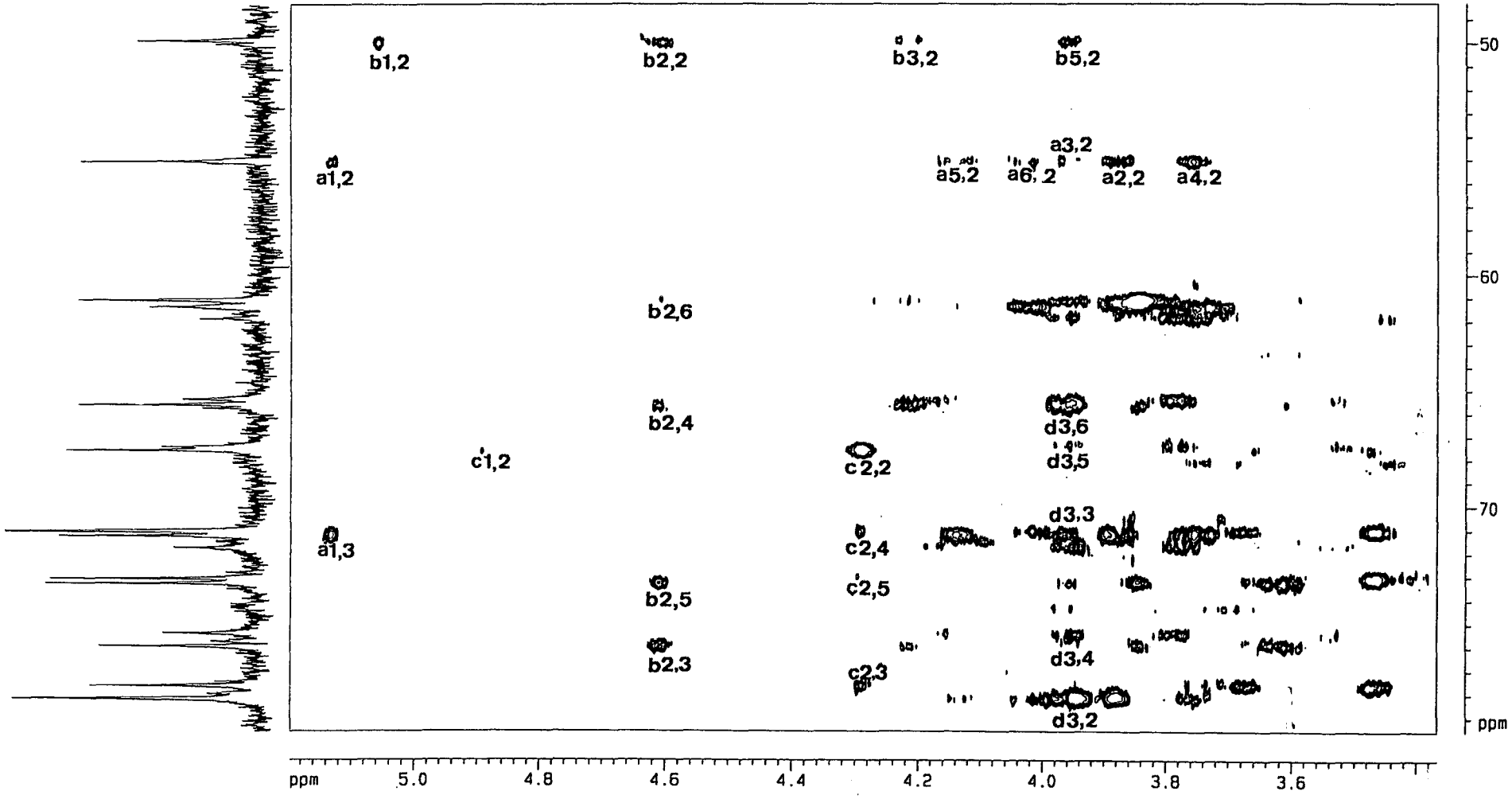


Figure 4.17 HMQC-TOCSY ^{1}H-^{13}C contour plot of the spectral region f_1 49 - 80 ppm (^{13}C) and f_2 δ 3.40 - 5.18 (^{1}H) for the PS. a1,2 connotes the cross-peak between H-1 and C-2 of residue a, etc. The ^{13}C NMR spectrum is projected along the f_1 axis

Residue **c** [→3)-L-Rha*p*]: Since the H-1 resonances of residues **c** and **d** exactly coincide (Table 4.5), the H-6 resonance at δ 1.33 was used as the starting point for tracing the ^{1}H resonances for residue **c** in the COSY and HOHAHA spectra. The ^{13}C resonances for residue **c**, with the exception of C-1, were then assigned from the ^{1}H-^{13}C correlation data obtained in the HMQC experiment.

Residue **d** [→2,4,6)-D-Man*p*]: The ^{1}H resonances for residue **d** were traced via their cross-peaks in the COSY spectrum. The connectivities were confirmed from the H-3 track in the HMQC-TOCSY spectrum, which showed relays to all the carbon signals for the residue, with the exception of C-1, and from the C-3 track which showed relays to H-2/6. The ^{13}C resonances thus obtained were confirmed from the HMQC spectrum.

The configurations of the glycosidic linkages were assigned by measuring the $^{2}J_{C\text{-}1,H\text{-}1}$ values obtained from a proton-coupled HMQC experiment. The values obtained were 175.7 Hz for residue **a**, 173.3 Hz for residue **b**, 163.6 Hz for residue **c** and 162.8 Hz for residue **d**, indicating α configurations for residues **a** and **b** and ß configurations for residues **c** and **d**[251].

The signal at 100.76 ppm in the ^{13}C spectrum which was absent from the HMQC spectrum, since it has no proton associated with it, could now be assigned to the acetal carbon of the pyruvate substituent. Comparison of the chemical shift data for residues **a** - **d** in Table 4.5 with those reported for methyl glycosides[144,254,284] permitted identification of residue **a** as 4-substituted α-D-GlcNAc, residue **b** as 3-substituted α-D-ManNAc, residue **c** as 3-substituted ß-D-Rha and residue **d** as 2,4,6-substituted ß-D-Man. These data are in agreement with the methylation results for **PS**.

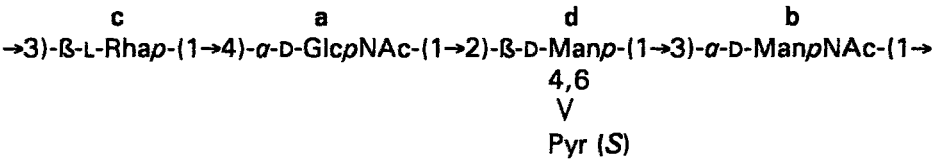
The sequence of the residues in the repeating unit was established by a Heteronuclear Multiple Bond Correlation (HMBC) experiment[195]. The first easily identifiable correlation was between C-1 of GlcNAc and H-2 of Man. A second correlation was visible between H-4 of GlcNAc and C-1 of either residue **c** or **d**. Since a link between GlcNAc and Man (residue **d**) had already been established, it was clear that this carbon signal belonged to residue **c**, thus enabling assignment

of the C-1 resonances for residues **c** and **d**. Other relevant correlations identified were: H-1 of ManNAc to C-3 of Rha, H-1 of Man to C-3 of ManNAc, C-1 of Man to H-3 of ManNAc, and H-6 and H-4 of Man to the acetalic carbon of the pyruvate moiety. It was also possible to assign the carbonyl signals to appropriate residues, since correlations were visible from H-2 of GlcNAc to the carbonyl resonance at 175.46 ppm, from H-2 of ManNAc to the carbonyl resonance at 174.89 ppm, and from the methyl protons of the pyruvate group to the carbonyl signal at 174.07 ppm. HMBC correlations are listed in Table 4.6.

The chiral carbon of the pyruvate moiety was assigned the *S* configuration based on the ^{1}H and ^{13}C chemical shifts of its methyl group[296]. This assignment accords with the recent finding of Miertus *et al.* that pyruvate acetals 4,6-linked to D-Glc*p* and D-Man*p* residues prefer the *S* configuration[189]. This preference is believed to be related to conformational stability.

4.3.4 Conclusion

The combined chemical and NMR data permit the structure of the tetrasaccharide repeating unit of the *E. coli* K50 capsular polysaccharide to be written as:



This is only the third *E. coli* capsular polysaccharide found to contain a pyruvate moiety as its sole acidic function[57,58]. Four capsular polysaccharides having phosphate groups as their sole acidic function have also been reported[60,67,297,298]. In common with three of the six other *E. coli* capsular antigens which contain a pyruvate substituent, the pyruvate acetal in the K50 capsular polysaccharide is 4,6-linked[48]. In addition, it is only the second capsular polysaccharide in the *E. coli* series to date found to contain a ManNAc residue, the other being *E. coli* K84, which contains 3-linked ß-D-ManNAc[50].

Table 4.5 **NMR data[a] for the *E. coli* K50 polysaccharide**

Residue		Proton or Carbon											
		1	2	3	4	5	6	6′	NAc CH_3	NAc C=O	Pyr CH_3	Pyr COOH	Pyr C-2
a													
→4)-α-D-Glc*p*NAc	H	5.12	3.87	3.97	3.76	4.14	4.03	3.72	2.07				
	C	99.43	54.96	71.17	78.16	70.97	61.30		22.84[b]	175.46			
b													
→3)-α-D-Man*p*NAc	H	5.05	4.61	4.21	3.61	3.95	3.84	3.84	2.08				
	C	96.18	49.79	75.89	65.55	73.22	61.00		22.99[b]	174.89			
c													
→3)-ß-D-Rha*p*	H	4.89	4.29	3.67	3.47	3.45	1.33						
	C	101.57	67.52	77.60	70.97	73.02	17.61						
d													
→2)-ß-D-Man*p*	H	4.89	3.93	3.96	3.78	3.52	4.15	3.80					
4,6 V Pyr	C	97.86	78.16	71.69	75.34	67.50	65.32						
Pyr	H										1.56		
	C										25.55	174.07	100.76

[a]Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 ppm for ^1H and ^{13}C respectively. [b]These values are interchangeable.

Table 4.6 **Two- and three bond ^{1}H-^{13}C correlations for the PS**

Residue		Proton	Correlation to
a	→4)-*α*-D-Glc*p*NAc	H-1	78.16(d;C-2), 71.17(a;C-3), 70.97(a;C-5)
		H-2	99.43(a;C-1), 175.46(C=O)
		H-3	54.96(a;C-2)
		H-4	101.57(c;C-1), 70.97(a;C-5)
b	→3)-*α*-D-Man*p*NAc	H-1	77.60(c;C-3), 75.89(b;C-3), 73.22(b;C-5)
		H-2	96.18(b;C-1), 75.89(b;C-3), 65.55(b;C-4), 174.89(C=O)
		H-3	97.86(d;C-1), 65.55(b;C-4)
c	→3)-ß-D-Rha*p*	H-1	78.16(a;C-4), 77.60(c;C-3), 67.52(c;C-2)
		H-2	77.60(c;C-3), 70.97(c;C-4)
		H-4	73.02(c;C-5)
d	→2)-ß-D-Man*p*	H-1	75.89(b;C-3), 71.69(d;C-3), 67.50(d;C-5)
	4,6	H-2	99.43(a;C-1), 71.69(d;C-3), 75.34(d;C-4)
	V	H-4	100.76(Pyr;C-2), 78.16(d;C-2)
	Pyr	H-6	100.76(Pyr;C-2)
Pyr		CH_3	174.07(Pyr;C=O)

4.3.5 Experimental

General methods.— Instrumentation used was as described for *E. coli* K101. A J & W Scientific fused-silica DB-17 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was used for separating partially methylated alditol acetates (programme I), and alditol acetates and acetylated octyl glycosides (programme II). A J & W Scientific fused-silica DB-Wax bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.15 μm) was used for separating alditol acetates of ManNAc, GlcNAc and GalNAc (130 kPa, 240°C isothermal). A J & W Scientific DB-225 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was also used for separating acetylated octyl glycosides (130 kPa, 240°C isothermal). The temperature programmes used were: I, 180°C for 2 min, then 3°C.min^{-1} to 240°C, 100 kPa; II, 180°C for 2 min, then 2°C.min^{-1} to 240°C, 100 kPa. The identities of all derivatives were determined by comparison with authentic standards and confirmed by GLC-MS. GPC was performed on a dextran-calibrated column (1.6 x 65 cm) of Sephacryl™ S500 using 0.1 M NaOAc buffer (pH 5.00) as eluent. Material was detected by refractive index.

Isolation and purification of the K50 polysaccharide.— An authentic culture of *E. coli* O8:K50:H$^-$ was obtained from Dr I. Ørskov (Copenhagen) and propagated on Mueller-Hinton agar as described for *E. coli* K101. The capsular polysaccharide was extracted with aqueous 1% phenol and separated from cellular debris by ultracentrifugation (35 000 rpm, 3 h). The supernatant was dialysed overnight against running water (12-14 000 mw cut-off) and freeze-dried. The buff-coloured powder obtained was reconstituted in a minimum quantity of water (100 mL) and the K-antigen isolated by precipitation with cetyltrimethylammonium bromide. Purification was achieved by gel permeation-ion exchange chromatography on a DEAE-Sepharose™ CL-6B column (2.6 x 27 cm) using gradient elution with 0 - 1 M NaCl in 0.01 M Tris-HCl (pH 8.50) as eluent. Fractions were assayed for carbohydrate using the phenol-H_2SO_4 reagent[85]. The yield of capsular polysaccharide was 470 mg.

Monosaccharide composition, absolute configuration, and methylation analysis.— Glycose analysis and methylation analysis were performed as for the *E. coli* K101 polysaccharide. Absolute configurations were determined by GLC analysis of the derived (-)-2-octyl glycosides. Since there was no uronic acid present in the polysaccharide, the carboxyl reduction step was omitted. The purified polysaccharide (10 mg) was hydrolysed with 4 M TFA (125°C, 1 h) and re-N-acetylated as described for the K45 polysaccharide. Acetylated (-)-2-octyl glycosides were then prepared as described for the K101 polysaccharide.

NMR Spectroscopy.— Samples were prepared as described previously. Spectra were recorded at 50°C on a Bruker AMX-400 spectrometer equipped with an X32 computer. The parameters used for 2D experiments were as follows: COSY [256 x 2048 data matrix, zero-filled to 1024 data points in t_1; 80 scans per t_1 value; spectral width 2007.9 Hz; recycle delay 1.0 s; unshifted sine-bell filtering in t_1 and t_2]. HOHAHA [256 x 4096 data matrix, zero-filled to 1024 data points in t_1; 96 scans per t_1 value; spectral width 2007.9 Hz; recycle delay 1.0 s; mixing time 89 ms; shifted sine-squared filtering in t_1 and t_2]. HMQC, HMQC-TOCSY, HMBC [512 x 4096 data matrix, zero-filled to 1024 data points in t_1; 52, 56 or 64 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.45 ms; mixing time 89 ms (HMQC-TOCSY); spectral width in t_1 11 068 Hz (HMQC and HMQC-TOCSY), 20 828 Hz (HMBC) and in t_2 2008.0 Hz (HMQC and HMQC-TOCSY) and 2016.13 Hz (HMBC); shifted sine-squared filtering in t_1 and t_2].

4.4 STRUCTURAL INVESTIGATION OF THE CAPSULAR POLYSACCHARIDE OF *Escherichia coli* O101:K103:H⁻ USING BACTERIOPHAGE DEGRADATION AND NMR SPECTROSCOPY

4.4.1 Abstract

NMR spectroscopy was performed on the depyruvylated capsular antigen of *E. coli* K103 and on the oligosaccharide obtained by depolymerisation of the native polysaccharide with a viral-borne endoglycanase. This capsular polysaccharide is the only one to be co-expressed with O group 101 and joins a small group of unusual capsular polysaccharides which possess pyruvic acid as the only acidic function. The primary structure was shown to be composed of the repeating unit:

→3)-*α*-D-Gal*p*-(1→4)-*α*-D-Gal*p*-(1→3)-*α*-L-Rha*p*-(1→
2,3
V
Pyr (*S*)

4.4.2 Introduction

Escherichia coli O101:K103:H⁻ is unique in that it is the only member of the O101 series which produces a capsular polysaccharide. The capsular antigen K99, also co-expressed with O101, has been found to be proteinaceous[44]. *E. coli* capsular antigens have been subdivided into two groups based on their physical, chemical, and microbial characteristics[44]. Co-expression with O-antigen 101 places the K103 capsular antigen among the Group I polysaccharides[44]. Unlike the other polysaccharides of this group it has a low molecular mass. It is, however, acidic and its structure is *Klebsiella*-like, both characteristics of Group I polysaccharides. This capsular polysaccharide is thus a candidate for possible inclusion in the intermediate Group I/II proposed by Finke *et al.*[70]. *E. coli* O101:K103 has been implicated in human appendicitis and enterotoxigenic diarrhoea, and causes septicaemia in cattle[44,46].

4.4.3 Results and Discussion

Isolation, composition, and 1D NMR spectra of the capsular antigen.— *E. coli* K103 bacteria were grown on Mueller-Hinton agar and the acidic capsular polysaccharide (**PS**) was isolated by extraction of the cells using aqueous 1% phenol, and was purified by precipitation with cetyltrimethylammonium bromide. Gel permeation chromatography (GPC) of the **PS** on Sephacryl™ S500 showed that it was polydisperse, with the two main fractions having average molecular masses M_r 1.32 x 10^7 and M_r 3.63 x 10^4 (Figure 4.18). Sugar analysis showed the two fractions to be identically constituted. Polydispersity has been reported for other *E. coli* capsular polysaccharides and has been ascribed to the formation of large micellar aggregates by these lipid-bound polysaccharides[71,298]. When the polysaccharides are treated with dilute acid the labile phosphodiester bridges linking the polysaccharide chains to lipid are hydrolysed, the micelles are broken down and much lower molecular masses are observed on GPC (see formation of depyruvylated polysaccharide **DPS** later; Figure 4.19). Hydrolysis of the **PS** followed by GLC-MS examination of the derived alditol acetates showed that Gal and Rha were present in the molar ratio 2:1. Methanolysis of the **PS** followed by reduction of the methoxycarbonyl groups, hydrolysis, and GLC examination of the derived alditol acetates showed no alteration of the sugar ratio, indicating the absence of uronic acid. Reaction of the products of methanolysis of the **PS** with trifluoroacetic anhydride, followed by GLC analysis on a chiral FS-Lipodex™ A column[154] established the configuration of both Gal residues as being D and that of Rha as L.

The ^{1}H NMR spectrum of the sodium salt of the **PS** in D_2O (Figure 4.20a) contained H-1 signals at δ 5.61, 5.09, and 5.02, a signal for the methyl protons of pyruvate at δ 1.59, and a signal for H-6 of a deoxy sugar at δ 1.33. The acid sensitivity of the pyruvate substituent, particularly at elevated temperatures, precluded the NMR spectroscopic examination of the **PS** in its acidic form. Interpretation of the NMR spectra of the sodium form of the **PS** proved difficult due to a significant degree of line broadening caused by the high viscosity of the sample. The **PS** was therefore depyruvylated by treatment with 1% aqueous acetic acid at 100°C for 1 h, followed by purification by GPC on Sephacryl™ S400. All subsequent analyses and 2D NMR spectroscopic studies were

carried out on the depyruvylated polysaccharide (DPS) thus obtained. The DPS was found by GPC on Sephacryl™ S400 to have a M_r of 2.2 x 10^4 (Figure 4.19). The ^{1}H NMR spectrum of the DPS (Fig. 4.20b) confirmed the presence of three monosaccharides in the repeating unit, one being a 6-deoxy sugar. The ^{13}C NMR spectrum complemented the ^{1}H NMR data with signals for C-1 at 102.63, 101.23, and 96.67 ppm, and a signal for C-6 of a deoxy sugar at 17.60 ppm.

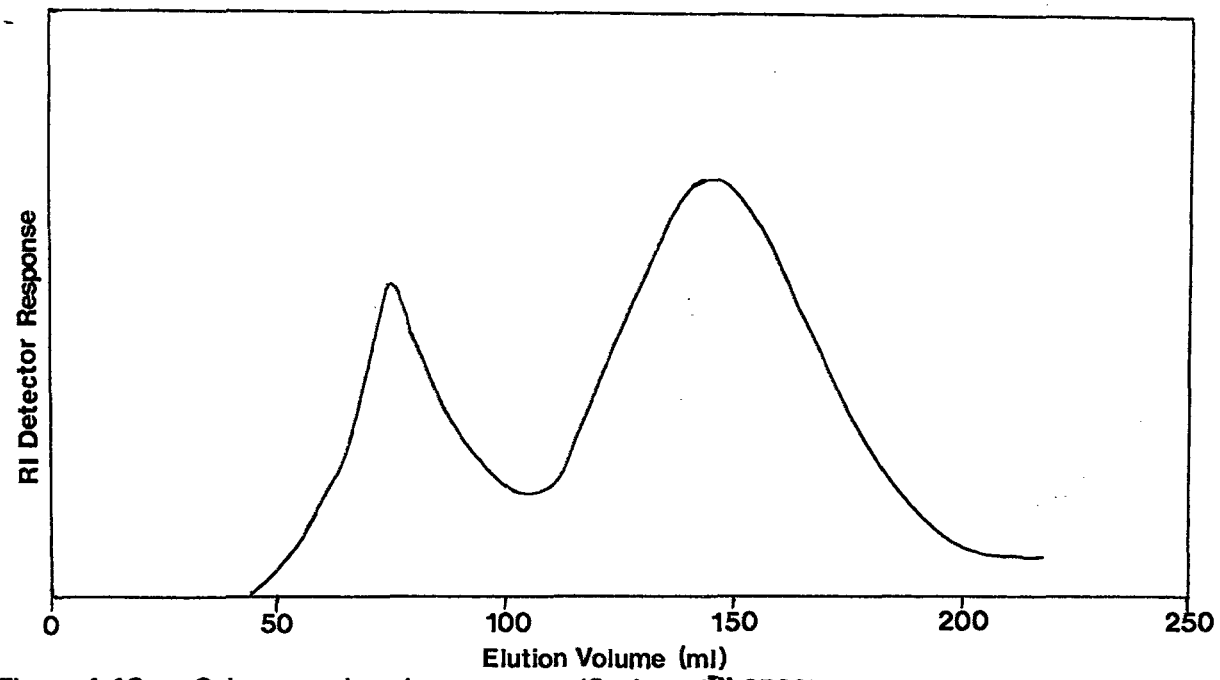


Figure 4.18 Gel-permeation chromatogram (Sephacryl™ S500) of the PS

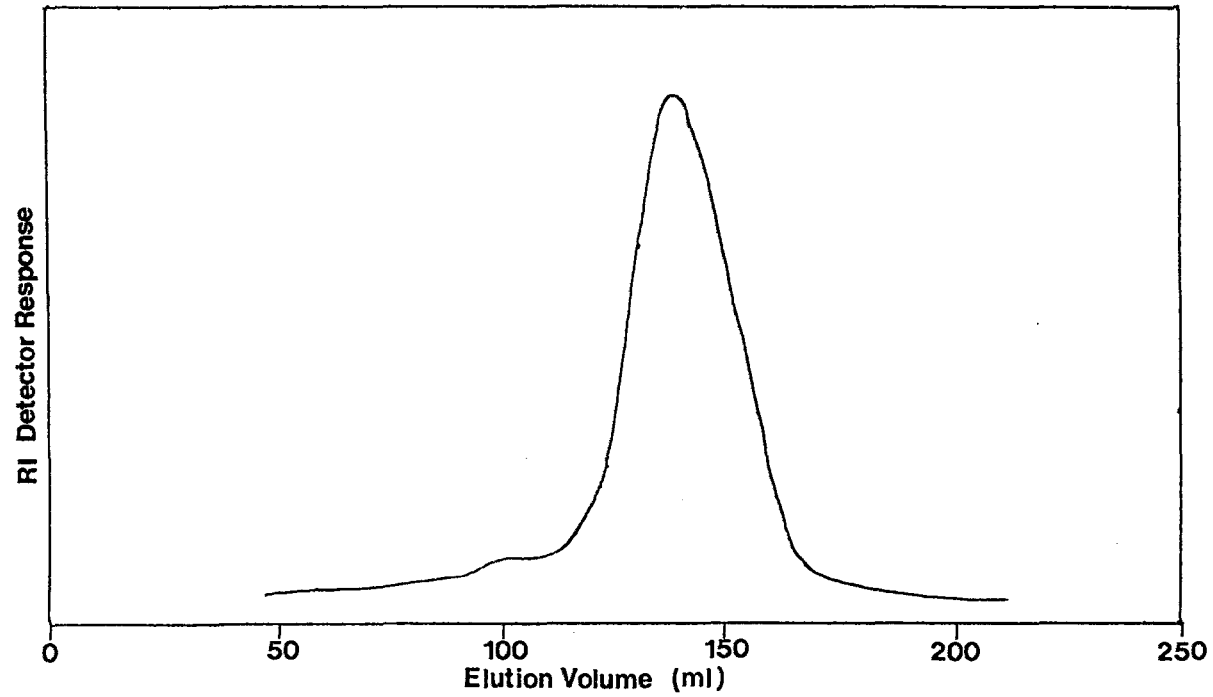


Figure 4.19 Gel-permeation chromatogram (Sephacryl™ S400) of the DPS

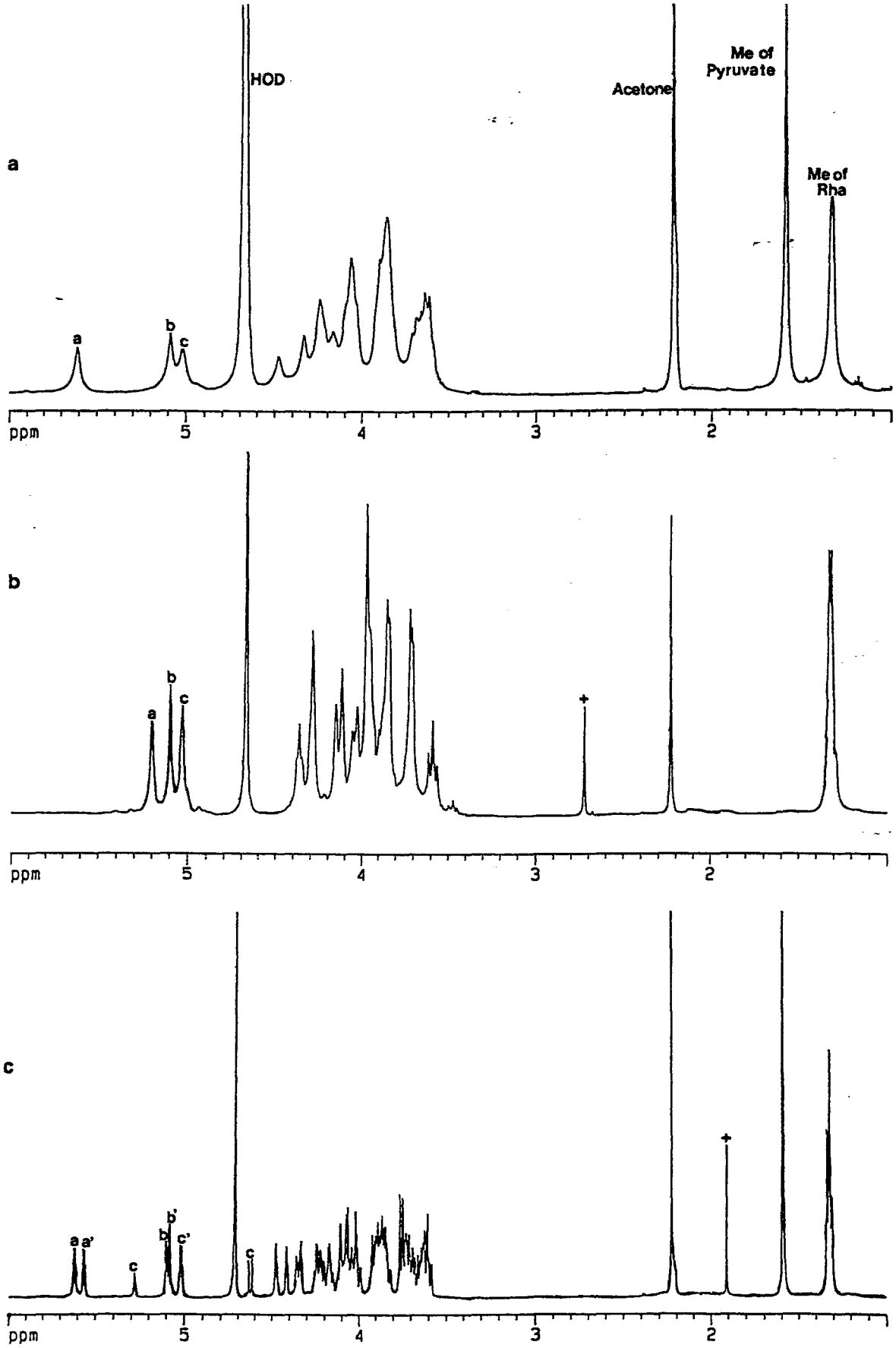
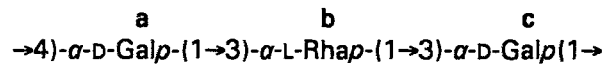


Figure 4.20 ^{1}H NMR spectra of a, K103 polysaccharide PS; b, depyruvylated K103 polysaccharide DPS; c, phage degradation product P2 (400 MHz, 303K).
+Unassigned signals

Methylation analysis.— Methylation analysis of the **DPS** showed the presence of 2,4-di-*O*-methylrhamnose, 2,4,6-tri-*O*-methylgalactose, and 2,3,6-tri-*O*-methylgalactose, indicating that the **DPS** contained 3-linked Rha, 3-linked Gal, and 4-linked Gal. The **DPS** was incompletely methylated by the Hakomori procedure[169] and the partially methylated polysaccharide was therefore remethylated by the method of Kuhn[179].

2D NMR studies of the E. coli *K103 DPS*.— The sequence of the residues in the repeating unit was established by 2D NMR experiments, which also confirmed the identity of the constituent residues and the glycosylation sites in the polysaccharide. The residues in the repeating unit were labelled **a** - **c** in order of decreasing chemical shift of their anomeric protons, as shown in Figure 4.20b. The ^{1}H resonances of all three residues were readily traced via their cross-peaks in the COSY[259] and 2D Homonuclear Hartmann-Hahn (HOHAHA)[261] spectra, while the ^{13}C resonances were assigned by comparing the ^{1}H assignments with the ^{1}H-^{13}C correlation data obtained from an HMQC[283] experiment. These data are shown in Table 4.7. Comparison of the chemical shift data for residues **a** - **c** with those reported for methyl glycosides[144,254] permitted identification of residue **a** as 4-linked Gal, residue **b** as 3-linked Rha, and residue **c** as 3-linked Gal. The sequence of the residues in the repeating unit was established by an HMBC[195] experiment. Correlations between H-1 of **a** and C-3 of **b**, between H-1 of **b** and C-3 of **c**, and between H-1 of **c** and C-4 of **a** were clearly visible. The structure of the repeating unit of the **DPS** is thus:



The position of the pyruvate acetal in the repeating unit of the **PS** and the configuration of its chiral carbon atom were established from a study of the hexasaccharide isolated from the bacteriophage depolymerisation the **PS**.

Table 4.7 NMR data[a] for *E. coli* K103 DPS

Residue		Proton or carbon						
		1	2	3	4	5	6	6'
→4)-α-D-Gal (a)	H	5.20	3.95	4.05	4.15	4.28	3.85	3.85
	C	96.67	69.31	69.89	79.57	71.81	60.98	
→3)-α-L-Rha (b)	H	5.10	4.28	3.96	3.58	3.87	1.32	
	C	102.63	67.71	76.49	71.20	70.03	17.60	
→3)-α-D-Gal (c)	H	5.05	3.97	3.98	4.13	4.36	3.72	3.72
	C	101.23	68.78	77.85	69.70	71.81	61.33	

[a]Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 ppm for ^1H and ^{13}C, respectively.

Bacteriophage-mediated degradation of the PS.— A bacteriophage isolated from sewage water and propagated on the K103 bacteria was used to depolymerise the **PS**. The hexasaccharide (**P2**) obtained, representing two repeating units of the polysaccharide, was purified by GPC on Sephacryl™ S200 and Bio-Gel™ P4. The ^1H NMR spectrum of **P2** (Figure 4.20c) contained five H-1 signals each integrating for one proton, and two fractional H-1 signals in the ratio of 1:3 integrating for one proton. These fractional signals at δ 5.25 and 4.63, relating to the α and ß forms of the reducing end of **P2**, were further identified as Gal by their $^3J_{1,2}$ values of 3.1 and 7.4 Hz[144,254]. The anomeric signals were labelled, in order of decreasing chemical shift, as **a**, **a'**, $\mathbf{c}_\alpha$, **b**, **b'**, **c'**, and $\mathbf{c}_\beta$, in order to relate them to the anomeric resonances of the residues already identified in the **DPS**. Figure 4.20 shows the ^1H NMR spectra of the **PS**, the **DPS**, and **P2** respectively. The H-1 signal for **b** showed slight broadening, as did the H-6 signal for this residue, due to its proximity to the reducing end of the oligosaccharide. A signal for the methyl protons of pyruvate occurred at δ 1.59 (2 x CH_3) and signals for H-6 of the deoxy sugars occurred at δ 1.33 and 1.32 respectively. The ^{13}C spectrum of **P2** contained C-1 signals at 102.57, 102.49, 100.48, 96.89 (fractional), 94.14, 93.92, and 92.90 (fractional) ppm as well as two signals for the pyruvic acetalic carbons at 108.72 and 108.60 ppm. Methyl signals for deoxy sugars occurred at 17.42 and 17.36 ppm, and for pyruvate at 23.37 and 23.17 ppm. Carbonyl signals for pyruvate occured

at 177.18 ppm and 176.45 ppm. Due to the acid lability of the pyruvate substituent, no methylation analysis was carried out on **P2**.

2D NMR studies of P2.— The sequence of the residues in the **P2** unit was established by 2D NMR experiments in the same way as for the **DPS**, using COSY, HOHAHA, DEPT-HMQC[300] (acquired with a read pulse angle of 60° to show CH_2 peaks negative and CH and CH_3 peaks positive), HETCOR[301], HMQC-TOCSY[267], and HMBC[195] experiments. The COSY, HETCOR, and HMQC-TOCSY spectra are shown in Figures 4.21, 4.22, and 4.23 respectively. NMR data for **P2** are listed in Table 4.8.

Residues **a** [→2,3,4)-*α*-D-Gal] and **a'** [2,3-*α*-D-Gal]: The ^1H resonances for both of these residues were easily traced via their crosspeaks in the COSY and HOHAHA spectra. The H-4/5 crosspeaks were small, as expected for Gal residues, due to small $^3J_{4,5}$ values. Carbon resonances were assigned by comparing the ^1H assignments with the ^1H-^{13}C correlation data obtained from both the HMQC and HETCOR experiments. It was necessary to use both of these correlation experiments due to the amount of signal overlap occurring in various regions of both spectra. The only marked difference in the chemical shift values for the two residues was in those for the C-4 resonance. Residue **a** is linked at this position and the carbon resonance is accordingly at a much lower field than that for residue **a'**.

Residues **b** and **b'** [→3)-*α*-L-Rha*p*]: The ^1H and ^{13}C resonances for these two residues were very similar. The ^1H anomeric signal for residue **b** showed slight broadening due to its proximity to the reducing end of **P2**. All ^1H resonances were assigned readily from the COSY spectrum despite the partial overlap from the H-3/4 crosspeak onwards, and the carbon resonances were assigned as for residues **a** and **a'**.

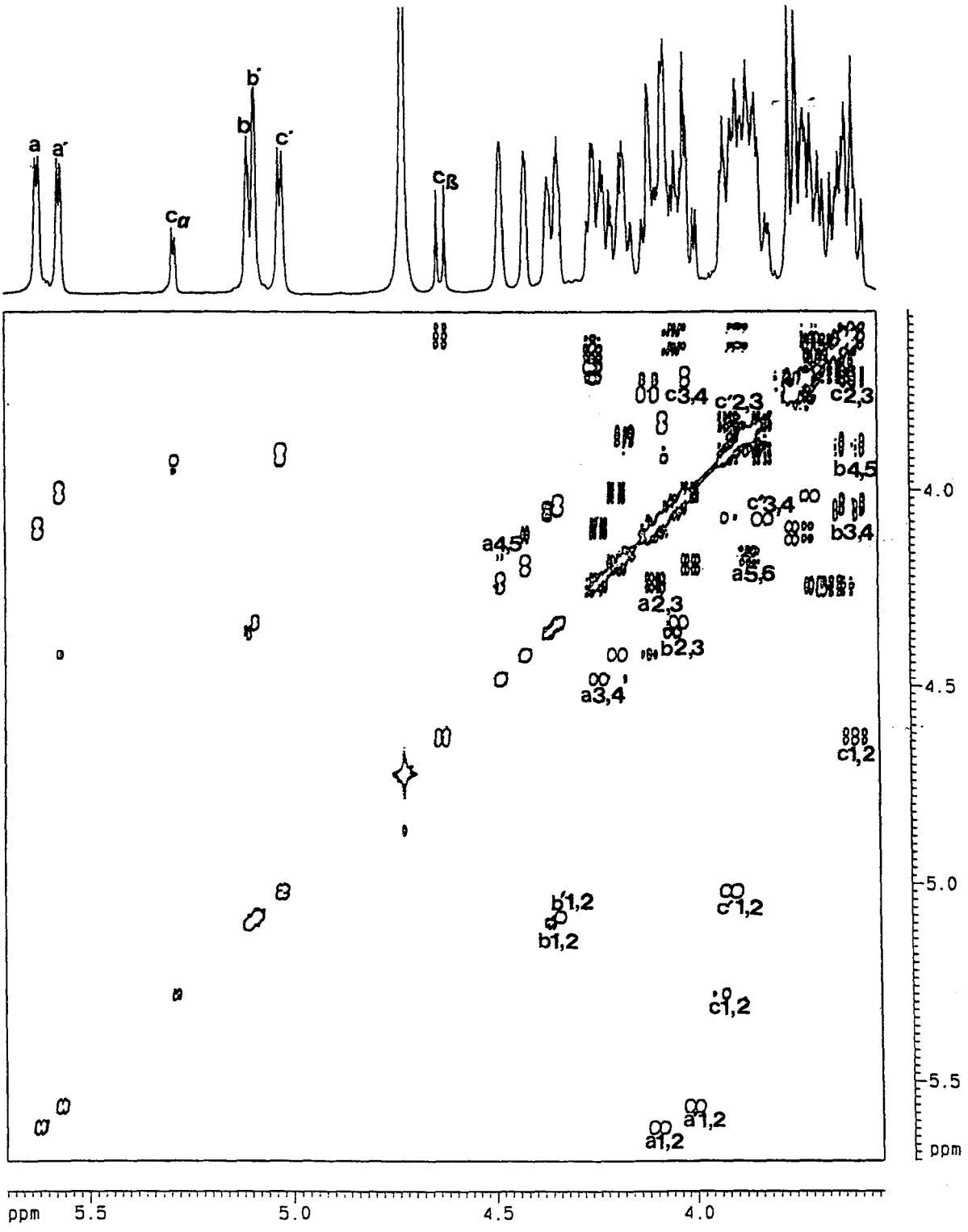


Figure 4.21 COSY contour plot for P2 showing the region δ 3.5 - 5.7 for f_1 and f_2. a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc. The ^{1}H NMR spectrum is projected along the f_2 axis.

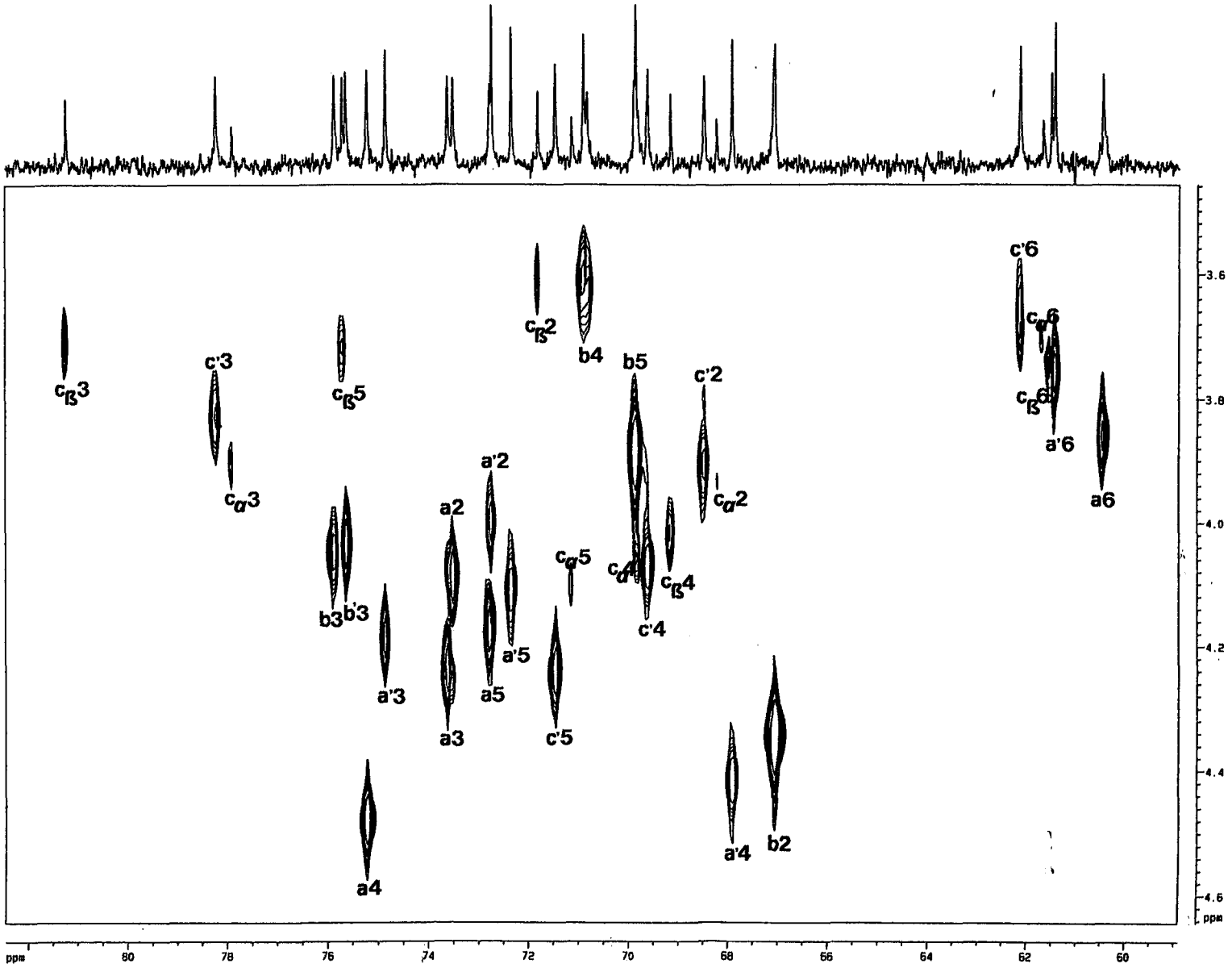


Figure 4.22 HETCOR contour plot for P2 showing the region f_1 δ 3.44 - 4.64 (^{1}H) and f_2 59 - 82 ppm (^{13}C). The ^{13}C NMR spectrum is projected along the f_2 axis.

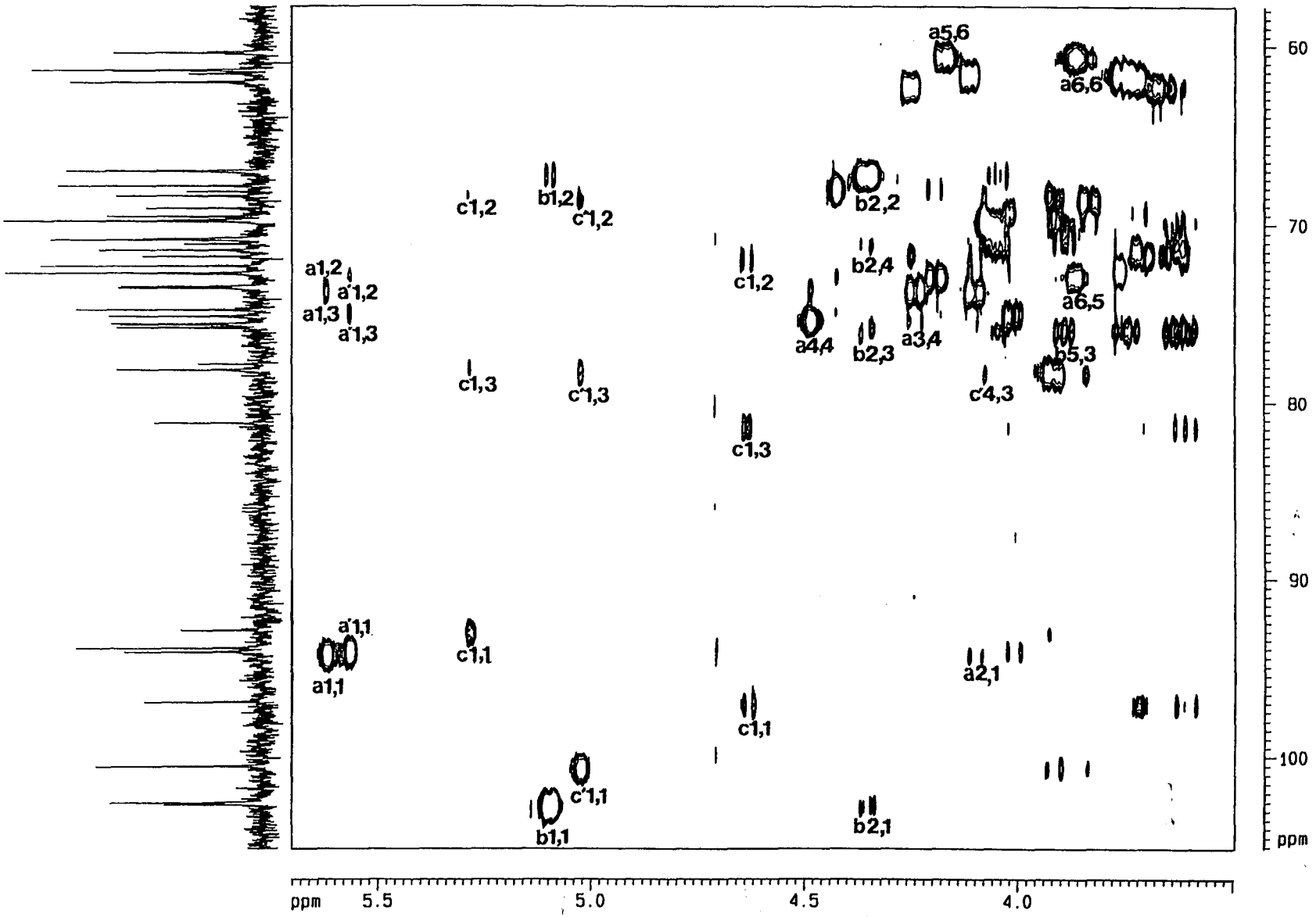


Figure 4.23 HMQC-TOCSY ^{1}H-^{13}C contour plot of the spectral regions f_1 56 - 105 ppm (^{13}C) and f_2 δ 3.50 - 5.70 (^{1}H) for P2 The ^{13}C NMR spectrum is projected along the f_1 axis.

Residue **c'** and $\mathbf{c}_{\alpha}$ [→3)-*α*-D-Gal*p*] and [→3)-*α*-D-Gal*p*OH]: The COSY and HOHAHA coupling patterns for these two residues were very similar. ^{1}H resonances for both residues could be traced as far as H-4 from the crosspeaks in the COSY spectrum, and H-5 and H-6 for **c'** could be identified from the HOHAHA spectrum as a small H-4/5 crosspeak was present. However, no H-4/5 crosspeak was visible for $\mathbf{c}_{\alpha}$. Since $\mathbf{c}_{\alpha}$ is present in smaller quantity than the other residues, the ^{1}H and ^{13}C peaks for this residue are significantly smaller than those for the other residues. Once all other ^{13}C signals had been assigned, it was possible to identify C-5 and C-6 signals for $\mathbf{c}_{\alpha}$ from the ^{13}C spectrum using their comparative heights for confirmation, the corresponding ^{1}H resonances being assigned from the HETCOR spectrum by correlation. Other ^{13}C resonances were also assigned by correlation using both the HETCOR and HMQC spectra.

Residue $\mathbf{c}_{\beta}$ [→3)-ß-D-Gal*p*OH]: The COSY and HOHAHA coupling patterns for this residue differ markedly from those for residues **a**, **a'**, **c'** and $\mathbf{c}_{\alpha}$, due to the ß-anomeric configuration. ^{1}H resonances could be identified as far as H-4 from both the COSY and HOHAHA spectra and ^{13}C resonances were assigned from the HMQC and HETCOR data by correlation. The C-5 and C-6 resonances were assigned by difference, permitting the H-5 and H-6 signals to be identified from the ^{1}H-^{13}C correlation data. A further distinction between residue $\mathbf{c}_{\beta}$ and the other Gal residues occurred in the shift position for C-5, which occurs downfield relative to the C-5 signals for the *α*-Gal residues, as expected[254].

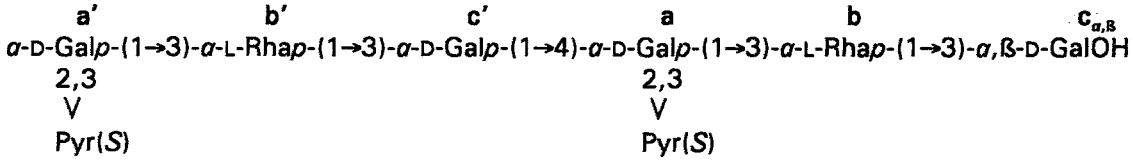
The sequence of the residues in **P2** was determined using an HMBC experiment. Clear correlations could be seen between H-1 of **a** and C-3 of **b**, between H-1 of **a'** and C-3 of **b'**, between H-1 of **b** and C-3 of both $\mathbf{c}_{\alpha}$ and $\mathbf{c}_{\beta}$, between H-1 of **b'** and C-3 of **c'**, between H-1 of **c'** and C-4 of **a**, and between C-1 of **c'** and H-4 of **a**, indicating a linkage pattern as shown in the structure below. Other two- and three-bond correlations are listed in Table 4.9.

The pyruvate substituent: The two 4-linked Gal residues, **a** and **a'** in **P2**, were shown to be substituted at *O*-2 and *O*-3 by the downfield shifts of their C-2 and C-3 resonances as compared

to those of residue **a** in the **DPS**. *O*-2 and *O*-3 must therefore be the location of the pyruvate substituent in **a** and **a'**. Unfortunately no correlations between the acetalic carbons of the pyruvate substituents and H-2 and H-3 of residues **a** and **a'** were visible in the HMBC spectrum. For this reason, and because the ^{1}H resonances for the methyl protons of the pyruvic moieties were identical, it was not possible to distinguish between the two pyruvate substituents. The absolute configuration of the acetalic carbon was established from a ROESY[264] experiment performed on **P2**. NOEs were clearly visible between the H-3 signals of both **a** and **a'** and the methyl protons of the pyruvic acetals, indicating that the pyruvic acid had the *S* configuration[199]. Additional NOEs, listed in Table 4.10, served to confirm the linkages assigned from the HMBC data.

4.4.4 Conclusion

The structure of **P2** can thus be written as:



and the structure of the repeating unit of the *E. coli* K103 polysaccharide is therefore as shown in the Abstract.

This is the fourth *E. coli* capsular polysaccharide, including K50 (this thesis), found to contain pyruvate as its sole acidic function[57,58]. Although pyruvate is a common component of bacterial polysaccharides[296], this is only the second 2,3-linked pyruvate to be reported in the *E. coli* series, the other being found in *E. coli* K33[46,302]. In this case the pyruvate is 2,3-linked to a ß-Glc*p*A residue. The K103 capsular polysaccharide is one of nine *E. coli* capsular polysaccharides which have a linear trisaccharide repeating unit[48].

Table 4.8 NMR data[a] for P2

Residue		Proton or Carbon						
		1	2	3	4	5	6	6'
a								
→4)-α-D-Gal*p*	H	5.62	4.09	4.24	4.48	4.17	3.86	3.86
2,3	^3J[b]	3.4		3.6				
V	C	94.14	73.49	73.60	75.20	72.77	60.44	
Pyr								
a'								
α-D-Gal*p*	H	5.56	4.01	4.19	4.42	4.10	3.75	3.75
2,3	^3J[b]	3.9	9.6	2.6				
V	C	93.92	72.73	74.83	67.89	72.33	61.41	
Pyr								
b								
→3)-α-L-Rha*p*	H	5.10	4.36	4.05	3.64	3.89	1.32	-
	^3J[b]	1.8	3.4					
	C	102.57	67.06	75.84	70.80	69.83	17.36	
b'								
→3)-α-L-Rha*p*	H	5.08	4.33	4.03	3.62	3.88	1.33	-
	^3J[b]	1.9	2.6					
	C	102.49	67.03	75.62	70.87	69.83	17.42	
c'								
→3)-α-D-Gal*p*	H	5.02	3.91	3.83	4.07	4.24	3.66	3.66
	^3J[b]	4.0						
	C	100.48	68.45	78.21	69.59	71.44	62.10	
c_α								
→3)-α-D-Gal*p*OH	H	5.28	3.93	3.90	4.06	4.10	3.70	3.70
	^3J[b]	3.1						
	C	92.90	68.20	77.89	69.82	71.11	61.65	
c_β								
→3)-ß-D-Gal*p*OH	H	4.63	3.60	3.72	4.02	3.73	3.76	3.76
	^3J[b]	7.4						
	C	96.89	71.79	81.22	69.14	75.69	61.48	

[a]Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 ppm respectively for ^1H and ^{13}C.
[b] ^1H-^1H Coupling constants in Hz.

Table 4.9 **Two and three-bond ^{1}H-^{13}C correlations (HMBC) for P2**

Residue	Proton	Correlation to
a	H-1	75.84(b;C-3), 73.60(a;C-3), 72.77(a;C-5)
	H-2	73.60(a;C-3)
	H-3	73.49(a;C-2), 75.20(a;C-4)
	H-4	100.48(c';C-1), 73.49(a;C-2), 73.60(a;C-3)
	H-5	75.20(a;C-4), 60.44(a;C-6)
a'	H-1	75.62(b';C-3), 74.83(a';C-3), 72.33(a';C-5)
	H-2	74.83(a';C-3)
	H-4	72.73(a';C-2), 74.83(a';C-3)
b	H-1	77.89(c_{α};C-3), 81.22(c_{β};C-3), 67.06(b;C-2), 75.84(b;C-3), 69.83(b;C-5)
	H-2	75.84(b;C-3), 70.80(b;C-4)
	H-3	94.14(a;C-1), 70.80(b;C-4)
b'	H-1	78.21(c';C-3), 67.03(b';C-2), 75.62(b';C-3), 69.83(b';C-5)
	H-3	93.92(a';C-1)
c'	H-1	75.20(a;C-4), 78.21(c';C-3), 71.44(c';C-5)
	H-2	78.21(c';C-3)
	H-3	102.49(b';C-1), 69.59(c';C-4)
c_{α}	H-1	77.89(c_{α};C-3), 71.11(c_{α};C-5)
	H-3	102.57(b;C-1)
c_{β}	H-1	75.69(c_{β};C-5)
	H-3	102.57(b;C-1)

Table 4.10 NOE data for P2

Residue	Proton	NOE to
a	H-1	4.36(b;H-2), 4.05(b;H-3), 4.09(a;H-2), 4.24(a;H-3)
	H-3	5.62(a;H-1), 4.48(a;H-4), 3.86(a;H-6), 1.59(pyr)
	H-4	5.02(c';H-1), 4.17(a;H-5), 3.86(a;H-6)
a'	H-1	4.33(b';H-2), 4.03(b';H-3), 4.01(a';H-2), 4.19(a';H-3)
	H-3	5.56(a';H-1), 4.42(a';H-4), 1.59(pyr)
	H-4	4.10(a';H-5), 3.75(a';H-6)
b	H-1	3.93(c_α;H-2), 3.90(c_α;H-3), 4.06(c_α;H-4), 3.60(c_β;H-2), 3.72(c_β;H-3), 4.02(c_β;H-4), 4.36(b;H-2)
	H-2	5.62(a;H-1), 5.10(b;H-1), 4.05(b;H-3)
	H-6	4.05(b;H-3), 3.64(b;H-4), 3.87(b;H-5)
b'	H-1	3.91(c';H-2), 3.83(c';H-3), 4.07(c';H-4), 4.33(b';H-2)
	H-2	5.56(a'-H-1), 5.08(b';H-1), 4.03(b'-H-3)
	H-6	4.03(b';H-3), 3.62(b';H-4), 3.87(b';H-5)
c'	H-1	4.48(a;H-4), 3.91 (c';H-2), 3.83(c';H-3)
	H-5	3.83(c'-H-3), 4.07(c';H-4)
c_α	H-1	3.93(c_α;H-2), 3.90(c_α;H-3)
	H-4	5.10(b;H-1), 3.93 (c_α;H-2), 3.90(c_α;H-3)
c_β	H-1	3.60(c_β;H-2), 3.73(c_β;H-3)
	H-4	5.10(b;H-1), 3.60(c_β;H-2), 3.72(c_β;H-3)

4.4.5. Experimental

General methods.— Instrumentation was essentially as described in Section 4.1.4. A J & W Scientific fused-silica DB-17 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was used for separating alditol acetates (100 kPa, temperature programme: 180°C for 2 min, then 2°C.min^{-1} to 240°C). A J & W Scientific fused-silica DB-225 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was used for separating partially methylated alditol acetates (100 kPa, 205°C, isothermal). A Machery-Nagel fused-silica FS-Lipodex™ A capillary column (50 m x 0.25 mm, film thickness 0.25 μm) was used for separating trifluoroacetates of methyl glycosides (150 kPa, temperature programme: 80°C for 1 min, then 2°C.min^{-1} to 150°C). The identities of all derivatives were determined by comparison with authentic standards and confirmed by GLC-MS. GPC was performed on dextran-calibrated columns (1.6 x 65 cm) of Sephacryl™ S500 and Sephacryl™ S400 using 0.1 M sodium acetate buffer (pH 5.00) as eluent. Compounds were detected by refractive index.

Isolation and purification of the K103 polysaccharide.— An authentic culture of *E. coli* O101:K103:H$^-$ was obtained from Dr I. Ørskov (Copenhagen) and propagated on Mueller-Hinton agar. The capsular polysaccharide was isolated as described for *E. coli* K50 and the yield obtained from a single growing was 270 mg. The growth and isolation procedure was therefore repeated twice to ensure sufficient material for analysis. Depyruvylation of the polysaccharide was carried out by treatment of the material with aqueous 1% acetic acid (100°C, 1 h) with subsequent dialysis against running water (overnight, 12-14 000 mw cut-off), centrifugation (25 000 rpm, 1 h) and freeze-drying.

Glycose and methylation analysis.— Glycose analysis was carried out as described for the K101 polysaccharide. The depyruvylated polysaccharide (**DPS**) was methylated as described for the K45 polysaccharide, except that the 1,1,1,1-tetramethyl urea was omitted. The permethylated product was analyzed as described for the K101 polysaccharide.

Absolute configuration.— For determination of absolute configuration of the constituent monosaccharides, a sample of the **DPS** (10 mg) was methanolysed for 24 h, the product was dissolved in THF (0.5 mL) and two 600 μL additions of trifluoroacetic anhydride were made at 10 min intervals with stirring[303]. The excess reagent was evaporated off under a stream of N_2 and the derived trifluoroacetates were analysed by GLC as described.

Bacteriophage isolation and purification.— A bacteriophage that could be propagated on *E. coli* K103 bacteria was isolated by incubating a mixture of Grahamstown sewage water (25 mL) and a 6 h culture (5 mL) of the host bacterium at 37°C for 12 h. Bacterial growth was terminated by the addition of $CHCl_3$ (15 mL), the cellular material was removed by low-speed centrifugation (2000 rpm, 20 min) and the clear supernatant was retained. The resulting impure bacteriophage suspension was assayed using the "lawn assay technique"[224] and the titre found to be 1 x 10^8 PFU/mL. The bacteriophage was purified and the titre increased for the purposes of depolymerisation as follows:

A single plaque was picked from one of the agar plates and added to nutrient broth (5 mL) which had been freshly inoculated with host bacteria. This was incubated with shaking at 37°C for 6 h after which the bacterial cells were lysed by the addition of a few drops of $CHCl_3$ and removed by centrifugation as described above. This process was repeated to ensure a pure bacteriophage. The titre and volume of the pure bacteriophage suspension was increased by a series of tube lyses. Nine sterile culture tubes containing 5 mL sterile nutrient broth were inoculated with 0.5 mL of bacterial culture and bacteriophage solution (0.1 mL) was successively added to the tubes at 15 minute intervals. The tubes were incubated at 37°C for 6 h, after which the fifth tube was assayed as described above and the titre found to be 9.6 x 10^9 PFU/mL. To increase the titre to the desired level[234], the process was repeated on a larger scale, using the bactriophage suspension from the fifth tube as the inoculant. 50 mL of bacteriophage solution containing 4 x 10^{14} PFU was prepared and purified exhaustively by dialysis against tap water (12-14 000 mw cut-off). The solution was assayed by the "lawn-technique" before and after dialysis and the titre was unaltered.

Bacteriophage-mediated depolymerization of PS.— A sample of K103 **PS** (sodium salt, 300 mg) was dissolved in the bacteriophage solution (50 mL) and was gently agitated at 37°C in the presence of $CHCl_3$ (1 mL) to prevent bacterial growth. After 3 days, the solution was shaken with $CHCl_3$ to effect cell lysis and centrifuged (2000 rpm, 15 min) to remove cellular debris. The aqueous supernatant was removed and freeze-dried, and the resulting solid was redissolved in water and dialysed (mw cut-off < 3500) against 8 x 30 mL distilled water fractions for 24 h periods. The combined diffusate was freeze-dried, redissolved in water and subjected to successive chromatographic separations on Sephacryl™ S200 using NaOAc buffer (0.1 M, pH 5.00) as eluent and on Bio-Gel™ P4 (both columns 2.6 cm x 100 cm) using distilled water as eluent. Material was detected by refractive index. 28 mg of **P2** was obtained.

NMR spectroscopy.— Samples were prepared as described previously. Spectra were recorded at 30°C on a Bruker AMX-400 spectrometer equipped with an X32 computer. The parameters used for 2D experiments were as follows: COSY and HOHAHA [512 x 2048 data matrix, zero-filled to 1024 data points in t_1; 48 or 56 scans per t_1 value; spectral width 2008 Hz; recycle delay 1.2 s (COSY) or 1.0 s (HOHAHA), mixing time 89 ms (HOHAHA); unshifted sine-bell filtering in t_1 and t_2 (COSY); shifted sine-squared filtering in t_1 and t_2 (HOHAHA)]. ROESY [256 x 4096 data matrix, zero-filled to 1024 data points in t_1; 104 scans per t_1 value; spectral width 3846 Hz; recycle delay 2.0 s; shifted sine-squared filtering in t_1 and t_2; carrier frequency placed at far left hand side of spectrum to minimize COSY and HOHAHA cross-peaks[304]]. DEPT-HMQC and HMQC-TOCSY [256 x 4096 data matrix, zero-filled to 1024 data points in t_1; 64 or 76 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.45 ms; mixing delay 89 ms (HMQC-TOCSY); spectral width in t_1 14 086 Hz and in t_2 2016 Hz; shifted sine-squared filtering in t_1 and t_2; read pulse angle 60°]. HETCOR [128 x 4096 data matrix, zero-filled to 1024 data points in t_1; 1500 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.45 ms; spectral width 2200 Hz in t_1 and in t_2 10 638 Hz; shifted sine-squared filtering in t_1 and t_2]. HMBC [512 x 4096 data matrix, zero-filled to 1024 data points in t_1; 64 scans per t_1 value; recycle delay 1.0 s; fixed delay 3.45 ms; spectral width 20 828 Hz in t_1 and 2024 in t_2; shifted sine-squared filtering in t_1 and t_2].

4.5 STRUCTURAL INVESTIGATION OF THE CAPSULAR POLYSACCHARIDE OF *Escherichia coli* O8:K43:H11 USING HIGH RESOLUTION NMR SPECTROSCOPY

4.5.1 Abstract

The primary structure of the acidic capsular antigen of *Escherichia coli* O8:K43:H11 was shown by monosaccharide analysis, methylation analysis, ß-elimination, and by 1D and 2D ^{1}H and ^{13}C NMR spectroscopy to be composed of branched pentasaccharide repeating units with the structure:

→4)-*α*-D-Glc*p*A-(1→3)-ß-D-Man*p*-(1→4)-ß-D-Man*p*-(1→3)-ß-D-Glc*p*NAc-(1→
2
↑
1
ß-D-Glc*p*NAc

4.5.2 Introduction

More than 40% of the known capsular (K) polysaccharide antigens of *E. coli* occur in association with only two O antigens *viz.* O8 and O9[84]. The structures of the repeating oligosaccharide units of all sixteen of the K antigens associated with O9 and all but two (K41 and K43) of the sixteen K antigens associated with O8 have been determined. The capsular antigen of *E. coli* O8:K43:H11 may be classified as a Group I polysaccharide since it is a high-molecular mass acidic polysaccharide which contains amino sugars[84].

4.5.3 Results and Discussion

Isolation, purification, and composition of the capsular polysaccharide.— *E. coli* O8:K43:H11 bacteria were grown on Mueller-Hinton agar, the harvested bacterial slime was diluted with aqueous 2% phenol, and ultracentrifuged to remove the cells. The acidic capsular polysaccharide was precipitated from the supernatant solution as the cetyltrimethylammonium complex. The recovered capsular polysaccharide was found to be lipid-bound and was released by treatment with hot aqueous 1% acetic acid. After removal of the lipid by ultracentrifugation the capsular polysaccharide was purified by chromatography on a DEAE-Sepharose™ column. GLC examination of the alditol acetates derived from an acid hydrolysate of the purified polysaccharide

(PS) showed the presence of Man and GlcN. Prior methanolysis of the PS, reduction of the methoxylcarbonyl groups formed, and GLC examination of the derived alditol acetates revealed the presence of Glc in addition to Man and GlcN, thereby establishing GlcA as the acidic component of the repeating unit. The D configuration was established for all the constituent monosaccharides by GLC examination of the derived acetylated (-)-2-octylglycosides[149].

1D NMR studies of the **PS**.— The ^{1}H NMR spectrum of the PS (Figure 4.23) contained discrete H-1 signals at δ 5.352 ($^{3}J_{1,2}$ 4.2 Hz) and 4.689 ($^{3}J_{1,2}$ 8.2 Hz) for an α and a ß-linked residue respectively, and three overlapping H-1 signals between δ 4.735 and 4.718. In addition signals for the methyl protons of two NAc groups at δ 2.070 and 2.055 were present. The ^{13}C NMR spectrum (Figure 4.24) of the PS had C-1 signals at 102.05, 102.00, 101.64, 101.50, and 100.42 ppm, signals at 23.33 and 23.19 ppm for the methyl carbons of NAc groups, signals at 55.65 and 56.55 ppm for acetamido substituted carbons, signals for the carbonyl carbons of NAc groups at 175.13 and 174.92 ppm, and a signal at 173.95 ppm for the carbon of a carboxyl group. The combined results thus far indicated that the repeating unit of the PS consisted of pentasaccharide repeating units composed of GlcNAc:Man:GlcA in the ratios 2:2:1.

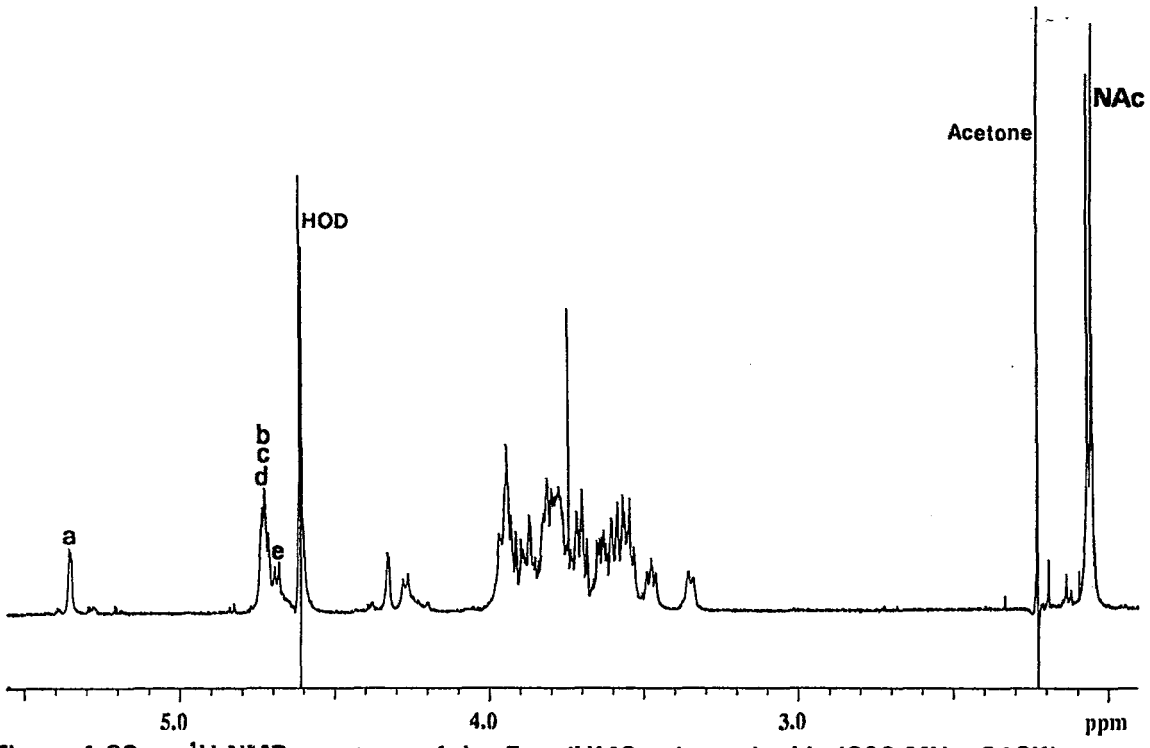


Figure 4.23 ^{1}H NMR spectrum of the *E. coli* K43 polysaccharide (600 MHz, 316K)

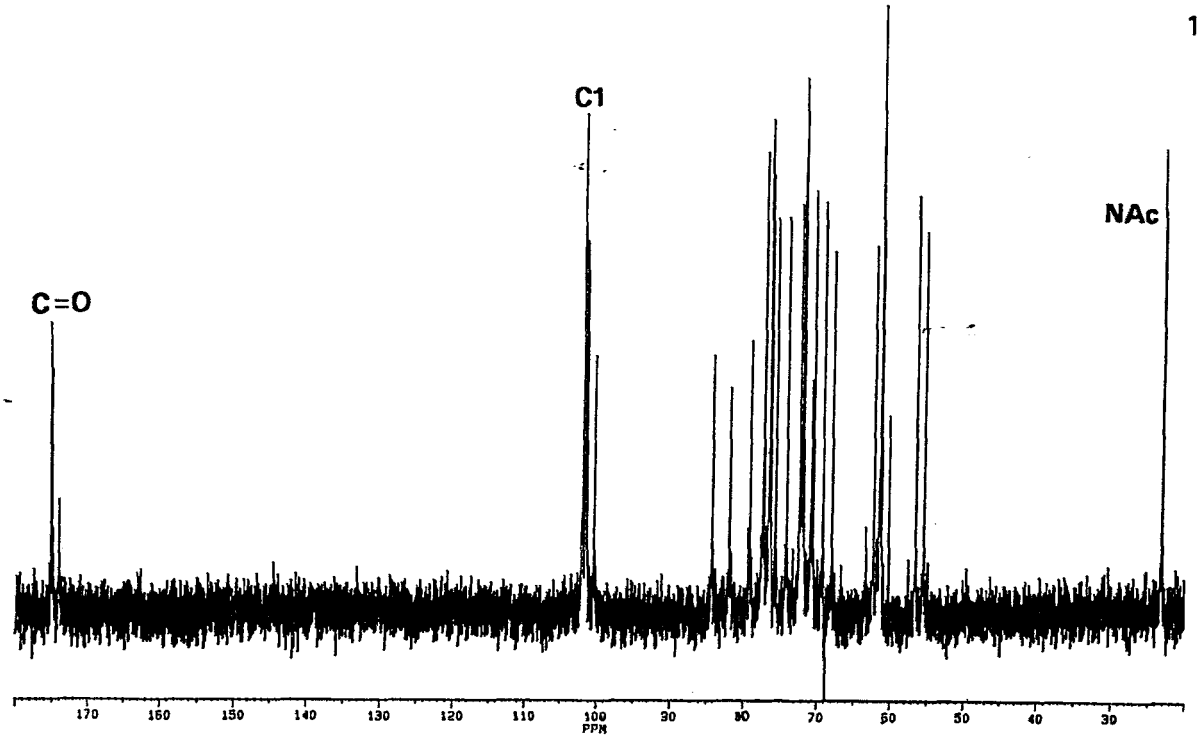


Figure 4.24 **^{13}C NMR spectrum of *E. coli* K43 polysaccharide**

Methylation and ß-elimination analyses.— Methylation of the **PS** by a modified Hakomori procedure[169] followed by Kuhn methylation[179] and GLC-MS analysis of the partially methylated alditol acetates derived from an acid hydrolysate of the methylated **PS** showed the presence of 4,6-di-*O*-methyl-mannose, 2,3,6-tri-*O*-methylmannose, 2,3-di-*O*-methylglucose (after carboxyl reduction), 2-deoxy-2-methylacetamido-4,6-di-*O*-methylglucose and 2-deoxy-2-methylacetamido-3,4,6-tri-*O*-methylglucose. Treatment of the methylated **PS** with sodium methylsulphinylmethanide followed by remethylation of the product and GLC examination of the derived partially methylated alditol acetates showed the presence of 3,4,6-tri -*O*-methylmannose and a great reduction in the ratio of 4,6-di-*O*-methylmannose. These results indicated that the repeating unit of the **PS** is a branched pentasaccharide with the following partial structure:

→4)-D-GlcA-(1→3)-D-Man-(1→
2
↑

The full sequence of the residues in the repeating unit of the **PS** was established by 2D-NMR experiments.

2D NMR studies of the **PS**.— The identity of the residues in the repeating unit, the configurations of the glycosidic linkages, and the glycosylation sites were established by ^{1}H-^{1}H correlation spectroscopy (COSY[305], HOHAHA[306], and NOESY[307]) and by ^{1}H-^{13}C correlation spectroscopy (HMQC)[283]. The COSY experiment was run in the phase-sensitive mode in order to eliminate dispersion peaks and so allow accurate measurement of coupling constants[308]. The residues in the repeating unit were denoted a - e in order of decreasing chemical shift of the H-1 resonances. The ^{1}H and ^{13}C chemical shifts are listed in Table 4.11 and the COSY, HOHAHA, NOESY, and HMQC spectra are shown in Figures 4.26 to 4.33.

Table 4.11 NMR data[a] for *E. coli* K43 polysaccharide

Residue		Proton or carbon						
		1	2	3	4	5	6	6′
→4)-α-D-Glc*p*A (a)	H	5.352	3.642	3.917	3.778	4.274		
	C	100.42	72.06	72.14	82.03	72.06	173.95	
→2,3)-ß-D-Man (b)	H	4.735	4.332	3.823	3.704	3.479	3.628	3.961
	C	102.00[b]	76.73	77.48	68.09	77.36	62.40	
→4)-ß-D-Man (c)	H	4.726	3.943	3.836	3.723	3.583	3.887	3.758
	C	102.05[b]	70.94	72.52	79.25	75.73	61.58[b]	
ß-D-GlcNAc (d)	H	4.724	3.701	3.571	3.549	3.351	3.804	3.940
	C	101.64	56.65	74.22	70.65	76.61	61.40	
→3)-ß-D-GlcNAc (e)	H	4.689	3.870	3.787	3.608	3.548	3.789	3.959
	C	101.50	55.50	84.38	69.33	76.46	61.56[b]	

[a]Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 ppm for ^{1}H and ^{13}C, respectively.
[b]Values may have to be exchanged.

Residue a [→4)-α-D-Glc*p*A]: Since the signal for the anomeric proton of residue a is well separated from the H-1 signals of the other residues it was possible to assign all the chemical shifts for this residue from the COSY (Figures 4.26 and 4.27) and HOHAHA (Figures 4.28 and 4.29) spectra. Magnetism relayed well through the spin system, as expected for a *gluco*-type residue, and none of the crosspeaks were overlapped to any significant degree. The ^{13}C resonances were assigned

by comparison of the ^{1}H assignments with ^{1}H-^{13}C correlation data obtained from the HMQC experiment (Figure 4.33). The downfield shift of the C-4 carbon signal indicated that it was the linkage position for this residue. The anomeric linkage of residue **a** was established as *α* from its H-1 and C-1 chemical shifts, its $^{3}J_{1,2}$ value, and from the intramolecular H-1/H-2 NOE observed in the NOESY spectrum.

Residue **b** [→2,3)-ß-D-Man*p*]: Due to the partial overlap of the anomeric signals of residues **b**, **c**, and **d**, some difficulty was encountered in assigning the resonances for these residues. No H-1/H-2 cross-peak was observed in the COSY and HOHAHA spectra for residue **b**, due to the small $^{3}J_{1,2}$ value typical of a *manno*-type residue. The chemical shifts for these resonances were obtained from the H-1/H-2 NOE observed in the NOESY spectrum (Figures 4.30 to 4.32), and the remaining resonances could then be traced in the COSY and HOHAHA spectra. Carbon resonances were again assigned from the HMQC spectrum. The intramolecular NOEs observed between H-1 and H-3 and between H-1 and H-5, together with the chemical shift of the C-5 resonance, established that the anomeric linkage of residue **b** had the ß configuration.

Residue **c** [→4)-ß-D-Man*p*]: As for residue **b**, no H-1/H-2 cross-peak could be observed for this residue in either the COSY or the HOHAHA spectrum and resonances for these protons were established from the H-1/H-2 NOE observed in the NOESY spectrum. The chemical shifts for H-3 and H-5 of residue **c** were established from the H-1/H-5 and H-3/H-5 NOEs present in the NOESY spectrum. The ^{1}H chemical shifts for the remaining resonances were traced in the COSY and HOHAHA spectra and carbon resonances were assigned from the HMQC spectrum, with the exception of the C-5 resonance as the HMQC spectrum showed no H-5/C-5 cross-peak. Once all other carbon resonances had been assigned, the one remaining resonance was identified as belonging to C-5 of residue **c**. The chemical shift of the C-5 resonance together with the intramolecular NOEs observed between H-1 and H-5 of residue **c** established its anomeric configuration as ß.

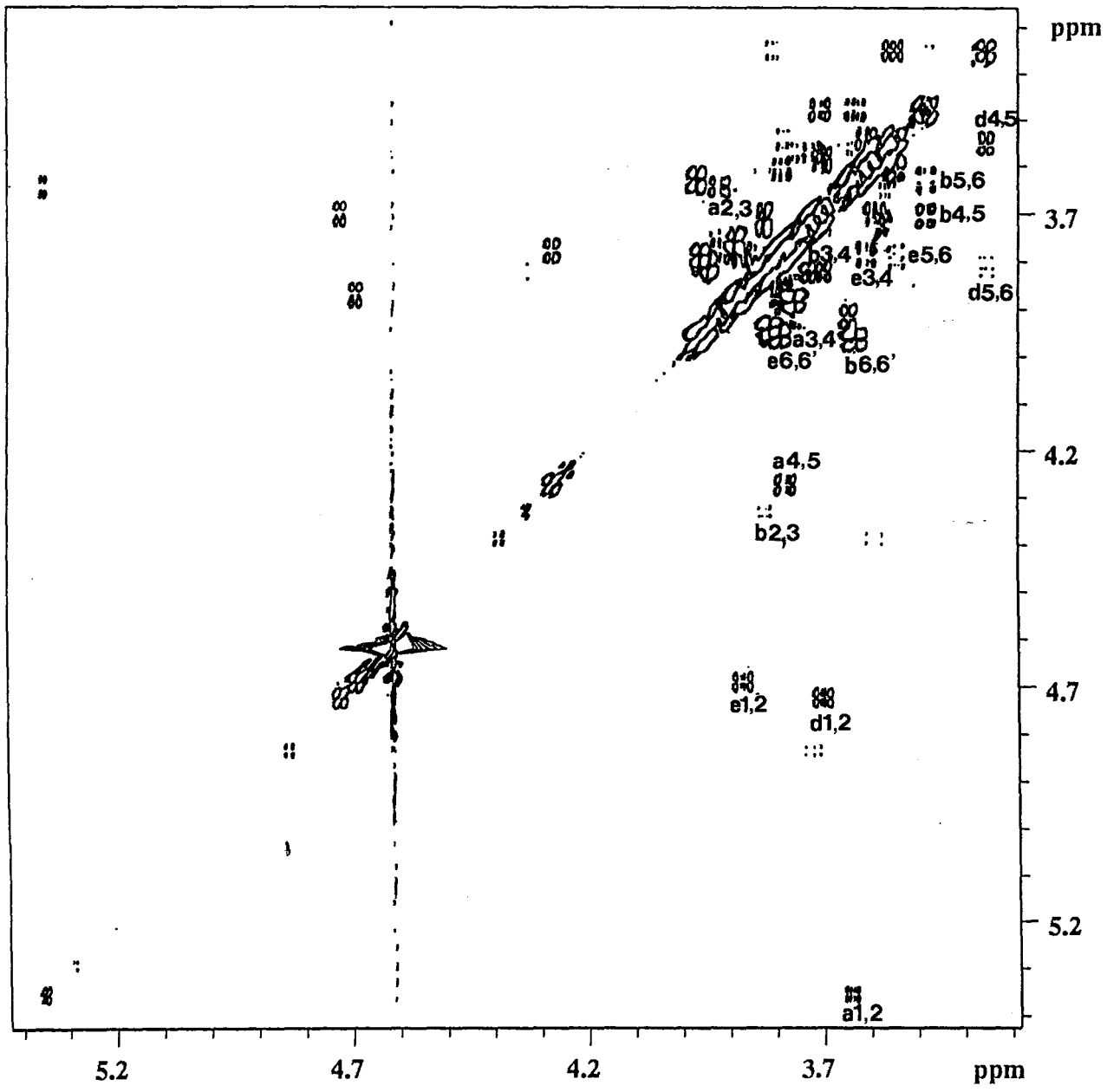


Figure 4.26 COSY contour plot of the region δ 3.20 - 5.50 for the K43 PS. a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc.

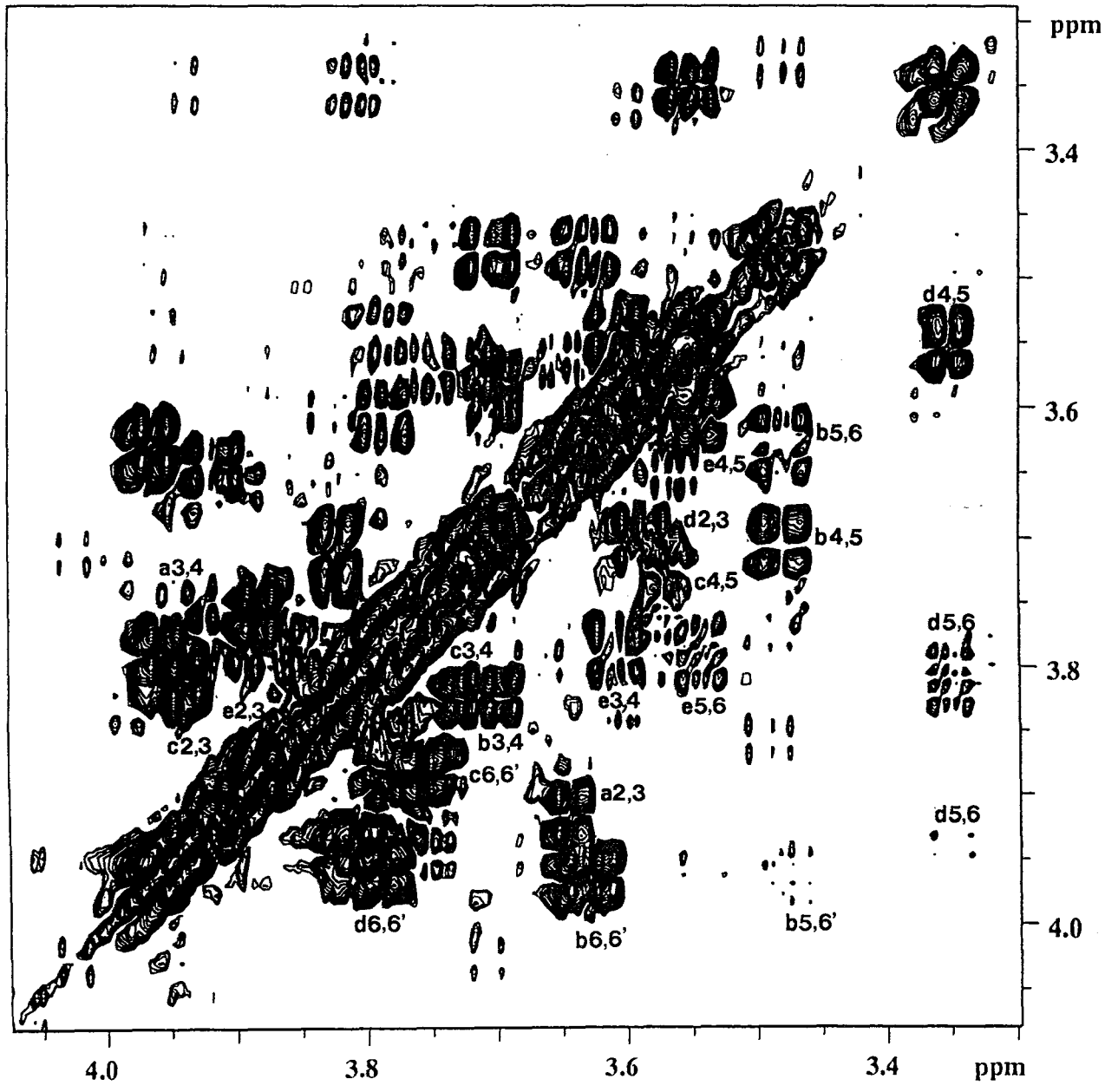


Figure 4.27 Expanded view of the COSY contour plot for the K43 PS showing the spectral region δ 3.30 - 4.05.

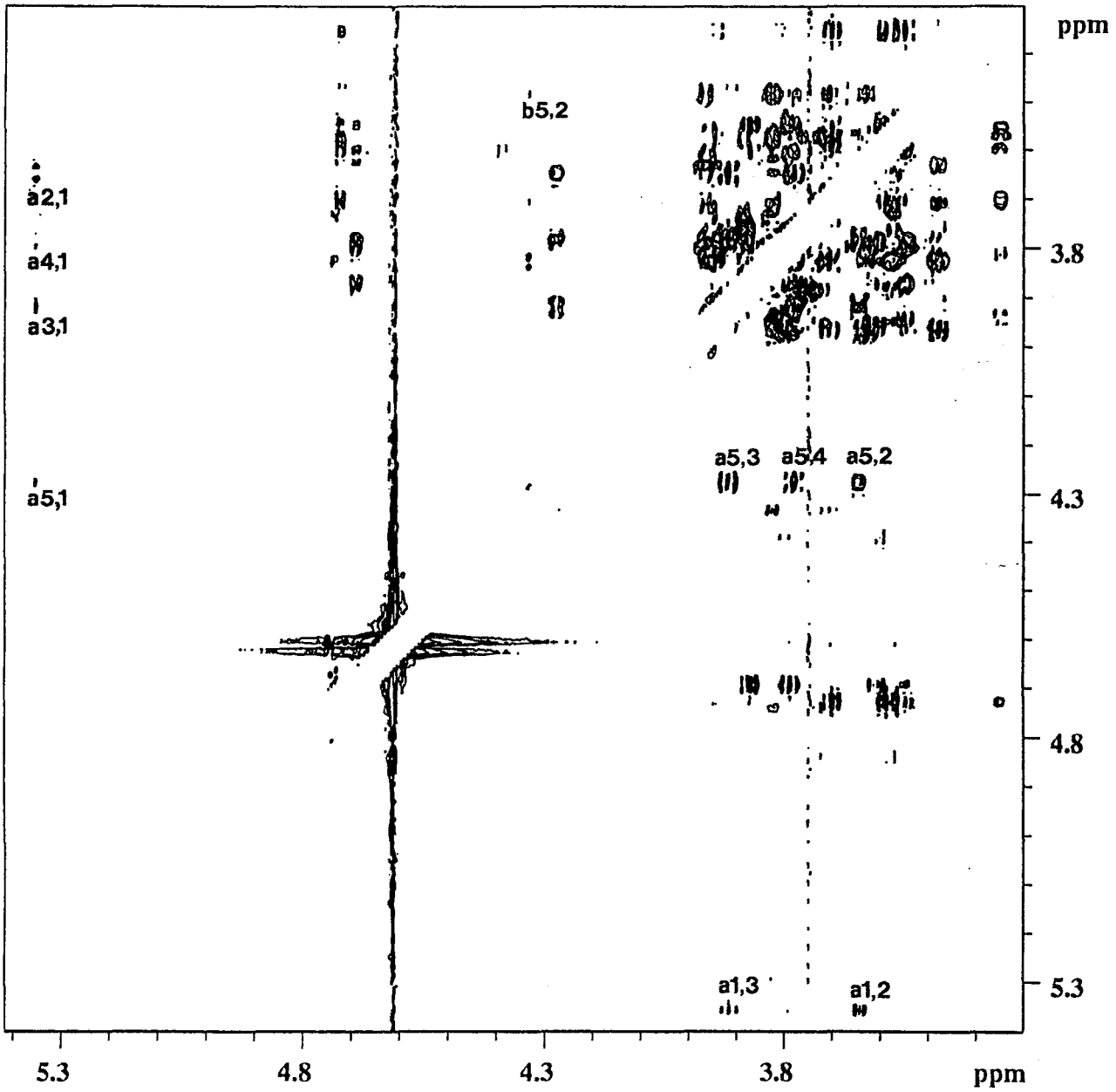


Figure 4.28 HOHAHA contour plot of the region δ 3.30 - 5.40 for the K43 PS. a1,2 connotes the cross-peak between H-1 and H-2 of residue a, etc.

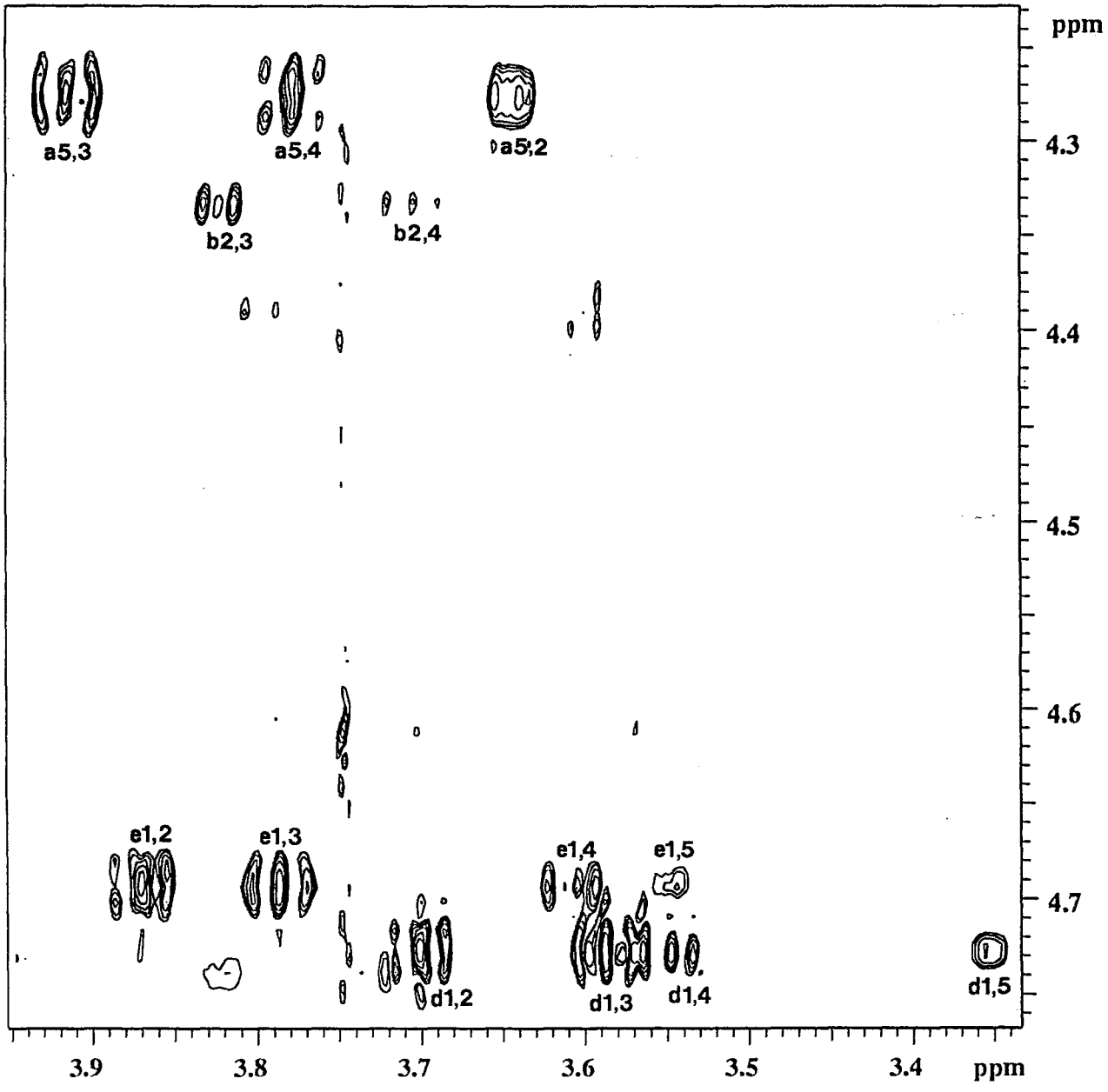


Figure 4.29 Expanded view of the HOHAHA contour plot for the K43 PS, showing the spectral regions δ 4.23 - 4.76 in f_1 and δ 3.34 - 3.95 in f_2

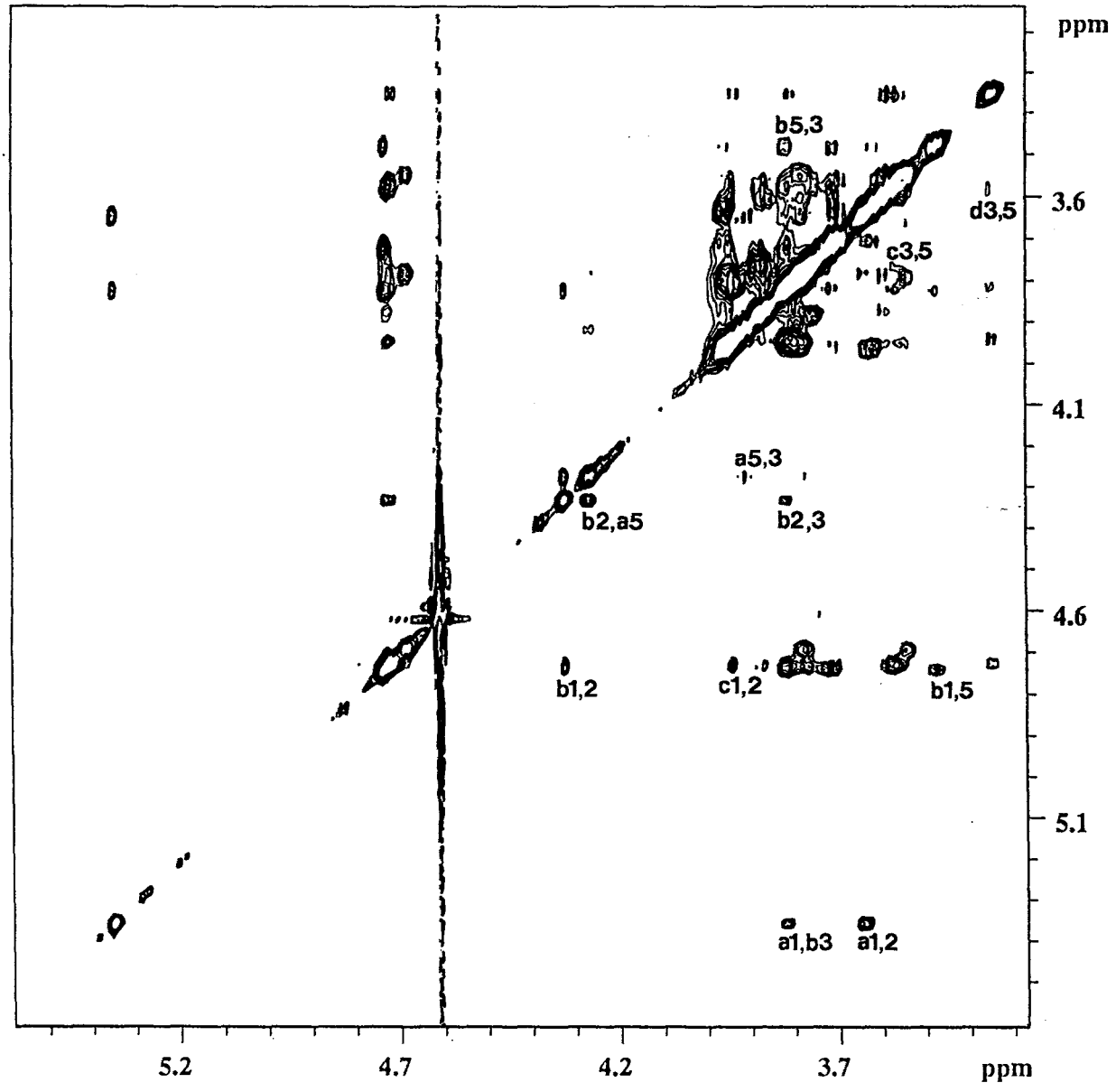


Figure 4.30 Two-dimensional NOESY spectrum for the K43 PS, showing the spectral region δ 3.30 - 5.50. a1,2 connotes the cross-peak between H-1 and H-2 of a, etc.

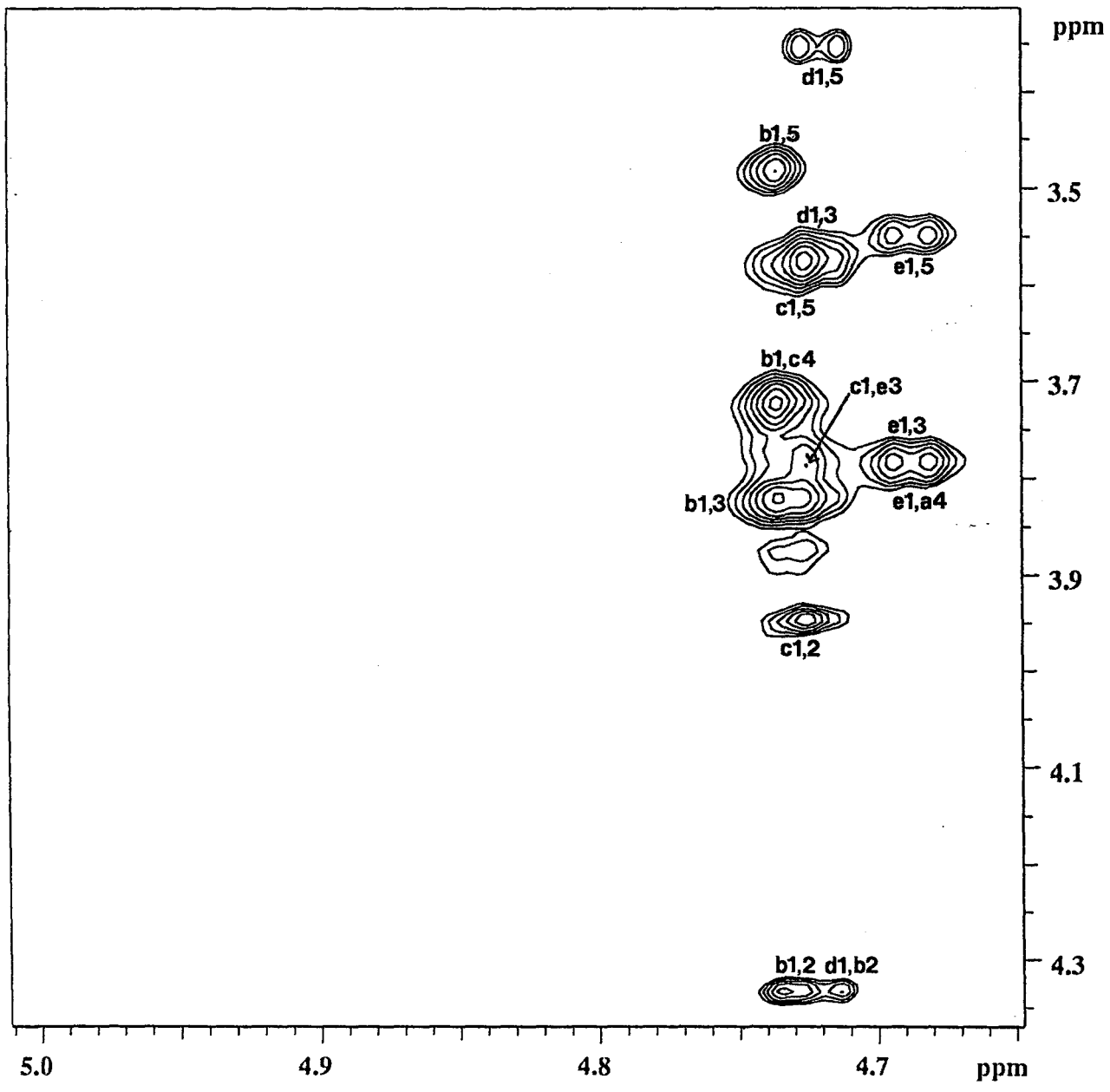


Figure 4.31 Expanded view of the 2D NOESY spectrum for the PS, showing the spectral regions δ 3.35 - 4.35 in f_1 and δ 4.65 - 5.01 in f_2

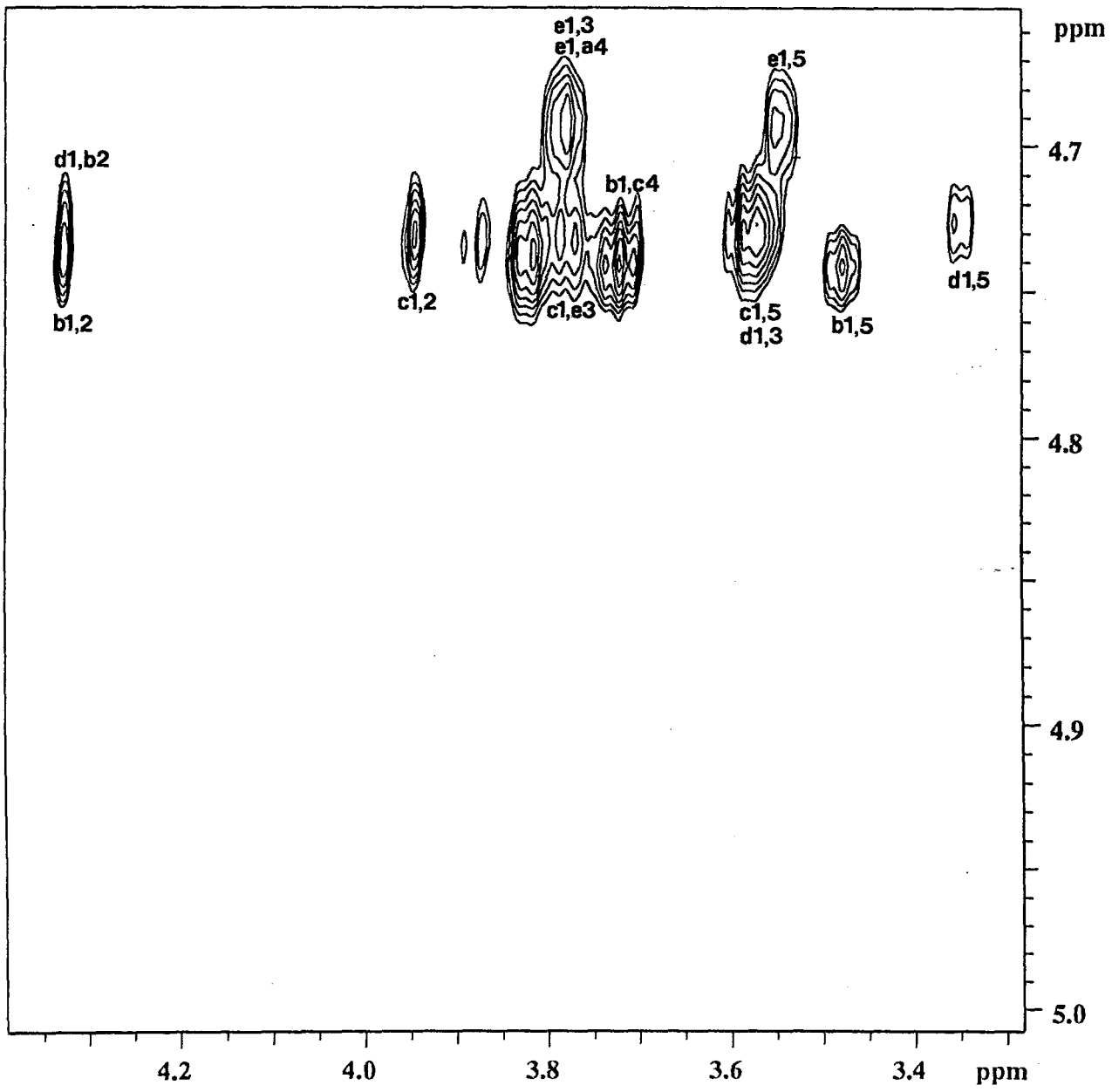


Figure 4.32 Expanded view of the 2D NOESY spectrum for the PS, showing the spectral regions δ 4.66 - 5.00 in f_1 and δ 3.30 - 4.35 in f_2

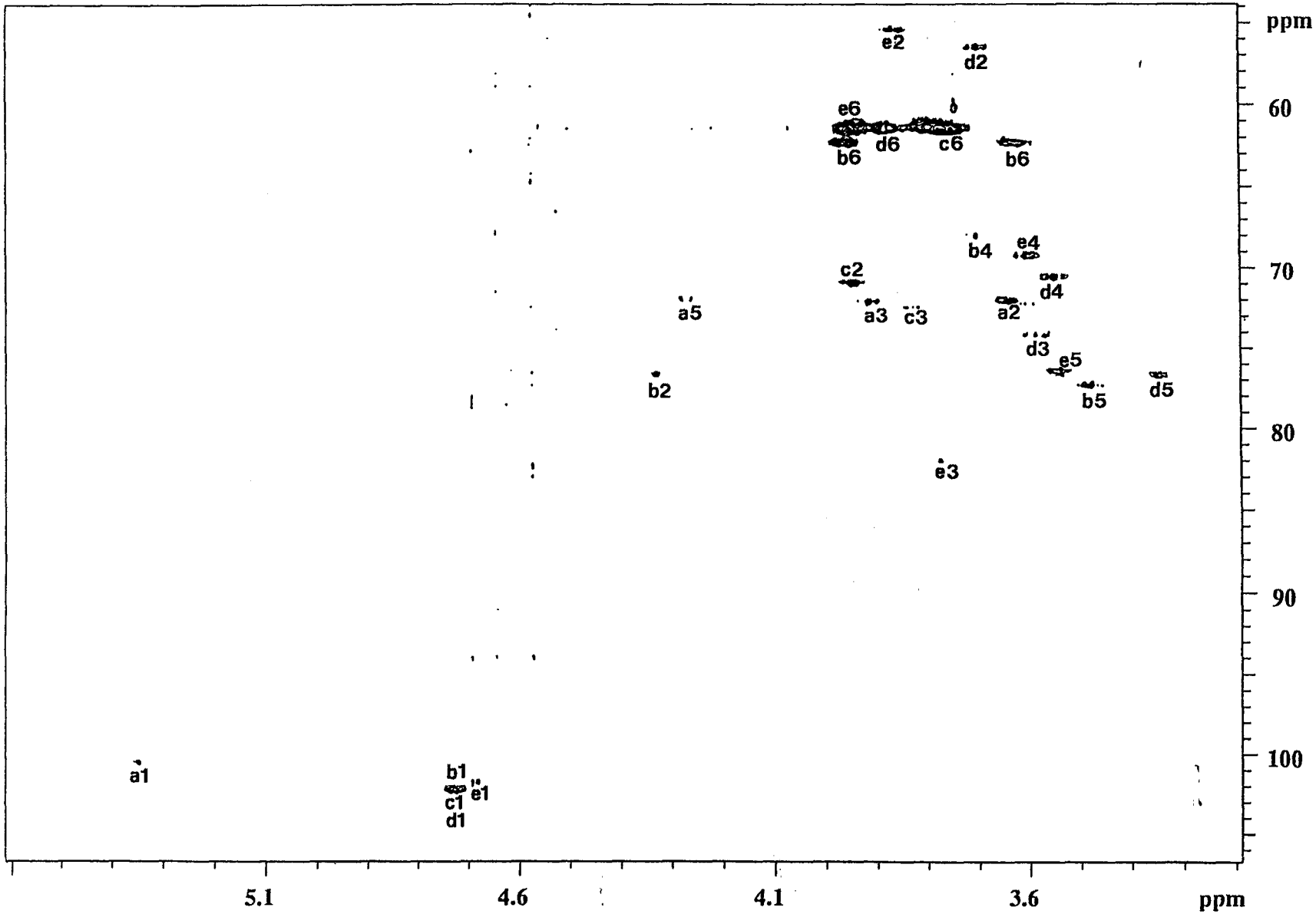


Figure 4.33 HMQC ^{1}H-^{13}C shift correlation map of the spectral regions f_1 54 - 106 ppm (^{13}C) and f_2 δ 3.20 - 5.60 (^{1}H) for the PS
Correlations for a4, b3, and c4 are only visible if a different window function is used

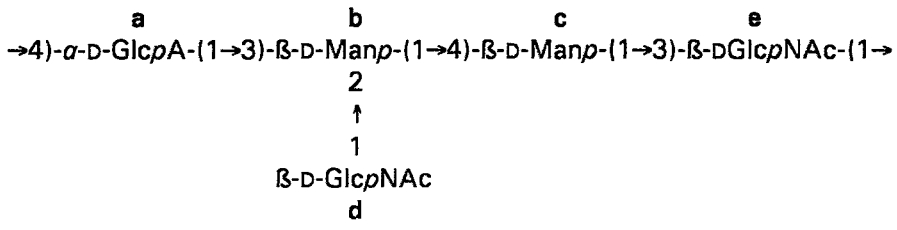
Residues **d** [ß-D-Glc*p*NAc] and **e** [→3)-ß-D-Glc*p*NAc]: The ^{1}H chemical shifts for these two residues were established from the COSY and HOHAHA spectra, since magnetism relayed well through the *gluco*-type spin systems and all cross-peaks were easily visible. The anomeric configurations for residues **d** and **e** were established as ß from their H-1 chemical shift values, their intramolecular H-1/H-3 and H-1/H-5 NOEs, and their $^3J_{1,2}$ values.

Comparison of the ^{1}H and ^{1}C chemical shifts for residues **a** - **e** with literature values for methyl glycosides[144,284,285] identified the residues in the repeating unit as indicated in Table 4.11. In agreement with the methylation results for the glycosylation sites were established as C-4 for **a**, C-2 and C-3 for **b**, C-4 for **c**, and C-3 for **e** by the significant deshielding experienced by these carbons.

The sequence of the residues in the repeating unit was established from the NOESY experiment. The NOEs observed are listed in Table 4.12. Assignment of the NOEs was greatly facilitated by the PRONTO software program[309] which permits simultaneous viewing and interrogation of NMR spectra. Intermolecular NOEs were observed between H-1 of **a** and H-3 of **b**, between H-1 of **b** and H-4 of **c**, between H-1 of **c** and H-3 of **e**, between H-1 of **d** and H-2 of **b**, and between H-1 of **e** and H-4 of **a**.

4.5.4 Conclusion

The combined chemical and NMR results support the following structure for the repeating unit of the capsular polysaccharide of *E. coli* K43:



The branched pentasaccharide repeating unit has a "4 + 1" structure, which is shared by only two other capsular polysaccharides of the *E. coli* series, *viz.* K39[221] and K87[310]. This is only the fourth *E. coli* capsular polysaccharide found to contain two GlcNAc residues, the others being K40[81], K85[311], and K45 (Section 4.2).

Table 4.12 NOE data for the PS

Residue	Proton	Correlation to
→4)-*α*-D-Glc*p*A (a)	H-1	3.823 (b, H-3), 3.642 (a, H-2)
	H-3	4.274 (a, H-5)
→2,3)-ß-D-Man (b)	H-1	3.723 (c, H-4), 4.332 (b, H-2), 3.823 (b, H-3), 3.479 (b, H-5)
	H-2	4.274 (a, H-5), 3.823 (b, H-3)
	H-5	3.823 (b, H-3)
	H-6	3.961 (b, H-6')
→4)-ß-D-Man (c)	H-1	3.787 (e, H-3), 3.943 (c, H-2), 3.583 (c, H-5)
	H-5	3.836 (c, H-3)
	H-6	3.758 (c, H-6')
ß-D-GlcNAc (d)	H-1	4.332 (b, H-2), 3.571 (d, H-3), 3.351 (d, H-5)
	H-3	3.351 (d, H-5)
	H-6	3.940 (d, H-6')
→3)-ß-D-GlcNAc (e)	H-1	3.778 (a, H-4), 3.787 (e, H-3), 3.548 (e, H-5)
	H-6	3.959 (e, H-6')

4.5.5 Experimental

General methods.— Instrumentation was essentially as described in Section 4.1.4. A J & W Scientific fused-silica DB-17 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was used for separating partially methylated alditol acetates (programme I), and alditol acetates and acetylated octyl glycosides (programme II). A J & W Scientific DB-225 bonded-phase capillary column (30 m x 0.25 mm, film thickness 0.25 μm) was also used for separating acetylated octyl glycosides (130 kPa, 240°C isothermal). The temperature programmes used were: I, 180°C for 2 min, then 3°C.min^{-1} to 240°C, 100 kPa; II, 180°C for 2 min, then 2°C.min^{-1} to 240°C, 100 kPa. The identities of all derivatives were determined by comparison with authentic standards and confirmed by GLC-MS on a Hewlett-Packard 5988A instrument, using the appropriate column.

Isolation and purification of the K43 polysaccharide.— An authentic culture of *E. coli* O8:K43:H11 was obtained from Dr I. Ørskov (Copenhagen) and propagated on Mueller-Hinton agar. The capsular polysaccharide was isolated as described for *E. coli* K50 (Section 4.3.5) 460 mg of capsular polysaccharide was obtained. The polysaccharide was then delipidated by treatment with aqueous 1% acetic acid (100°C, 1 h) after which it was converted to the acid form by passage down a column of Amberlite IR 120 [H^+] and purified by GPC-ion exchange chromatography on a column (2.6 x 27 cm) of DEAE-Sepharose™ CL-6B using gradient elution with 0 - 1 M NaCl in 0.01 M Tris-HCl (pH 8.50). Fractions were assayed for carbohydrate using the phenol-H_2SO_4 reagent[85].

Monosaccharide composition, absolute configuration, and methylation analysis.— Glycose analysis was carried out as described for the K101 polysaccharide. Methylation analysis was performed as described for the K50 polysaccharide, the modified Hakomori method[169] being followed by a 72 h Kuhn methylation[179]. The absolute configurations of the monosaccharides were determined by GLC analysis of the derived acetylated (-)-2-octyl glycosides, as described for the K101 polysaccharide.

NMR spectroscopy.— Samples were prepared as described previously. Spectra were recorded at 43°C on a Bruker AMX 600 spectrometer. The 2D pulse programmes used were as follows: a) Phase-sensitive COSY using time proportional phase increments (TPPI)[312] to preserve phase information with double quantum filter and presaturation during relaxation delay[305]; b) phase-sensitive NOESY using the method of States *et al.* with presaturation during relaxation delay[306]; c) HOHAHA using the method of States *et al.* with presaturation during relaxation delay and employing the DIPSI-2 sequence for mixing[307]; d) HMQC using the States-TPPI method with presaturation during the relaxation delay and the BIRD pulse delay[313] (to remove proton signals associated with ^{12}C) and using GARP1 decoupling during acquisition. The parameters for the 2D experiments were as follows: COSY, NOESY, HOHAHA [1024 x 4096 data matrix, zero-filled to 2048 data points in t_1; 16, 24, or 32 scans per t_1 value; spectral width 7.574 ppm; mixing time 200 ms (NOESY) and 62.6 ms (HOHAHA); shifted sine-squared window function applied prior to fourier transformation]. HMQC [512 x 4096 data matrix, zero-filled to 1024 data points in t_1; spectral width 110 ppm in t_1 and 7.547 ppm in t_2; 32 scans per t_1 value; sine-squared (t_2) and sine (t_1) window functions were applied prior to fourier transformation].

CHAPTER FIVE: APPENDIX I

Formulae for Nutrient Media

Mueller-Hinton Agar (g/L) in water

Meat Infusion	5.0
Casein hydrolysate	17.5
Starch	1.5
Agar	14.0

pH = 7.4 (± 0.2)

Mueller-Hinton Broth (g/L in water)

Meat infusion	2.0
Casein hydrolysate	17.5
Starch	1.5

pH = 7.4 (± 0.2) at 25°C

Luria-Bertani Broth (g/L in water)

Bacto-tryptone	10.0
Yeast extract	5.0
NaCl	10.0
Maltose	2.0

Luria Bertani Agar (g/L in water)

Bacto-tryptone	10.0
Yeast extract	5.0
NaCl	10.0
Maltose	2.0
Agar	15.0

Nutrient Broth (g/L in water)

Meat extract	1.0
Peptone	5.0
Yeast extract	2.0
NaCl	8.0

CHAPTER SIX: REFERENCES

1. L. Kenne and B. Lindberg in *The Polysaccharides*, vol. 2, edited by G.O. Aspinall (Academic Press 1983) pp. 287-365.
2. S. M. Hammond, P.A. Lambert, and A.N. Rycroft, in *The Bacterial Cell Surface* (Kapitan Szabo Publishers, Washington DC. 1984).
3. G.D. Shockman and J.F. Barrett, *Ann. Rev. Microbiol.*, 37, (1983) 501-527.
4. I.W. Sutherland, in *Biotechnology of microbial expoloysaccharides* (Cambridge University Press 1983).
5. I W Sutherland in *Biomedical and Biotechnological Advances in Industrial Polysaccharides*, edited by V. Crescenzi, I.C.M. Dea, S. Paoletti, S.S. Stivala, and I.W. Sutherland (Gordon and Breach Science Publishers, 1989) pp. 123-132.
6. R. Bell and G. Torrigiani, in *Towards better carbohydrate vaccines* (Gordon and Breach, New York, 1987).
7. A.F.M. Verheul, G.J.P.H. Boons, G.A. Van der Marel, J.H. Van Boom, H.J. Jennings, H. Snippe, J. Verhoef, P. Hoogerhout, and J.T. Poolman, *Infect. Immun.*, 59 (1991) 3566-3573.
8. A. Jeanes, *Encycl. Polymer Sci. Technol.*, 4 (1968) 693-711.
9. F. Priest and B. Austin in *Modern Bacterial Taxonomy*, 2nd edition, (Chapman and Hall, London, 1993).
10. H.J. Rogers, H.R. Perkins, and J.B. Ward, *Microbial Cell Walls and Membranes* (Chapman and Hall, London, 1980).
11. D.J. Tipper, J-M. Ghuysen, and J.L. Strominger, *Biochemistry*, 4 (1965) 468-473.
12. D.J. Tipper and J.L. Strominger, *Proc. Natn. Acad. Sci. USA*, 54 (1965) 1133-1141.
13. W. Brumfitt, A.C. Wardlaw, and J.T. Park, *Nature*, 181 (1958) 1783-1784.
14. H. Hayashi, Y. Araki, and E. Ito, *J. Bacteriol.*, 113 (1973) 592-598.
15. J. Coley, A.R. Archibald, and J. Baddiley, *FEBS Letters*, (1977) 405-407.
16. A.W. Wyke and J.B. Ward, *J. Bacteriol.*, 130 (1977) 1055-1063.

17. S. Hase and Y. Matsushima, *J. Biochem.*, 72 (1972) 1117-1128.

18. T.E. Rohr, G.N. Levy, N.J. Stark, and J.S. Anderson, *J. Biol. Chem.*, 252 (1977) 3460-3465.

19. D.C. Ellwood and D.W. Tempest, *Biochem. J.*, 111 (1969) 1-5.

20. K.W. Knox and A.J. Wicken, *Bact. Rev.*, 37 (1973) 215-257.

21. D.A. Powell, M. Duckworth, and J. Baddiley, *Biochem. J.*, 151 (1975) 387-397.

22. P.A. Lambert, I.C. Hancock, and J. Baddiley, *Biochim. Biophys. Acta*, 472 (1977) 1-12.

23. A.J. Wicken and K.W. Knox, *Science*, 187 (1975) 1161-1167.

24. E.A. Hall and K.W. Knox, *Biochem. J.*, 96 (1965) 302-309.

25. O. Larm and B. Lindberg, *Adv. Carbohydr. Chem.*, 33 (1976) 295-322.

26. J. Hoffman, B. Lindberg, J. Lönngren, and T. Hofstad, *Carbohydr. Res.*, 47 (1976) 261-267.

27. Y. Habu, M. Nagaoka, T. Yokokura and I. Azuma, *J. Biochem.*, 102 (1987) 1423-1432.

28. M. Nagaoka, M. Muto, T. Yokokura, and M. Mutai, *J. Biochem.*, 103 (1988) 618-621.

29. P.H Mäkelä and H. Mayer, *Bacteriol. Rev.*, 40 (6) (1976) 591-632.

30. C. Lugowski, E. Romanowska, L. Kenne, and B. Lindberg, *Carbohydr. Res.*, 118 (1983) 173-181.

31. A. Dell, J. Oates, C. Lugowski, E. Romanowska, L. Kenne, and B. Lindberg, *Carbohydr. Res.*, 133 (1984) 95-104.

32. S. Basu, H.M. Kuhn, A. Neszmélyi, K. Himmelspach, and H. Mayer, *Eur. J. Biochem.*, 162 (1987) 75-81.

33. C. Lugowski and E. Romanowska, *FEMS Microbiol. Lett.* 77 (1991) 315-318.

34. E.V. Vinogradov, Y.A. Knirel, J.E. Thomas-Oates, A.S. Shashkov, and V.L. L'vov, *Carbohydr. Res.*, 258 (1994) 223-232.

35. H. Smith, *Can. J. Microbiol*, 38 (1992) 747-752.

36. E.J. McGuire and S.B. Binkley, *Biochemistry*, 3 (1964) 247-251.

37. M.A. O'Neill, P.D. Robison, K.J. Chou, A.G. Darvill, and P. Albersheim, *Carbohydr. Res.*, 226 (1992) 131-154.

38. P.J. Garegg, B. Lindberg, T. Onn, and T. Holme, *Acta Chem. Scand.*, 25 (1971) 1185-1194.

39. P.J. Garegg, B. Lindberg, T. Onn, and L.W. Sutherland, *Acta Chem. Scand.*, 25 (1971) 2103-2108.

40. L.A.S. Parolis, H. Parolis, G.G.S. Dutton, P.L. Wing, and B.J. Skura, *Carbohydr Res.*, 216 (1991) 495-504.

41. A. Linker and R.S. Jones, *J. Biol. Chem.*, 241 (1966) 3845-3851.

42. P.A.J. Gorin and J.F.I. Spencer, *Can. J. Chem.*, 44 (1966) 993-998.

43. B. Larsen and A. Haug, *Carbohydr. Res.*, 17 (1971) 287-296.

44. I. Ørskov, F. Ørskov, B. Jann and K. Jann, *Bacteriol. Rev.*, 41 (1977) 667-710.

45. M.M. Levine in *Bacterial Vaccines*, edited by R. Germanier (Academic Press, London, 1984) pp. 187-235.

46. F. Ørskov and I. Ørskov in *Methods Microbiol.*, vol. 14 (Academic Press, London, 1984) pp. 43-112.

47. G.G.S. Dutton and L.A.S Parolis, in *Biomedical and Biotechnological Advances in Industrial Polysaccharides*, edited by V. Crescenzi, I.C.M. Dea, S. Paoletti, S.S. Stivala and I.W. Sutherland, (Gordon and Breach, New York, 1989) pp. 223-240.

48. A.H. de Bruin, *MSc Thesis*, Rhodes University, Grahamstown, 1992.

49. M.R. Lifely, J.C. Lindon, J.M. Williams, and C. Moreno, *Carbohydr. Res.*, 143 (1985) 191-205.

50. D.V. Whittaker, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, in press.

51. F.-P. Tsui, R.A. Boykins, and W. Egan, *Carbohydr. Res.*, 102 (1982) 263-271.

52. L.Tarcsay, B. Jann, and K. Jann, *Eur. J. Biochem.*, 23 (1971) 505-514.

53. T. Dengler, B. Jann, and K. Jann, *Carbohydr. Res.*, 142 (1985) 269-276.

54. T. Dengler, K. Himmelspach, B. Jann, and K. Jann, *Carbohydr. Res.*, 178 (1988) 191-201.

55. D.V. Whitakker, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 256 (1994) 289-301.

56. B. Jann and K. Jann, *Current Topics Microbiol. Immunol.*, 150 (1990) 19-42.

57. A.N. Anderson, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 163 (1987) 81-90.

58. A.H. de Bruin, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 233 (1992) 195-204.

59. B. Jann, P. Hofmann, and K. Jann, *Carbohydr. Res.*, 120 (1983) 131-141.

60. P. Hofmann, B. Jann and K. Jann, *Eur. J. Biochem.*, 147 (1985) 601-609.

61. T. Dengler, B. Jann and K. Jann, *Carbohydr. Res.*, 150 (1986) 233-240.

62. L.M. Beynon, G.G.S. Dutton, and J.C. Richards, *Carbohydr. Res.*, 205 (1990) 347-359.

63. P. Hofmann, B. Jann, and K. Jann, *Carbohydr. Res.*, 139 (1985) 261-271.

64. K. Jann and B. Jann, *Reviews Inf. Dis.*, 1 (1987) S517-S526.

65. K. Jann and B. Jann, *Can. J. Microbiol.*, 38 (1992) 705-710.

66. A.N. Anderson and H. Parolis, *Carbohydr. Res.*, 188 (1989) 157-168.

67. F.P. Tsui, W. Egan, M.F. Summers, R.A. Byrd, R. Schneerson, and J.B. Robbins, *Carbohydr. Res.*, 173 (1988) 65-74.

68. M.L. Rodriguez, B. Jann, and K. Jann, *Carbohydr. Res.*, 173 (1988) 243-253.

69. G.G.S. Dutton, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 170 (1987) 193-206.

70. A. Finke, B. Jann, and K. Jann, *FEMS Microbiol. Lett.*, 69 (1990) 129-134.

71. M.A. Schmidt and K. Jann, *FEMS Microbiol. Lett.*, 14 (1982) 69-78.

72. B. Jann, H. Kochanowski, and K. Jann, *Carbohydr Res.*, 253 (1994) 323-327.

73. O. Westphal and K. Jann, *Methods Carbohydr. Chem.*, 5 (1965) 83-91.

74. G.O. Aspinall in *The Polysaccharides*, vol. 1, edited by G.O. Aspinall (Academic Press, 1983) pp. 19-34.

75. H.O. Bouveng and B. Lindberg, *Adv. Carbohydr. Chem.* 15 (1960) 53-89.

76. J.H. Pazur and L.S. Forsberg, *Methods Carbohydr. Chem.*, 8 (1980) 211-217.

77. H.E. Conrad, *Methods Carbohydr. Chem.*, 6 (1972) 154-157.

78. K. Okutani and G.G.S. Dutton, *Carbohydr. Res.*, 86 (1980) 259-271.

79. F. Ørskov, V. Sharma, and I. Ørskov, *J. Gen. Microbiol.*, 130 (1984) 2681-2684.

80. A.J. Erskine and J.K.N. Jones, *Can. J. Chem.*, 34 (1956) 821-826.

81. S.A. Barker, M. Stacey, and G. Zweifel, *Chem. and Ind.* (1957) 330.

82. S.C. Churms, *Adv. Carbohydr. Chem. Biochem.*, 25 (1970) 13-51.

83. R.L. Whistler and A.K.M. Anisuzzaman, *Methods Carbohydr. Chem.*, 3 (1963) 45-53.

84. S.C. Churms *C.R.C. Handbook of Chromatography*, edited by G. Zweig and J. Sherma (C.R.C. Press Inc., 1982).

85. M. Dubois, K.A. Gilles, J.K. Hamilton, P.A. Rebers, and F. Smith, *Anal. Chem.*, 28 (1956) 350-356.

86. G.O. Aspinall in *The Polysaccharides*, vol. 1, edited by G.O. Aspinall (Academic Press, 1982) pp. 35-132.

87. D.A. Bastin, G. Stevenson, P.K. Brown, A. Haase, and P.R. Reeves, *Molecular Microbiology*, 7(5) (1993) 725-734.

88. G.O. Aspinall, E.L. Hirst, E.G.V. Percival, and R.G.J. Telfer, *J. Chem. Soc.* (1953) 337-342.

89. P. Albersheim, D.J. Nevins, P.D. English, and A. Karr, *Carbohydr. Res.*, 5 (1967) 340-345.

90. M.P. Sanger and D.T.A. Lamport, *Anal. Biochem.*, 128 (1983) 66-70.

91. L. Hough, J.V.S. Jones, and P. Wusteman, *Carbohydr. Res.*, 21 (1972) 9-17.

92. S. Manna, B.H. McAnalley, and H.L. Ammon, *Carbohydr. Res.*, 243 (1993) 11-27.

93. J-R. Neeser and T.F. Schweizer, *Anal. Biochem.*, 142 (1984) 58-67.

94. J.D. Blake and G.N. Richards, *Carbohydr. Res.*, 8 (1968) 275-281.

95. E.J. Roberts, M.A. Godshall, M.A. Clarke, W.S.C. Tsang, and F.W. Parrish, *Carbohydr. Res.*, 168 (1987) 103-109.

96. R.L. Taylor and H.E. Conrad, *Biochemistry*, 11 (8) (1972) 1383-1388.

97. G.O. Aspinall, E.L. Hirst, and N.K. Matheson, *J. Chem. Soc.*, (1956) 989.

98. L. Kenne, B. Lindberg, M.M. Rahman, and M. Mosihuzzaman, *Carbohydr. Res.*, 242 (1993) 181-189.

99. T.J. Painter, *Methods Carbohydr. Chem.*, 5 (1965) 280-285.

100. Y.F. Wang, D.P. Wittner, C. Dorschel, and C.G. Hellerqvist, *Abstracts of XVIth International Carbohydrate Symposium*, Paris, France, 5-10 July 1992, p. C164.

101. M.L. Wolfram and A. Thompson, *Methods Carbohydr. Chem.*, 3 (1963) 150-153.

102. C.J. Bierman, *Adv. Carbohydr. Chem. and Biochem.*, 46 (1988) 251-271.

103. Y.A. Knirel, E.V. Vinogradov, and A.J. Mort, *Adv. Carbohydr. Chem. Biochem.*, 47 (1989) 167-202.

104. J-R. Neeser, *Carbohydr. Res.*, 138 (1985) 189-198.

105. A.T. James and A.J.P. Martin, *Biochem. J.*, 50 (1952) 679-690.

106. M. Novotny, *Anal. Chem.*, 50 (1) (1978) 16A-32A.

107. *High Resolution Gas Chromatography*, 2nd Edition, edited by R.R. Freeman (Hewlett Packard USA, 1981) 2-3.

108. B. Newton in *High Resolution Gas Chromatography*, 3rd Edition, edited by K.J. Hyver (Hewlett-Packard Corporation USA, 1989), Chapter 2.

109. M.L. Duffy, *International Laboratory* (1986) 78-87.

110. K. Grob and G. Grob, *J. Chromatogr.*, 213 (1981) 211-221.

111. M.F. Laker, *J. Chromatogr.*, 184 (1980) 457-470.

112. M. Wilson in *High Resolution Gas Chromatography*, 3rd Edition, edited by K.J. Hyver (Hewlett-Packard Corporation USA, 1989), Chapter 4.

113. D.P. Sweet, R.H. Shapiro, and P. Albersheim, *Carbohydr. Res.*, 40 (1975) 217-225.

114. *High Resolution Chromatography Products*, J&W Scientific, California, USA, 1988/89.

115. J. Drozd, *J. Chromatogr.*, 113 (1975) 303-316.

116. C.T. Bishop, *Adv. Carbohydr. Chem.*, 19 (1964) 95-147.

117. K. Bryn and E. Jantzen, *J. Chromatogr.*, 240 (1982) 405-413.

118. C.C. Sweeley, R. Bentley, M. Makita, and W.W. Wells, *J. Am. Chem. Soc.*, 85 (1963) 2497-2507.

119. R.A. Laine and C.C. Sweely, *Carbohydr. Res.*, 27 (1973) 199-213.

120. G.G.S. Dutton, *Adv. Carbohydr. Chem. Biochem.*, 28 (1973) 11-160.

121. T.P. Mawhinney, M.S. Feather, J.R. Martinez, and G.J. Barbero, *Anal. Biochem.*, 101 (1980) 112-117.

122. A.G.W. Bradbury, D.J. Halliday, and D.G. Metcalf, *J. Chromatogr.*, 213 (1981) 146-150.

123. S. Honda, N. Yamauchi, and K. Kakehi, *J. Chromatogr.*, 169 (1979) 287-293.

124. H.J. Chaves Das Neves, A.M.V. Riscado, and H. Frank, *Carbohydr. Res.*, 152 (1986) 1-6.

125. J.D. Blake and G.N. Richards, *Carbohydr. Res.*, 14 (1970) 375-387.

126. R.H. Furneaux, *Carbohydr. Res.*, 113 (1983) 241-255.

127. C.C. Chen and G.D. McGinnis, *Carbohydr. Res.*, 90 (1981) 127-130.

128. F.R. Seymour, E.C.M. Chen, and J.E. Stouffer, *Carbohydr. Res.*, 83 (1980) 201-242.

129. T.P. Mawhinney, M.S. Feather, J.R. Martinez, and G.J. Barbero, *Carbohydr. Res.*, 75 (1979) C21-C23.

130. K.B. Hicks, *Adv. Carbohydr. Chem. Biochem.*, 46 (1988) 17-72.

131. C.K. Lim in *HPLC of Small Molecules*, edited by C.K. Lim (IRL, Oxford, 1986) pp. 1-68.

132. G.D. McGinnes and P. Fang, *Methods Carbohydr. Chem.*, 8 (1980) 33-43.

133. A.A. Ben-Bassat and E. Grushka, *J. Liquid Chromatogr.*, 14(6) (1991) 1051-1111.

134. Y.C. Lee, *Methods Enzymol.*, 28 (1982) 63-73.

135. C.C. Yu-Ip, V. Manam, R. Hepler, and J.P. Hennessey, Jr., *Anal. Biochem.*, 201 (1992) 343-349.

136. M.R. Hardy, R.R. Townsend, and Y.C. Lee, *Anal. Biochem.*, 170 (1988) 54-62.

137. T.J. O'Shea, S.M. Lunte, and W.R. LaCourse, *Anal. Chem.*, 65 (1993) 948-951.

138. M. Gruter, B.R. Leeflang, J. Kuiper, J.P. Kammerling, and J.F.G. Vliegenthart, *Carbohydr. Res.*, 239 (1993) 209-226.

139. L.H. Slaughter and D.P. Livingston III, *Carbohydr. Res.*, 253 (1994) 287-291.

140. L. Hough and J.K.N. Jones, *Methods Carbohydr. Chem.*, 1 (1962) 21-31.

141. J. Sherma, *Anal. Chem.*, 60 (1988) 74R-86R.

142. L.D. Hall, *Adv. Carbohydr. Chem. Biochem.*, 29 (1974) 11-40.

143. T.A.W. Koerner, J.H. Prestgard, and R.K. Yu, *Methods Enzymol.*, 138 (1987) 38-59.

144. K. Bock and H. Thørgersen, *Annu. Rep. in NMR Spectrosc.*, 13 (1982) 1-57.

145. A.J. Charlson, P.A.J. Gorin, and A.S. Perlin, *Methods Carbohydr. Chem.*, 1 (1962) 419-426.

146. C. Hudson, *J. Am. Chem. Soc.*, 31 (1909) 66-86.

147. G.M. Bebault, J.M. Berry, Y.M. Choy, G.G.S. Dutton, N. Funnell, L.D. Hayward, and A.M. Stephen, *Can. J. Chem.*, 51 (1973) 324-326.

148. G. Avigad, D. Amaral, C. Asensio, and B.L. Horecker, *J. Biol. Chem.*, 237 (1962) 2736-2743.

149. K. Leontein, B. Lindberg, and J. Lönngren, *Carbohydr. Res.*, 62 (1978) 359-362.

150. G.J. Gerwig, J.P. Kamerling, and J.F.G. Vliegenthart, *Carbohydr. Res.*, 62 (1978) 349-357.

151. M.R. Little, *Carbohydr. Res.*, 105 (1982) 1-8.

152. H. Schweer, *J. Chromatogr.*, 243 (1982) 149-152.

153. R. Oshima, Y. Yamauchi, and J. Kumanotani, *Carbohydr. Res.*, 107 (1982) 169-176.

154. W.A. König, P. Mischnik-Lübbecke, B. Brassat, S. Lutz, and G. Wenz, *Carbohydr. Res.*, 183 (1988) 11-17.

155. A.S. Shashkov, G.M. Lipkind, Y.A. Knirel, and N.K. Kochetkov, *Magn. Reson. Chem.*, 26 (1988) 735-747.

156. H. Baumann, A.O. Tzianabos, J-R. Brisson, D.L. Kasper, and H.J. Jennings, *Biochemistry*, 31 (1992) 4081-4089.

157. J. Hoffman, B. Lindberg, and S. Svensson, *Acta Chem. Scand.*, 26 (1972) 661-666.

158. M.A. Schmidt and K. Jann, *Eur. J. Biochem.*, 131 (1983) 509-517.

159. J.J. Marshall, *Adv. Carbohydr. Chem. Biochem.*, 30 (1974) 276-370.

160. N.K. Matheson and B.V. McCleary in *The Polysaccharides*, vol. 3, edited by G.O. Aspinall (Academic Press New York, 1985) pp. 2-94.

161. Y.M. Choy, F. Fehmel, N. Frank, and S. Stirm, *J. Virol.*, 16 (1975) 581-590.

162. H. Parolis, L.A.S. Parolis, and D.V. Whittaker, *Carbohydr. Res.*, 231 (1992) 93-103.

163. D.R. Bundle and R. Lemieux, *Methods Carbohydr. Chem.*, 7 (1976) 79-86.

164. B. Lindberg, *Methods Enzymol.*, 28 (1972) 178-195.

165. H. Rauvala, J. Finne, T. Krusius, J. Kärkkäinen, and J. Järnefelt, *Adv. Carbohydr. Chem. and Biochem.*, 38 (1981) 389-407.

166. S-I. Hakomori, *J. Biochem. (Tokyo)*, 55 (2) (1964) 205-208.

167. P.A. Sandford and H.E. Conrad, *Biochemistry*, 5 (1966) 1508-1517.

168. J. Finne, T. Krusius, and H. Rauvala, *Carbohydr. Res.*, 80 (1980) 336-339.

169. L.R. Phillips and B.A. Fraser, *Carbohydr. Res.*, 90 (1981) 149-152.

170. A.B. Blakeney and B.A. Stone, *Carbohydr. Res.*, 140 (1985) 319-324.

171. T. Narui, K. Takahashi, M. Kobayashi, and S. Shibata, *Carbohydr. Res.*, 103 (1982) 293-295.

172. J-P. Joseleau, G. Chambat, and B. Chumpitazi-Hermoza, *Carbohydr. Res.*, 5 (1967) 340-345.

173. P.J. Harris, R.J. Henry, A.B. Blakeney, and B.A. Stone, *Carbohydr. Res.*, 127 (1984) 59-73.

174. I. Ciucanu and C. Luca, *Carbohydr. Res.*, 206 (1990) 71-77.

175. U. Zähringer and E.T. Rietschel, *Carbohydr. Res.*, 152 (1986) 81-87.

176. I. Ciucanu and F. Kerek, *Carbohydr. Res.*, 131 (1984) 209-217.

177. W.S. York, L.L. Kiefer, P. Albersheim, and A.G. Darvill, *Carbohydr. Res.*, 208 (1990) 175-182.

178. P.W. Needs and R.R. Selvendran, *Carbohydr. Res.*, 245 (1993) 1-10.

179. R. Kuhn, H. Trischmann, and I. Low, *Angew. Chem.*, 67 (1955) 32.

180. P. Prehm, *Carbohydr. Res.*, 78 (1990) 372-374.

181. P-E. Jansson, L. Kenne, H. Liedgren, B. Lindberg, and J. Lonngren, *Chem. Commun.*, 8 (1976) 1-75.

182. K. Stellner, H. Saito, and S.-I. Hakomori, *Arch. Biochem. Biophys.*, 155 (1973) 464-472.

183. E.T. Oakley, D.F. Magin, G.H. Bokelman, and W.S. Ryan, Jr., *J. Carbohydr. Chem.*, 4(1) (1985) 53-65.

184. P.W. Needs and R.R. Selvendran, *Carbohydr. Res.*, 254 (1994) 229-244.

185. D. Rolf and G.R. Gray, *J. Am. Chem. Soc.*, 104 (1982) 3539-3541.

186. G.R. Gray, *Methods in Enzymol.*, 138 (1987) 26-38.

187. A.J. D'Ambra and G.R. Gray, *Carbohydr. Res.*, 251 (1994) 115-125.

188. A.J. D'Ambra and G.R. Gray, *Carbohydr. Res.*, 251 (1994) 127-144.

189. S. Miertus, L. Navarini, A. Cesàro, *Carbohydr Res.*, 257 (1994) 227-238.

190. R. Albert, K. Dax, A.E. Stütz, and H. Weidmann, *J. Carbohydr. Chem.*, 2(3) (1983) 279-292.

191. J.A. Lomax, A.H. Gordon, and A. Chesson *Carbohydr. Res.*, 122 (1983) 11-22.

192. W.R.D. Leigh and Z.S. Krzeminski, *J. Chem Soc. (C)*, (1966) 1700-1703.

193. J.D. Stanikowski and S.G. Zeller, *Carbohydr. Res.*, 224 (1992) 237-241.

194. P.K. Agrawal, H.-J. Schneider, M.S. Malik, and S.N. Rostogi, *Org. Magn. Reson.*, 21 (1984) 146.

195. A. Bax and M.F. Summers, *J. Am. Chem. Soc.*, 108 (1986) 2092-2094.

196. S.G. Zeller and G.R. Gray, *Carbohydr. Res.*, 198 (1990) 285-303.

197. T. Fontaine, F. Talmont, G.G.S. Dutton, and B. Fournet, *Anal. Biochem.*, 199 (1991) 154-161.

198. P.J. Garegg, P.-E. Jansson, B. Lindberg, F. Lindh, J. Lönngren, I. Kvarnström, and W. Nimmich, *Carbohydr. Res.*, 78 (1980) 127-132.

199. C. Jones, *Carbohydr Res.*, 198 (1990) 353-357.

200. N. Ravenscroft, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 254 (1994) 333-340.

201. V.J. Basus, *Methods Enzymol.*, 177 (1989) 132-149.

202. P.S. Chen, T.Y. Toribara, and H. Warner, *Anal. Chem.*, 28(11) (1956) 1756-1758.

203. B.N. Ames, *Methods Enzymol.*, 8 (1966) 115-118.

204. A.J. Mort and D.T.A. Lamport, *Anal. Biochem.*, 82 (1977) 289-309.

205. D.R. Bundle, H.J. Jennings, and I.C.P. Smith, *Can. J. Chem.*, 51 (1973) 3812-3819.

206. B. Lindberg, J. Lönngren, and S. Svensson, *Adv. Carbohydr. Chem. Biochem.*, 31 (1975) 185-240.

207. A.J. Mort and W.D. Bauer, *J. Biol. Chem.*, 257 (1982) 1870-1875.

208. A.J. Mort, J-P. Utille, G. Torri, and A.S. Perlin, *Carbohydr. Res.*, 121 (1983) 221-232.

209. V.N. Reinhold, E. Coles, and S.A. Carr, *J. Carbohydr. Chem.*, 2(1) (1983) 1-18.

210. J.U. Bowif, P.V. Trescony, and G.R. Gray, *Carbohydr. Res.*, 125 (1983) 301-307.

211. D. Rolf, J.A. Bennek, and G.R. Gray, *Carbohydr. Res.*, 137 (1985) 183-196.

212. J-G. Jun and G.R. Gray, *Carbohydr. Res.*, 163 (1987) 247-261.

213. J.A. Bennek, M.J. Rice, and G.R. Gray, *Carbohydr. Res.*, 157 (1986) 125-137.

214. S.A. Vodonik and G.R. Gray, *Carbohydr. Res.*, 175 (1988) 93-102.

215. S.M.R. Stanley, *PhD Thesis*, Rhodes University, 1990.

216. P.L. Hackland, *PhD Thesis*, Rhodes University, 1992.

217. J.M. Lau, M. McNeil, A.G. Darvill, and P. Albersheim, *Carbohydr. Res.*, 168 (1987) 219-243.

218. M. Aeroboe, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 248 (1993) 213-223.

219. L.M. Beynon and G.G.S. Dutton, *Carbohydr. Res.*, 200 (1990) 457-468.

220. G.G.S. Dutton and A. Kumah-Mintah, *Carbohydr. Res.*, 169 (1987) 213-220.

221. H. Parolis, L.A.S. Parolis, and R.D. Venter, *Carbohydr. Res.*, 185 (1989) 225-232.

222. A.H. de Bruin, H. Parolis, and L.A.S. Parolis, *Carbohydr. Res.*, 235 (1992) 199-209.

223. D.V. Whitakker, *PhD Thesis*, Rhodes University, 1993.

224. J. Douglas, *Bacteriophages* (Chapman and Hall, London, 1975).

225. B. Kwiatkowski, B. Boschek, H. Thiele, and S. Stirm, *J. Virol.*, (1983) 367-374.

226. S. Stirm and E. Freund-Mölbert, *J. Virol.*, 8 (1971) 330-342.

227. D. Rieger-Hug and S. Stirm, *Virology*, 113 (1981) 363-378.

228. J.E.G. van Dam, H. van Halbeek, J.P. Kammerling, J.F.G. Vliegenthart, H. Snippe, M. Jansze, and J.M.N. Willers, *Carbohydr. Res.*, 142 (1985) 338-343.

229. N. Ravenscroft, A.M. Stephen, and E.H. Merrifield, *Carbohydr. Res.*, 167 (1987) 257-267.

230. H. Parolis, L.A.S. Parolis, and G.G.S. Dutton, *Carbohydr. Res.*, 182 (1988) 127-134.

231. E. Altman, G.G.S. Dutton, and A.M. Stephen, *S. Afr. J. Sci.*, 82 (1986) 45-46.

232. G.G.S. Dutton, K.L. Mackie, A.V. Savage, D. Rieger-Hug, and S. Stirm, *Carbohydr. Res.*, 84 (1980) 161-170.

233. K.R. Yamamoto, B.M. Alberts, R. Benzinger, L. Lawhorne, and G. Treiber, *Virology*, 40 (1970) 733-744.

234. G.G.S. Dutton, J-L. Di Fabio, D.M. Leek, E.H. Merrifield, J.R. Nunn, and A.M. Stephen, *Carbohydr. Res.*, 97 (1981) 127-138.

235. R. Benn and H. Günther, *Angew. Chem. Int. Ed. Engl.*, 22 (1983) 350-380.

236. I.H. Sadler, *Natural Products Reports*, (1988) 101-127.

237. A.E. Derome, *Natural Products Reports*, (1989) 111-141.

238. R.E. Ernst, *Angew. Chemm. Int. Ed. Engl.*, 31 (1992) 805-930.

239. P.K. Agrawal, *Phytochemistry*, 31 (1992) 3307-3330.

240. A.E. Derome, *Modern NMR Techniques for Chemistry Research*, edited by J.E. Baldwin (Pergamon Press, 1987).

241. B. Capon and D. Thacker, *Proc. Chem. Soc.* (1964) 369.

242. F. Inagaki, I. Shimada, D. Kohada, A. Suzuki, and A Bax, *J. Magn. Reson.*, 81 (1989) 186-190.

243. C.J. Bauer, R. Freeman, T. Frenkiel, J. Keeler, and A.J. Shaka, *J. Magn. Reson.*, 58 (1984) 442-457.

244. H. Kessler, H. Oschkinat, C. Griesinger, and W. Bermel, *J. Magn. Reson.*, 70 (1986) 106-133.

245. H. Kessler, U. Anders, G. Gemmecker, and S. Steuernagel, *J. Magn. Reson.*, 85 (1989) 1-14.

246. R. Barker, H. Nunez, P. Rosevear, and A.S. Serianni, *Methods Enzymol.*, 83 (1974) 58-69.

247. S. Sørensen, R.S. Hansen, and H.J. Jakobsen, *J. Magn. Reson.*, 14 (1974) 243.

248. G.A. Morris and R. Freeman, *J. Am. Chem. Soc.*, 101 (1979) 760-762.

249. R. Freeman in *A Handbook of Nuclear Magnetic Resonance* (Longman Scientific and Technical, 1988) pp. 92-96.

250. E. Wenkert, A.O. Clouse, D.W. Cochran, and D. Doddrell, *J. Am. Chem. Soc.*, 91 (1969) 6879-80.

251. K. Bock and C. Pedersen, *J. Chem. Soc. Perkin Trans.*, 2 (1974) 293-297.

252. H.J. Jennings and I.C.P. Smith, *Methods Enzymol.*, 50 (1978) 39-50.

253. K. Bock and C. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 27-66.

254. J.H. Bradbury and G.A. Jenkins, *Carbohydr. Res.*, 126 (1984) 125-156.

255. S.L. Patt and J.N. Shoolery, *J. Magn. Reson.*, 46 (1982) 535-539.

256. D.M. Doddrell, D.T. Pegg, and M.R. Bendall, *J. Magn. Reson.*, 48 (1982) 323-327.

257. P.J. Hore, E.R.P. Zuiderweg, K. Micolay, K. Dijkstra, and R. Kaptein, *J. Am. Chem. Soc.*, 104 (1982) 4286-4288.

258. G.A. Morris, *Magn. Reson. in Chemistry*, 24 (1986) 371-403.

259. A. Bax and R. Freeman, *J. Magn. Reson.*, 46 (1982) 535-539.

260. A. Bax and G. Drobny, *J. Magn. Reson.*, 61 (1985) 306-320.

261. A. Bax and D.G. Davis, *J. Magn. Reson.*, 65 (1985) 355-360.

262. A. Bax, L.G. Marzilli, and M.F. Summers, *J. Am. Chem. Soc.*, 109 (1987) 566-574.

263. G. Wider, S. Macura, A. Kumar, R.R. Ernst, and K. Wuethrich, *J. Magn. Reson.*, 56 (1984) 207-234.

264. J. Breg, D. Romun, J.F.G. Vliegenthart, G. Strecker, and J. Montreuil, *Carbohydr. Res.*, 183 (1988) 19-34.

265. A. Bax and G.A. Morris, *J. Mag. Reson.*, 42 (1981) 501-505.

266. G.E. Martin and R.C. Crouch, *J. Nat. Prod.*, 54 (1991) 1-70.

267. L. Lerner and A. Bax, *J. Magn. Reson.*, 69 (1986) 365-380.

268. N.K. Kochetkov and O.S. Chizhov, *Methods in Carbohydr. Chem.*, 6 (1972) 540-554.

269. N.K. Kochetkov and O.S. Chizhov, *Adv. Carbohydr. Chem.*, 21 (1966) 39-93.

270. J. Lönngren and S. Svensson, *Adv. Carbohydr. Chem. Biochem.*, 29 (1974) 41-106.

271. A. Dell, *Adv. Carbohydr. Chem. Biochem.*, 45 (1987) 19-72.

272. H. Björndal, C.G. Hellerqvist, B. Lindberg, and S. Svensson, *Angew. Chem. Int. Ed. Engl.*, 9 (1970) 610-619.

273. B. Munson, *Anal. Chem.*, 49 (1977) 772A-778A.

274. V.N. Reinhold, *Methods Enzymol.*, 138 (1987) 59-86.

275. G.G.S. Dutton, Z. Lam, and A.V.S. Lim, *Carbohydr. Res.*, 183 (1988) 123-125.

276. R.S. Pappas, B.J. Sweetman, S. Ray, and C.G. Hellerqvist, *Carbohydr. Res.*, 197 (1990) 1-14.

277. A. Dell and P.R. Tiller, *Biochem. Biophys. Res. Commun.*, 135 (1986) 1126-1134.

278. P.L. Hackland, H. Parolis, A. Dell, and P.R. Tiller, *Carbohydr. Res.*, 181 (1988) 153-162.

279. S. Santikarn, G.R. Her, and V.N. Reinhold, *J. Carbohydr. Chem.*, 6 (1987) 141-154.

280. M.L. Coates and C.L. Wilkins, *Anal. Chem.*, 59 (1987) 197.

281. Z. Lam, M.B. Comisarow, G.G.S. Dutton, D.A. Weil, and A. Bjarnason, *Carbohydr. Res.*, 180 (1988) C1-C7.

282. Z. Lam, M.B. Comisarow, G.G.S. Dutton, H. Parolis, L.A.S. Parolis, A. Bjarnason, and D.A. Weil, *Analytica Chimica Acta*, 241 (1990) 187-199.

283. A. Bax and S. Subramanian, *J. Magn. Reson.*, 67 (1986) 565-569.

284. K. Izumi, *Carbohydr. Res.*, 170 (1987) 19-25.

285. P.A.J. Gorin and M. Mazurek, *Can. J. Chem.*, 53 (1975) 1212-1223.

286. E. Altman, M.B. Perry, and J-R. Brisson, *Carbohydr Res.*, 191 (1989) 295-303.

287. P-E. Jansson, H. Lenholm, B. Lindberg, U. Lindquist, and S. B. Svenson, *Carbohydr. Res.*, 161, (1987) 273-279.

288. J. L. Di Fabio, M. B. Perry, and D. R. Bundle, *Biochem. Cell. Biol.*, 65, (1987) 968-977.

289. P-E. Jansson, B. Lindberg, M. Spellman, T. Hofstad, and N. Skaug, *Carbohydr. Res.*, 137 (1985) 197-203.

290. E. V. Vinogradov, Z. Sidorczyk, A. Swierzko, A. Rożalski, E. D. Daeva, A. S. Shashkov, Y. A. Knirel, and N. K. Kochetkov, *Eur. J. Biochem.*, 197 (1991) 93-103

291. J. Hickman and G. Ashwell, *J. Biol. Chem.*, 241 (1966) 1424-1428.

292. A. C. Richardson, *Carbohydr. Res.*, 4 (1967) 415-421.

293. W.E. Trevelyan, D.P. Procter, and J.S. Harrison, *Nature*, 166, (1950) 444.

294. S.C. Churms, in *Handbook of Chromatography*, edited by G. Zweig and J. Sherma (CRC Press Inc. Boca Raton, Florida, 1982) vol 1, p193.

295. C.T. Bishop and H.J Jennings, in *The Polysaccharides*, vol 2, edited by G.O. Aspinall (Academic Press Inc., 1982) 291-330.

296. P.J. Garegg, B. Lindberg, and I. Kvarnström, *Carbohydr. Res.*, 77 (1979) 71-78.

297. B. Jann, T. Dengler, and K. Jann, *FEMS Microbiol. Lett.*, 29 (1985) 257-261.

298. M.L. Rodriguez, B. Jann, and K. Jann, *Carbohydr. Res.*, 196 (1990) 101-109.

299. A.N. Anderson, H. Parolis and L.A.S. Parolis, *Carbohydr. Res.*, 163 (1987) 81-90, and references cited therein.

300. H. Kessler, P. Schmieder, and M. Kurtz, *J. Magn. Reson.*, 85 (1989) 400-405.

301. D.L. Rabenstein and W. Guo, *Anal. Chem.*, 60 (1988) 1R-28R.

302. B.A. Lewis, unpublished results.

303. A. Berthold, W.Y. Li, and D. Armstrong, *Carbohydr. Res.*, 201 (1990) 175-184.

304. A. Bax and D.G. Davis, *J. Magn. Reson.*, 63 (1985) 207-213.

305. U. Piantini, O.W. Sørensen, and R.R. Ernst, *J. Am. Chem. Soc.*, 104 (1982) 6800-6801.

306. S.P. Rucker and A.J. Shaka, *Mol. Phys.*, 68 (1989) 509-517.

307. R. Baumann, G. Wider, R.R. Ernst, and K. Würthrich, *J. Magn. Reson.*, 44 (1981) 402-406.

308. D.J. States, R.A. Haberkorn, and D.J. Ruben, *J. Magn. Reson.*, 48 (1982) 286-292.

309. M. Kjaer, K.V. Anderson, and F.M. Poulsen, *Methods Enzymol.*, (1994) in press.

310. H. Parolis, L.A.S. Parolis, S.M.R. Stanley, and G.G.S. Dutton, *Carbohydr. Res.*, 205 (1990) 361-370.

311. K. Jann, B. Jann, F. Ørskov, and I. Ørskov, *Biochem. Z.*, 346 (1966) 368-385.

312. G. Bodenhausen, R. Freeman, G.A. Morris, R. Niedermeyer, and D.L. Turner, *J. Magn. Reson.*, 25 (1977) 559.

313. A. Bax, *J. Magn. Reson.*, 52 (1983) 330-334.