Structural analysis of capsular polysaccharides produced by some emerging *Klebsiella pneumoniae* serotypes

A dissertation submitted to the

University of Cape Town

in fulfilment of the requirements for the degree of

Master of Science in Chemistry

By

Siwaphiwe Mfana

Supervisors:

A/Prof Neil Ravenscroft &

Dr Cesarina Edmonds-Smith

Department of Chemistry

February 2024

Plagiarism declaration

I, Siwaphiwe Mfana, hereby declare that the work on which this thesis “Structural analysis of capsular polysaccharides produced by some emerging *Klebsiella pneumoniae* serotypes” is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorise the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature…….

Date: February 2024

Acknowledgements

I would like to extend my sincere appreciation to the following individuals and institutions whose unwavering support and contributions have given rise to the accomplishment of this work:

First and foremost, I express my deepest gratitude to my supervisors, A/Prof Neil Ravenscroft and Dr Cesarina Edmonds-Smith, for sharing their expertise, providing outstanding guidance and exceptional supervision, and for their remarkable patience throughout the duration of this project.

I would like to sincerely thank Nicole Richardson (PhD), a member of our research group, for providing insightful advice, willingness to assist, and for meticulously proofreading this work.

Special thanks to our collaborators at the University of Trieste, Italy (Paola Cescutti), and the GVGH Vaccines Institute for Global Health in Siena, for providing the capsular polysaccharide samples of the strains investigated in this work.

I am deeply grateful to my funders: National Research Foundation (NRF) scholarship; Cipla SA bursary; University of Cape Town, Department of Chemistry (Supervisors), for their financial assistance in funding this MSc degree.

I would also like to express my sincere and heartfelt appreciation to my family for their continuous unwavering support, encouragement, and motivation throughout my academic journey. Firstly, special thanks to my late father, Mr Ngoniwa Mfana, whose memory continues to inspire me to do the best I can, and my mother, Mrs Thembela Mfana, for her enduring support. Additionally, I appreciate my siblings: Mandisi, Nkosikho, Zenande, and Chwayita Mfana, as well as my nieces: Avuzwa and Iminathi Mfana, for all their support.

Dlamini, Zizi, Jama kas'Jadu, Fakade, iZiz'elimnyama nenkomo zalo… Ndiyacamagwisha!!

Abstract

Klebsiella pneumoniae (*K. pneumoniae*), a Gram-negative pathogenic bacterium, is a leading cause of neonatal sepsis in low- and middle-income countries (LMICs). Due to rapidly increasing anti-microbial resistance (AMR), it has been ranked among the "critical-priority 1" pathogens to human health by the World Health Organisation (WHO). *K. pneumoniae* produces a capsular polysaccharide (CPS or K-antigen), which constitutes an important virulence factor and a potential vaccine antigen. To date, there are up to 77 known distinct *K. pneumoniae* K-antigen serotypes, that have been classified through serotyping. Furthermore, there are additional K-locus (KL) serotypes that have been recently identified using genotyping.

This project involved the characterisation and structural elucidation of the CPS repeating unit (RU) structures, known as chemotyping, of four emerging clinically significant strains of *K. pneumoniae*, which were identified through genotyping as novel serotypes: KL102; KL112; KL122; and KL107. NMR spectroscopy provides a non-destructive method for the structural elucidation of polysaccharides on relatively small amounts of samples. A combination of 1D and 2D homo and heteronuclear NMR experiments can be used to determine the monosaccharide composition, anomeric configurations (α or β), linkage positions and sequence of residues in the polysaccharide RU. NMR can also identify the presence and positions of non-carbohydrate substituents including O-acetyl and pyruvate groups.

Complete NMR characterisation and structural elucidation was successfully achieved for all four capsular polysaccharides investigated. Their proposed RU structures were compared to those of known *K. pneumoniae* K-antigens, using a database that was constructed during this investigation, to confirm if they were novel serotypes. KL102 CPS has a novel hexasaccharide RU, made up of a trisaccharide backbone, and is doubly-branched with a disaccharide and a single terminal residue side chains, this can be categorised as a 3 + 2 + 1 RU type. KL112 CPS has a novel and unusual pentasaccharide RU, consisting of a disaccharide backbone chain and a trisaccharide side chain, categorised as a 2 + 3 RU type. KL122 CPS also has a novel hexasaccharide RU containing three uronic acids and is made up of a tetrasaccharide backbone chain with two terminal monosaccharide residues, thus a 4 + 1 + 1 RU type. In contrast, KL107 CPS was found to be identical to serotype K2, having a tetrasaccharide RU, with a trisaccharide backbone chain and a single terminal residue with O-acetylation on the branched backbone residues. The diagnostic NMR data acquired for these CPSs can be used for serotype identity testing, and to monitor the structures of K-antigens during the process of glycoconjugate vaccine development.

Table of Contents

List of Figures

List of Tables

Abbreviations and symbols

^{13}C	Carbon-13
1D or 2D	One- or Two-Dimensional
^{1}H (H)	Proton
α	Alpha
AMR	Anti-Microbial Resistant (Resistance)
β	Beta
CASPER	Computer-Assisted Spectrum Evaluation of Regular Polysaccharides
CDC	Centers for Disease Control and Prevention
COSY	Correlation Spectroscopy
CPS	Capsular polysaccharide
D_2O	Deuterium Oxide
DEPT	Distortionless Enhancement by Polarization Transfer
DOSY	Diffusion Ordered Spectroscopy
EOS	Early-Onset Sepsis
ESBLs	Extended-Spectrum β-Lactamases
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter species*
ExPEC	Extraintestinal Pathogenic *Escherichia Coli*
Fuc	Fucose
Gal	Galactose
GalA	Galacturonic acid
GC	Gas Chromatography
Glc	Glucose
GLC	Gas-Liquid Chromatography
GlcA	Glucuronic acid
HDO	Hydrogen-Deuterated Oxide (Partially deuterated water)
HMBC	Heteronuclear Multiple Bond Correlation
HPAEC-PAD	High-Performance Anion Exchange Chromatography with Pulsed Amperometric Detection
HPLC	High-Performance Liquid Chromatography
HSQC	Heteronuclear Single-Quantum Correlation Spectroscopy
Hz	Hertz
IUPAC	International Union of Pure and Applied Chemistry

K	Kelvin
KetoA	Keto acid
KL	K-locus
KPC	*K. pneumoniae* Carbapenemases
LMICs	Low- and Middle-Income Countries
LOS	Late-Onset Sepsis
LPS	Lipopolysaccharide
Man	Mannopyranose
MDR	Multi-Drug Resistant/Resistance
mg	milligrams
MHz	Mega Hertz
MLST	Multi-locus Sequence Typing
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
NUS	Non-Uniform Sampling
PCR	Polymerase Chain Reaction
PFGE	Pulsed-Field Gel Electrophoresis
PMAAs	Partially methylated alditol acetates
Pre-sat	Pre-saturation
PyrA	Pyruvic acid
RFLP	Restriction Fragment Length Polymorphism
Rha	Rhamnose
RU	Repeating Unit
SNFG	Symbol Nomenclature for Glycans
STs	Sequence Types
TOCSY	Total Correlation Spectroscopy
UTIs	Urinary Tract Infections
WGS	Whole Genome Sequencing
WHO	World Health Organization

Chapter 1: Introduction

Klebsiella pneumoniae (K. pneumoniae) is a Gram-negative, encapsulated, and non-motile bacterium, and a leading cause of hospital-acquired infections and neonatal sepsis in low- and middle-income countries (LMICs).[1] This pathogen has gained recognition as a concerning threat to human health, due to the rapid emergence of multi-drug resistant (MDR) and hypervirulent strains, causing high rates of morbidity and mortality globally.[2] The external surface of the bacteria is comprised of a capsular polysaccharide (CPS, K-antigen) and a lipopolysaccharide (LPS, O-antigen), which are the main virulence factors and the basis for the classification of strains.[3] There is an urgent need for an effective *K. pneumoniae* glycoconjugate vaccine to provide protective immunisation to neonates, the elderly and immunocompromised individuals, to help reduce the use of antimicrobials, and consequently, the spread of resistance.[4] The CPS is not only established as the most important virulence factor but has been demonstrated to be an effective immunological and potential vaccine antigen.[5]

Up to 77 *K. pneumoniae* serotypes have been identified based on the unique and surface-exposed K-antigens produced by different strains using a technique termed serotyping.[6] However, due to its challenges and limitations, serotyping is no longer widely used, now genotyping has become established as a more effective approach to identify K-antigen serotypes and distinguish clinical isolates.[7] Using genotyping, a set of additional K-antigen serotypes, known as K-locus (KL) series, have been identified based on the analysis of specific genes (K-loci) responsible for the biosynthesis of the CPS.[7] The chemical structures of the 77 K-antigens, identified by the traditional serotyping approach, have been investigated and their repeating unit (RU) structures elucidated in various chemical studies in the literature (as shown in the **Appendix**, **Table 1A**), this process is referred to as chemotyping. Chemotyping specifically defines each K-antigen at a chemical-molecular level, enabling the understanding of their immunological behaviour. This investigation involved the characterisation and chemotyping of capsular polysaccharides of some emerging clinically important strains of *K. pneumoniae*, that have been identified by genotyping, using Nuclear Magnetic Resonance (NMR) spectroscopy.

This chapter introduces *K. pneumoniae*, highlighting the prevalence and effects of its infections, association with drug resistance, a brief discussion of the traditional serotyping and the modern genotyping methods used in the identification of isolates, as well as the importance of chemotyping the capsular polysaccharides. Moreover, the structures of *K. pneumoniae* CPSs are described, followed by the structural elucidation methods, highlighting the advantages of NMR spectroscopy compared to

wet-chemical analysis, and finally, a discussion of the need and current stage in the development of a conjugate vaccine against this pathogen.

1.1 Brief history of *K. pneumoniae* and its significance

In 1882, Carl Friedlander first discovered and described *K. pneumoniae* as an encapsulated bacillus, that he isolated from the lung tissues of patients deceased from pneumonia.[8] As a result, it was originally named Friedlander's bacillus and ultimately classified and renamed as a species of the *Klebsiella* genus in 1886, a name which honours the German microbiologist, Edwin Klebs.[9–11] *Klebsiella* species are members of the *Enterobacteriaceae* family and are present in various environments in nature.[11] *K. pneumoniae* is the most clinically important species within this genus and the species name, *pneumoniae*, was derived from its association with pneumonia, one of the primary diseases it is responsible for.[12] However, it can also cause other diseases, including urinary tract infections (UTIs) bloodstream infections (bacteraemia), and sepsis.[4,9,13] *K. pneumoniae* bacteria naturally form part of the human flora, particularly colonizing mucosal surfaces of the gastrointestinal tract and oropharynx, where they are non-symptomatic.[4,11,12] Moreover, they are also found in environmental water, especially sewage-contaminated water, and in the soil and plants.[12]

1.1.1 Classical *K. pneumoniae*

Historically, *K. pneumoniae* has been classified as an opportunist pathogen that is associated with nosocomial infections in neonates, the elderly, and other immunosuppressed individuals who are suffering from other significant comorbidities.[4,13] This is because classical *K. pneumoniae* does not typically cause disease in healthy individuals but uses the opportunity to spread in healthcare environments and infect patients with weakened immune systems. These patients are usually admitted for other reasons and do not have the infection at the time of admission. The types of hospital admissions include underlying diseases such as diabetes mellitus, hypertension, chronic lung disease, malignancy, liver and gallbladder diseases, renal failure, and alcoholism.[10] It has been reported that *K. pneumoniae* is responsible for about 11.8% of nosocomial pneumonia globally.[14]

This is directly correlated to an observed increase in carrier rates of this pathogen in humans who have been exposed to hospital environments.[12] Approximately 5 - 38% of humans in community settings carry *K. pneumoniae* as part of their gut microbiome, as detected from their stool samples, and 1 - 6% in their nasopharynx, commensally.[12] However, these colonisation rates increase dramatically in healthcare environments, and are directly proportional to the length of hospitalisation period, with 77% and 19% of patients reported to carry *K. pneumoniae* in their gut and pharynx, respectively.[12] Additionally, the patient's previous antibiotic therapy, especially with broad-spectrum or multiple

antibiotics, is significantly associated with acquiring *K. pneumoniae* in their gastrointestinal flora.[15] This hospital colonisation is a significant risk factor for *K. pneumoniae* diseases, as it has been identified that patients who have acquired hospital colonisation possess up to a four-fold increased infection rate compared to non-carriers.[15]

1.1.2 Hypervirulent *K. pneumoniae*

From the 1980s, reports about the emergence of hypervirulent *K. pneumoniae* strains began to rise, which, in contrast to classical *K. pneumoniae*, are significantly more invasive and cause community-acquired infections such as pneumonia.[9,13] These strains originated from Asia (Taiwan, China) and have continuously spread across the Asian Pacific Rim.[16] They are now increasingly reported in regions outside of Asia. The hypervirulent *K. pneumoniae* is described as strains evolved from classical *K. pneumoniae*, which have acquired hypervirulent genetic and phenotypic traits.[9] Their enhanced virulence and invasiveness enable them to cause infections in healthy, immunocompetent, individuals and allow them to establish more severe infections in other human organs, including pyogenic liver abscesses, central nervous system infections, and endophthalmitis.[9,16] Furthermore, they are capable of causing multiple site infections and may spread metastatically, as a result, they are associated with higher rates of mortality.

1.2 *K. pneumoniae* diseases

1.2.1 Prevalence of infections and antibiotic resistance

K. pneumoniae infections have become a significant human health threat in the last few decades, due to a concerning rise in antibiotic resistance and the emergence of MDR strains.[9] This has led to increasingly high mortality rates associated with *K. pneumoniae* and significant challenges in the healthcare treatment and management of its infections.[5] Approximately 50% of global mortality has been reported to be attributed to *K. pneumoniae* infections.[12] The World Health Organization (WHO) has recognised and ranked *K. pneumoniae* as a "critical-priority 1" antibiotic-resistant pathogen to human health.[17] *K. pneumoniae* is also classified as one of the "ESKAPE" pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter species*), a group of both Gram-negative and Gram-positive bacteria which have been found as leading causes of nosocomial infections and has been associated with MDR since 2008.[18] Due to this MDR, even nosocomial infections that have previously been simple to treat, such as UTIs, and wound or surgical site infections, have become more complicated and intractable, while severe infections like pneumonia and bacteraemia are increasingly life-threatening.[9]

Antibiotic resistance is mainly associated with classical *K. pneumoniae,* however, strains which possess both the traits of hypervirulence and multi-drug resistance have been reported in China, which is even more alarming.[9,19,20]

Due to hypervirulent strains, in Western countries, *K. pneumoniae* is estimated to be responsible for approximately 3 - 5% of all community-acquired pneumonia, whereas it is approximately 15% in developing countries such as Africa.[14] These are highly significant rates for a pathogen such as *K. pneumoniae*, which has previously only been associated with hospital-acquired infections and highlights its threatening evolution as a pathogen. Neonatal sepsis is one of the most important *K. pneumoniae* hospital-acquired infections, claiming countless lives of neonates, especially in LMICs.[21,22]

1.2.2 *K. pneumoniae* neonatal sepsis

Neonatal sepsis is a severe condition where a pathogen causes a systemic infection in a newborn.[23] This has been categorised as either early-onset sepsis (EOS) or late-onset sepsis (LOS). In the context of neonatal sepsis caused by pathogens other than group B *Streptococcu*s (GBS), EOS occurs when the disease appears within 72 hours after birth*,* while LOS disease manifests after 72 hours of life.[23] Moreover, EOS is attributed to infections acquired by the infant before or during birth and is usually caused by pathogens transmitted vertically from mother to infant. In contrast, LOS is usually caused by pathogens acquired horizontally from the hospital environment or community after birth.

In LMICs, *K. pneumoniae* is the most common and critical cause of both EOS and LOS and associated mortality,[24,25] this is demonstrated by data from the National Neonatal Perinatal Database in India.[26] Sepsis has been reported to kill around 1.6 million neonates per year, mainly in LMICs.[24] Estimations have indicated that about 800,000 neonates die within the first month of their birth due to sepsis annually, and *K. pneumoniae* is one of the main pathogens responsible for the infections.[1] Unsurprisingly, the prevalence of antibiotic-resistant *K. pneumoniae* strains is a high-risk factor which increases the susceptibility of neonates to infections and poses critical challenges in the clinical management of its infections.[27]

1.3 Treatment and resistance mechanisms

The discovery of Penicillin G, in 1929, and its introduction to clinical practice in 1940, marked a significant milestone in medicine.[28] Penicillin G, the first of the β-lactam antibiotics class, established and initiated an effective therapeutic approach for the treatment of life-threatening bacterial infections in healthcare facilities, thus saving countless lives.[4] However, it was soon discovered that an enzyme produced by *Escherichia coli*, called penicillinase, can eradicate the ability of penicillin to kill bacteria.[29] This was alarming and a first encounter of resistance against penicillin

and β-lactam antibiotics. As this was still before the extensive use of penicillin, it showed the natural existence of enzymes that can help bacteria resist the effects of β-lactam antibiotics.[29] Furthermore, it became clear that the widespread use of antibiotics provides selective pressure for the evolution of drug resistance in bacteria.[30] Consequently, this has led to the development and discovery of four groups of β-lactam antibiotics over the last few decades, namely, penicillins, cephalosporins, monobactams, and carbapenems.[31] These were discovered in attempts to overcome drug resistance, and scientists found ways to make modifications to existing β-lactams to form new generations of drugs with improved efficacy. Extended-spectrum β-lactam antibiotics such as third-generation cephalosporins and carbapenems have been considered as the last resort for the treatment of MDR gram-negative bacteria such as *K. pneumoniae*.[29,31]

The multidrug resistance of *K. pneumoniae* strains is a product of two main kinds of β-lactamases: the extended-spectrum β-lactamases (ESBLs) and *K. pneumonia* carbapenemases (KPC).[31] ESBL-producing *K. pneumoniae* strains were first reported in Europe and the United States in 1983 and 1989, respectively.[32] Since then, they have increasingly spread across the world, leaving the carbapenems as the last hope for treatment of ESBL-producing *K. pneumoniae.* ESBLs are capable of hydrolysing most of the β-lactam antibiotics, particularly, third generation cephalosporins and monobactams, while carbapenemases can hydrolyse almost all the β-lactams.[9] In 2013, the Centers for Disease Control and Prevention (CDC) reported 9 000 infections caused by carbapenem-resistant *Enterobacteriaceae,* of which approximately 80% were due to *K. pneumoniae*.[33] These reports about *K. pneumoniae* isolates expressing carbapenemases have raised even more concerns as these strains make the bacteria resistant to nearly all the currently available β-lactam antibiotics.[9]

1.4 Prevalence of *K. pneumoniae* disease in South Africa

K. pneumoniae diseases are of great concern in LMICs such as African countries. For example, *K. pneumoniae* has been a significant pathogen implicated in hospital-acquired infections in South African paediatric intensive care units.[34–37] Additionally, a variety of South African-based studies have reported several outbreaks of MDR *K. pneumoniae* infections. Moreover, *K. pneumoniae* has recently been identified to constitute 25% and 28% of ESKAPE and *E. coli* isolates obtained from patients with bacteraemia in the South African public and private sectors, respectively.[38]

Antibiotic susceptibility studies have revealed a continuous increase in resistant *K. pneumoniae* strains to multiple antibiotics, including a two-fold increase of isolates which are non-susceptible to carbapenems.[38] This is also evidenced by the 2021 South African surveillance of AMR and consumption of antimicrobials, which has found *K. pneumoniae* as one of the most common bacterial pathogens isolated from the blood of patients with bloodstream infections in both the public and

private sectors.[39] The prevalence of *K. pneumoniae* strains expressing extended-spectrum-β-lactamases (ESBLs) was found to have increased from 65% in 2016 to 70% in 2020, thus demonstrating the increasing limitations in the effectiveness of cephalosporins as a treatment option. Furthermore, an alarming increase from 8% in 2016 to 25% in 2020 of carbapenemase-producing *K. pneumoniae* strains was observed.

There is an urgent need for new effective antibiotics and an alternative approach to prevent infections caused by *K. pneumoniae*. Immunisation with a glycoconjugate vaccine is a promising alternative approach for the prevention of *K. pneumoniae* infections.[40] This idea has been known for many years and arose from the discovery of surface carbohydrates as major virulence factors of encapsulated pathogenic bacteria.[41]

1.5 Surface polysaccharides of *K. pneumoniae*

Bacterial surface polysaccharides are complex carbohydrate polymers produced by bacteria for expression on their external cell surface. Polysaccharides are polymers comprised of monosaccharide units arranged in repeating units (RUs); they can also contain non-carbohydrate substituents such as acetyls, phosphates, or pyruvate groups at specific positions in the RUs. These surface polysaccharides have been identified as important virulence factors for pathogenic bacteria for many years, this is generally because they are the first point of contact with the host's immune system and protect the bacteria from innate immune cells and environmental conditions such as osmotic pressure and desiccation.[42]

As a Gram-negative bacteria, *K. pneumoniae* possesses two types of surface polysaccharides, a capsular polysaccharide, also known as K-antigen, and a lipopolysaccharide (LPS), also known as O-antigen (**Figure 1.1**), in contrast to Gram-positive bacteria which lack LPS.[42] Additionally, these surface polysaccharides have been used as the basis for the differentiation and classification of bacterial strains into distinct K- and O-antigen serotypes.[43] Due to the wide variety of existing monosaccharide units, with at least 100 monosaccharides identified,[44] and the numerous ways in which they can combine, bacteria have used this to their advantage to produce a vast structural diversity of surface polysaccharides.

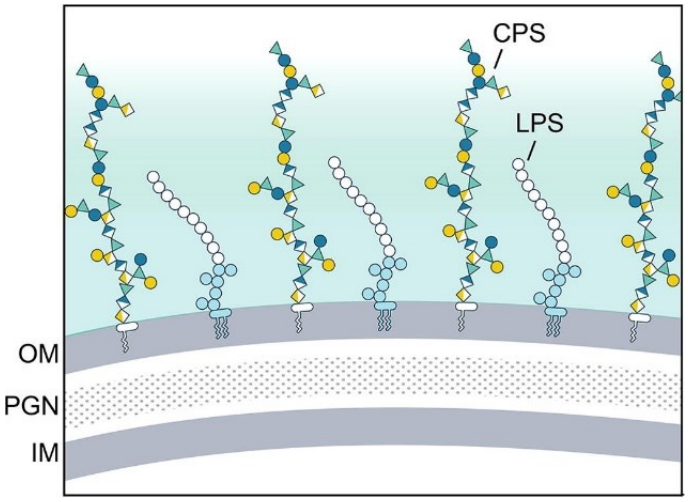


Figure 1.1: Schematic diagram of the extracellular structure (IM = inner membrane; PGN = peptidoglycan; OM = outer membrane) and surface polysaccharides (CPS = capsular polysaccharide; LPS = lipopolysaccharide) of *K. pneumoniae*. Modified from Opoku-Temeng et al. with permission.[4]

1.5.1 Capsular polysaccharides

Capsular polysaccharides (CPS) forms the outermost surface of *K. pneumoniae* and has been found as the main virulence factor as it allows the bacteria to evade innate immune complement activity and phagocytosis.[45] The CPS polymers are usually covalently linked to the cell surface of the bacteria and are capable of stimulating specific immune responses.[45] They are acidic and hydrophilic, thus, usually 95% hydrated, protecting the bacteria from desiccation, and shielding them from the environment.[12,46] Studies have shown that bacterial strains that are unencapsulated are less pathogenic and more susceptible to phagocytosis and serum killing compared to their encapsulated counterparts.[4,45]

To date, at least 77 serologically different CPS (K-antigen) serotypes have been identified for *K. pneumoniae*: K1 - K82 (excluding K73, K75, K76, K77, and K78, which were recognised to be serologically indistinguishable from previously described serotypes or to belong to another genus).[12,45,47] Moreover, additional K-antigens (known as the KL series) have been identified through a genotyping approach (described in **Section 1.6.2**).

Some *K. pneumoniae* strains, especially the hypervirulent *K. pneumoniae,* are heavily encapsulated with a hypermucoviscous CPS, forming a thick, highly protective and virulent polysaccharide coat (**Figure 1.2**).[45,48] The vast majority of these clinically isolated hypervirulent *K. pneumoniae* strains have been identified as belonging to two serotypes, K1 and K2, which account for about 70% of all hypervirulent *K. pneumoniae* infections worldwide.[49,50] However, it has been found that not all hypervirulent *K. pneumoniae* strains exhibit heavy hypermucoviscous CPS and that some classical *K.*

pneumoniae strains are also heavily encapsulated. This hypermucoviscous heavy CPS can be indicated by a positive string test, whereby an inoculation loop can successfully produce a viscous string longer than 5 mm when stretching the bacterial colonies grown on agar plates.[16]

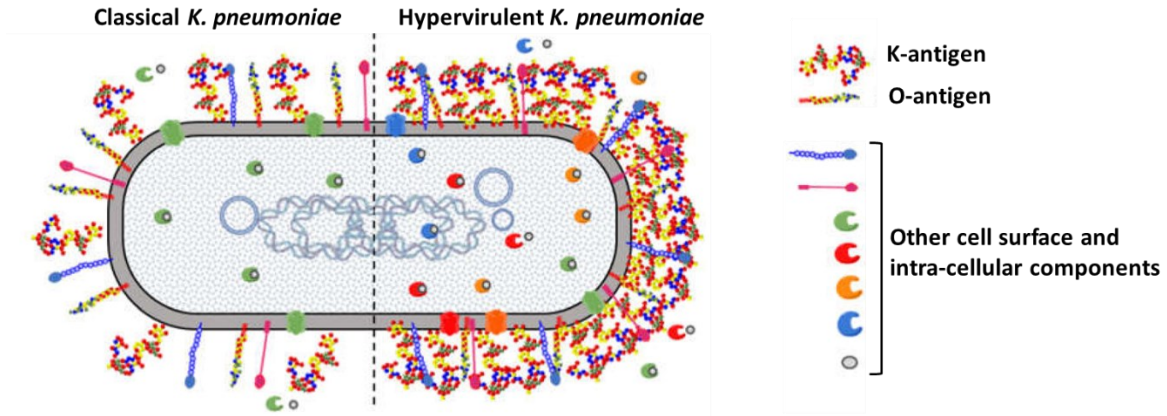


Figure 1.2: Schematic diagram representing the difference of the capsular polysaccharide (K-antigen) normally produced by classical *K. pneumoniae* compared to the hypermucoviscous and heavy capsular polysaccharide usually produced by hypervirulent *K. pneumoniae*. Modified from Arato et. Al. [51]

1.5.2 Lipopolysaccharides

The lipopolysaccharides (LPS) are made up of three components, Lipid A that attaches the whole LPS into the bacterial membrane, an oligosaccharide core in the middle, that is divided into an inner and outer core, and an outer polysaccharide chain to which the term, O-antigen, specifically refers (**Figure 1.3**).[12] The O-antigen is the most important component of the LPS because it is the part that interacts with the host's immune system. Moreover, the LPS is considered to play an important role in the pathology of septicaemia because of its endotoxic effects.[12] To date, there are at least 8 O-antigen serogroups of *K. pneumoniae* that have been identified with approximately 11 O-serotypes described.[52] However, historically, 12 O-serogroups were described, from which the serogroups: O6; O8; O9; and O11, have been removed as they have been found to be identical or closely related to previously identified serogroups.[3,45,53]

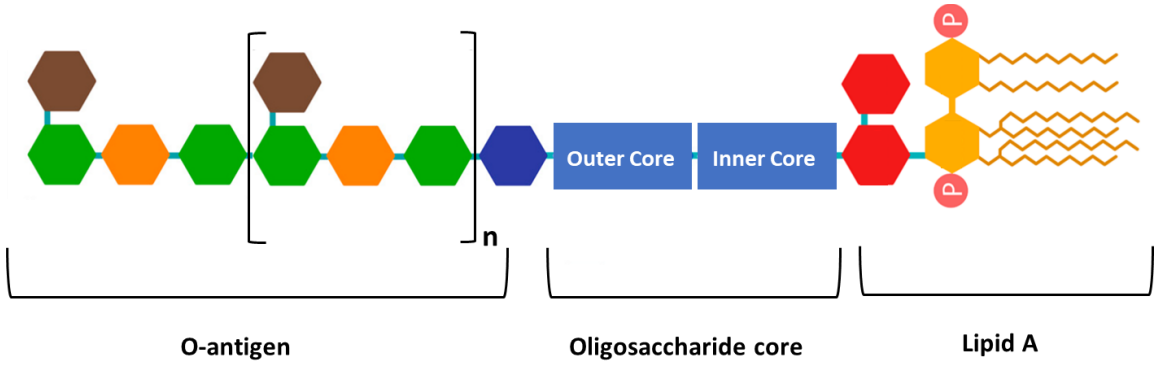


Figure 1.3: Schematic diagram of a lipopolysaccharide, showing its three structural components: Lipid A, oligosaccharide core, and the O-antigen. Modified from Mazgaeen L and Gurung P.[54]

1.6 Typing techniques of *K. pneumoniae* strains

1.6.1 Serotyping

Bacterial strains have been traditionally serologically classified into serotypes based on the interaction of surface polysaccharide antigens with specific typing sera.[12,55,56] This classification method is called serotyping, and since the CPS is commonly expressed by both Gram-negative and Gram-positive bacteria, it is the main surface antigen for serotyping. To achieve this, specific antisera must be produced using a suitable animal model.[57] The produced antisera is used in serological tests which are based on the recognition of the K-antigen with corresponding antibodies in a specific typing antiserum.[43,57,58] Similarly, in Gram-negative bacteria such as *K. pneumoniae* that produce LPSs, the O-antigens have been used for serotyping. However, since the CPSs (K-antigens) usually mask the O-antigens and are heat-stable, they have been described to make the serological classification of O-antigens difficult. In contrast, the CPS typing technique has been found to give good replicability and the ability to differentiate most clinical isolates.[12] However, serological cross-reactions among some K-antigens can make it challenging to differentiate between them.[56] This is usually because of structural similarities among some K-antigens. Moreover, the process can be cumbersome, time-consuming, and in certain cases, the serological tests are prone to subjective interpretation due to poor K-antigen and anti-serum reactions.[12] As a result, reports have suggested that a few strains, especially clinical isolates, are impossible to type using serotyping.[58–60] Additionally, typing anti-sera are not commercially available.[56]

1.6.2 Genotyping

Recent advancements in molecular techniques have given rise to genotyping methods, allowing for a complementary or alternative approach to bacterial serotyping. These are based on the characterisation and sequencing of specific genes conserved between bacteria of the same species, while varying between different K- or O-antigen serotypes.[61] Common examples reported for the classification of *K. pneumoniae* strains include the K-antigen typing based on the determination of the sequence of the *wzi* and/or *wzc* genes found within the capsular polysaccharide biosynthesis locus (cps locus or K-locus),[59,62] and the O-antigen typing based on the *wzm* or *wzt* genes involved in the transport system of the O-antigen during biosynthesis.[53,59,63]

The advanced molecular techniques used for genotyping include the polymerase chain reaction (PCR)-based method,[59,64] restriction fragment length polymorphism (RFLP),[65] pulsed-field gel electrophoresis (PFGE), multi-locus sequence typing (MLST), and whole genome sequencing (WGS).[61,66]

Genotyping has become an established, more practical and widely used approach for classification of *K. pneumoniae* strains, providing a higher differentiating power of isolates compared to the traditional serotyping technique.[6,7,59,62] As a result, up to 141 distinct K-loci have been identified from *K. pneumoniae* WGS, suggesting the presence of at least 141 K-antigen serotypes.[63,66] These 141 K-serotypes are referred to as the K-locus (KL) series: KL1–KL81, KL101–KL149, KL151, KL153–KL155, and KL157–165; and are inclusive of the 77 K-antigens that have been identified serologically, whereby KL1-KL81 are synonymous with K1-K81.[63]

The genotyping methods are also important for gaining insight into the epidemiology, AMR, and virulence factors of strains. For example, MLST has been widely used in epidemiological studies of *K. pneumoniae,* revealing the diversity of sequence types (STs),[67] such as the carbapenem-resistant *K. pneumoniae* strain identified as ST258, known as a high-risk clone which has spread globally.[68] Additionally, PCR methods and WGS have been utilised as means to determine drug-resistant genes and genetic traits of highly virulent strains, such as hypervirulent *K. pneumoniae*.[69]

Genotyping, however, has some limitations in that it focuses solely on the genotype (for example, K-loci), whereas it is important to note that the genotype may not always manifest fully in the phenotype (for example, the actual K-antigen), which may be due to mutations.[70] Further analysis can be done by in-depth investigation of the chemical structures of the K- and/or O-antigens, which provide absolute differentiation of closely related serotypes and specifically define each of them, using the method called chemotyping.

1.6.3 Chemotyping

Chemotyping refers to the classification and differentiation of bacterial strains through the study of their chemical composition, which can be based on their cell wall molecular components. This involves the direct analysis and characterisation of the chemical structures of the K-antigen or O-antigens using chemical and physiochemical analytical methods. Through a combination of these methods, the chemical composition and repeating unit (RU) structure of the polysaccharide antigen, known as the chemotype, is elucidated. The polysaccharide RU is defined by constituent monosaccharide residue types, stereochemistry, positions of linkages and the presence of non-carbohydrate substituents including O- or N-acetyls, pyruvate, or phosphate groups (**Figure 1.4**). The chemotypes of the traditionally serologically identified K- and O-antigens have been elucidated and reported in the literature. Chemotyping is of vital importance as the serotype-specificity of the immunological effects of K- and O-antigens is determined by their chemical structures. These can differ by even small structural features such as glycosidic linkage positions, or just the presence, position, and amount of non-carbohydrate substituents. Elucidation of the structure can provide an understanding of the observed serological cross-reactions among some antigens, which may be due to similar structural features and epitopes.

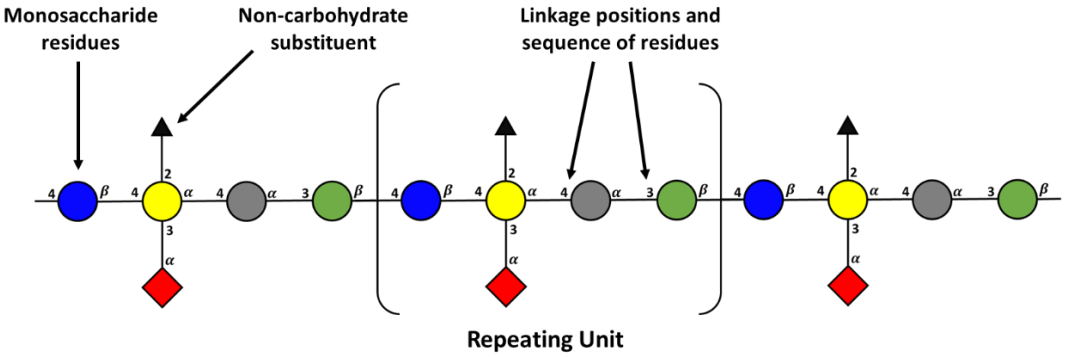


Figure 1.4: Representation of a typical bacterial polysaccharide repeating unit structural features. The blue, yellow, grey, and green circles, as well as the red square, represent distinct monosaccharide residues, with their α or β anomeric configurations and linkage positions indicated. The small black triangle represents a non-sugar substituent (for example, an O-acetyl group).

1.7 CPS repeating unit structures of known *K. pneumoniae* serotypes

The capsular polysaccharide structures of the 77 *K. pneumoniae* serotypes, originally identified serologically, have been elucidated and reported in several chemical studies in the literature, while this is unknown for the remaining genotypically identified K-antigens of the KL series.[66] To facilitate this investigation of characterisation and elucidation of the capsular polysaccharides of some new emerging *K. pneumoniae* strains, a database was constructed by compiling the CPS structural information of the known 77 *K. pneumoniae* serotypes (**Table 1A** in the **Appendix**). The database, as summarised according to the sugar compositions of the capsular polysaccharides in **Table 1.1**, shows that the capsular polysaccharides of *K. pneumoniae* generally consist of groups of neutral sugar and uronic acid residues that are linked to each other in a variety of patterns to form the array of K-antigens. The most common monosaccharide residues that form the CPSs are the neutral aldohexoses: glucose (Glc), galactose (Gal), and mannose (Man); the 6-deoxysugars: rhamnose (Rha) and fucose (Fuc); and uronic acid residues: glucuronic acid (GlcA), and galacturonic acid (GalA). Typically, all these monosaccharides have a D-configuration, except for the 6-deoxysugar residues, which are usually found in L-configuration. Unusual monosaccharides including modified uronic acids or keto acids (KetoA) are rare but have been found in the capsular polysaccharides of serotypes K22, K37, K38, and K66 (**Table 1A** in the **Appendix**). Furthermore, several K-antigens contain ketal-linked pyruvic acids and a few of them have O-acetyl groups, which may be important immunological epitopes, as indicated in **Table 1.1**. Due to the general presence of uronic acid residues, pyruvic acid (PyrA) substituents, or both, the capsular polysaccharides of *K. pneumoniae* serotypes are usually acidic in nature.

The three carbon pyruvic acid molecule usually forms the ketal-linkage through its C2, to the 4,6 hydroxyl groups of an aldohexopyranose residue, resulting in a six-membered ring, or across the 2,3 or the 3,4 hydroxyls, resulting in a 5-membered ring.[71,72] Additionally, this linkage can either form an R or S chiral centre at C2 of the pyruvate.

The capsular polysaccharides of distinct serotypes can differ based on one or a combination of the following structural features: the types or number of monosaccharide constituents, anomeric configurations, linkage positions, and the presence, positions, and extent of non-carbohydrate substituents (pyruvate and acetyl substituents). The sizes of the repeating unit structures can range from relatively simple trisaccharides to more complex heptasaccharides, as summarised by **Table 1.2**. Apart from the trisaccharide repeating units, most longer RUs often consist of terminal branches and residues from the backbone chain. This is indicated by the RU type column in **Table 1.2**, for example, a RU type of 3 + 1 denotes a tetrasaccharide repeating unit with a trisaccharide backbone and a single terminal branching residue.

Table 1.1: Qualitative analysis and grouping of known *Klebsiella pneumoniae* capsular polysaccharides based on their chemical monosaccharide compositions, extracted from the database in **Table 1A** in the **Appendix**. The superscript P indicates pyruvated and Ac indicates O-acetylated K-antigens, respectively.

Chemical composition	K-serotypes
GlcA; Gal; Glc	$8^{P,}$ 11^{P}, 15, 25, 27^{P}, 51, 82
GlcA; Gal; Man	20, 21^{P}, 29^{P}, 42^{P}, 43, 66, 74^{P}
GlcA; Gal; Rha	9,9*, 47, 52, 81, 83
GlcA; Glc; Man	2, 4, 5^{P}, 24, 39
GlcA; Glc; Rha	17, 23, 44, 45, 71
GlcA; Glc; Fuc	1, 54
GlcA; Gal; Glc; Man	7^{p}, 10, 13^{P}, 26p, 28, 30^{P}, 31^{P}, 33^{P}, 35^{P}, 39, 46^{P}, 50, 59, 60, 61, 62, 69^{P}
GlcA; Gal; Glc; Fuc	16, 58^{P}
GlcA; Gal; Glc; Rha	12^{P}, 18, 19, 36^{P}, 41, 55^{P}, 70^{P}, 79
GlcA; Gal; Man; Rha	40, 53, 80^{P}
GlcA; Glc; Man; Fuc	6^{P}
GlcA; Glc; Man; Rha	64^{P}, 65^{P}
GlcA; Gal; Glc; Man; Fuc	68^{P}
GlcA; Gal; Glc; Man; Rha	14^{P}, 67
GalA; Gal; Man	3^{P}, 49, 57
GalA; Glc; Rha	34, 48
GalA; Gal; Fuc	63
PyrA; Glc; Rha	72
PyrA; Gal; Rha	32
PyrA; Gal; Glc; Rha	56
KetoA; Gal; Glc	22, 37, 38

The full structural characterisation and elucidation of the repeating unit structures of these polysaccharide antigens are crucial to understanding their immunological characteristics and epidemiological distribution to facilitate the fight and prevention of infections. Importantly, as it was noted, some of these K-antigens can be serologically distinct due to even small structural differences, which is not always evident from serotyping and genotyping. As an example, K30 and K33 produce chemically identical capsular polysaccharides and only differ based on the amount of O-acetyl substituents present in their chemical structures, K30 is 33% O-acetylated while K33 is 100%.[73–75]

Table 1.2: Summary of the repeating unit size and repeating unit type (branching structure) of *K. pneumoniae* capsular polysaccharides, extracted from the database given in **Table 1A** in the **Appendix**.

Repeating unit size	Repeating unit type	K-serotypes
Trisaccharide	3	1^{P}, $5^{P, Ac}$, 63
Tetrasaccharide	4	4, 6, 32^{P}, 72^{P}
	3 + 1	2, 11^{P}, 16, 49^{Ac}, 54^{Ac}, 57, 58^{P}, 68^{P}, 82
	2 + 2	20, $22^{P, Ac}$, 23, 25, 37, 47, 51, 55^{Ac}
Pentasaccharide	5	44
	4 + 1	3^{P}, 8^{P}, 9, 17, 21^{P}, $21b^{P, Ac}$, 24, 35^{P}, 45, 48, 56^{P}, 59^{Ac}, 61, 62, 66
	3 + 2	13^{P}, 31^{P}, 43, 74^{P}, 80^{P}
	3 + 1 + 1	$30^{P, Ac}$, 33^{P}, 38, $69^{P,Ac}$
Hexasaccharide	6	39, 70^{P}, 81
	5 + 1	7^{P}, 10, 19, 34, 52, 53
	4 + 2	12^{P}, 14^{P}, 28, 36^{P}, 46^{P}, 64
	4 + 1 + 1	15, 27
	3 + 3	18
Heptasaccharide	5 + 2	40, 50, 79
	4 + 3	26^{P}, 67, 71
	4 + 1 + 1 + 1	60

1.8 Structural elucidation of polysaccharides

Classically, the analysis of polysaccharide structures relied on wet chemical methods and analytical techniques, involving acid hydrolysis of the polysaccharide into its constituent monosaccharides, followed by high-performance liquid chromatography (HPLC) or gas chromatography (GC) usually coupled with mass spectrometry (MS), or high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD).[44,76] These chemical steps and analytical techniques allow the determination of the monosaccharide chemical composition - the constituent monosaccharides present and their stoichiometric ratio. Thereafter, linkage analysis of the monosaccharides can be performed using methylation analysis, a process whereby the free hydroxyl groups in the polysaccharide are methylated, (using methyl iodide in basic conditions), followed by acid hydrolysis of the permethylated polysaccharide into partially methylated monosaccharides units.[44] Finally, the obtained partially methylated monosaccharides are reduced (using $NaBH_4$ or

$NaBD_4$), and the hydroxyl groups are acetylated with acetic anhydride in pyridine.[44] This leads to volatile, partially methylated alditol acetates (PMAAs) which are amenable for separation and identification with gas-liquid chromatography (GLC), coupled to MS for detection. The MS spectra of the partially methylated alditol acetates give characteristic peak patterns, which are indicative of the positions of the O-methyl and O-acetyl groups on each monosaccharide derivative, thus elucidating the linkage positions of the corresponding monosaccharide alditols.[44] However, uronic acids are not amenable to this analysis, due to the hydrolysis of their methyl ester groups. Furthermore, this chemical technique cannot provide the anomeric configurations of the monosaccharides and their sequence in the RU.[76] For the uronic acids, this can be overcome by the reduction of these residues into their corresponding aldoses, using deuterated reducing reagents ($NaBD_4$) to distinguish the derived hydroxyl group.[44,76] Although it is time-consuming and complex, a variety of degradation methods such as partial acid hydrolysis, acetolysis, periodate degradation/Smith degradation, or enzymatic degradation, can be used to overcome some of the described limitations of linkage analysis.[44,76]

Nuclear Magnetic Resonance (NMR) spectroscopy was initially only used as a tool for the confirmation of structural elucidation of oligosaccharides and some polysaccharides, to supplement chemical analysis.[77] Structural elucidation of polysaccharides has now become possible to achieve using NMR spectroscopy as the main technique. NMR spectroscopy can elucidate polysaccharide repeating unit structural features including the number and types of monosaccharides present, and their linkage positions, provides the sequence at which the residues are linked and whether they are α- or β-sugars, and also identifies the presence, positions, and extent of non-carbohydrate substituents.[78] NMR spectroscopy can elucidate all these structural features non-destructively, requiring a relatively small amount of sample (3 - 5 mg), and without too much manipulation or complex sample preparation.[79] This has been made possible by the significant advancement in NMR technology over the years.[77]

1.9 Fundamental history of NMR

Nuclear Magnetic Resonance is a spectroscopic analytical technique that measures the absorption of electromagnetic radiation when nuclear spins, such as proton (^1H) or carbon-13 (^{13}C), are exposed to a magnetic field.[80] The absorbed frequency results in a transition between two energy levels, characteristic of the electronic environment (local magnetic fields) around the nucleus, giving insight into the chemical structure, dynamics, and interactions of molecules.

In 1945, the phenomenon of Nuclear Magnetic Resonance was discovered independently by two groups of scientists led by Felix Bloch and Edward Purcell.[77] Years later, Proctor and Yu initiated the introduction of this phenomenon to the field of chemistry, when they discovered that the nitrogen

atoms of the NH_4NO_3 molecule gave rise to two different NMR chemical shift signals.[81] This led to the development of the early NMR spectrometers, which used continuous-wave radiofrequency, but were significantly insensitive and had limited resolution, thus resulting in challenges in the detection of ^{13}C NMR spectra due to the low natural abundance of ^{13}C.[77] In 1957, this technology was explored to investigate small biological molecules including common amino acids for the first time, however, it was found that when running the ^{1}H NMR experiment in water (H_2O), as a solvent, resulted in the lack of a visible spectrum. It was during this time that it was discovered that deuterium oxide (D_2O) could resolve this issue and yield a proton spectrum with well-resolved signals. A major challenge was the suppression of the significantly larger water residual signal compared to the signals of interest. In the early 1960s, superconducting magnets became available, and in 1966, Fourier-transformed NMR was introduced, providing a significant improvement in the sensitivity of this technique, thus allowing the detection of ^{13}C nuclei and formed the basis for the observation of other low natural abundance NMR active nuclei. This allowed the early NMR analysis of polysaccharides, which mainly relied on detecting ^{13}C as the proton NMR of polysaccharides gave broad lines which could not be analysed.

The next key advancement occurred in 1976 through the introduction of two-dimensional (2D) NMR experiments, which gave rise to experiments that can identify spin correlations through either scalar or dipolar couplings, allowing for the analysis of complex biomolecules including carbohydrates. Further technological advancements in superconducting materials, cryoprobes, and the increased power of computers continued to improve NMR as an analytical tool, especially, for biological molecules which can be highly complex.[77] Additional pulse sequences including non-uniform sampling (NUS), allowed a significant reduction of the acquisition time of 2D NMR experiments by acquiring data selectively, from few points of interest than from the entire experimental region, resulting in the possibility of suppression experiments,[82] particularly, diffusion-ordered spectroscopy (DOSY), which can remove signals of low molecular weight compounds such as the residual water signal and other solvents.[83] NMR has become the most powerful and important analytical technique in chemistry, especially for the structural studies of biological molecules, including polysaccharides.[79] This was the employed technique in this investigation and is discussed further in **Chapter 2**.

1.10 Carbohydrate-based vaccines

Vaccines can provide a cost-effective approach to the prevention of morbidity and mortality from infectious diseases caused by MDR pathogens. Glycoconjugate vaccines, where the polysaccharide antigens are covalently linked to a carrier protein, have proven to be efficacious in providing protective immunity against both pathogenic Gram-negative and Gram-positive bacteria.[84]

1.10.1 Overview of the concept and application of carbohydrate-based vaccine

The concept of carbohydrate-protein conjugate vaccines originated from the discovery of the immune-specific antibody-generating effects of bacterial capsular polysaccharides. This was initially used to develop purified capsular polysaccharide vaccines which were licensed during the 1970s to the early 1980s.[18,85] These were polysaccharide vaccines against pathogenic bacteria, for example, *Neisseria meningitidis*, *Streptococcus pneumoniae*, and *Haemophilus influenza type b*.[85] Due to the high diversity of CPSs produced by the different serotypes of these pathogens, their polysaccharide vaccines had to be multivalent, meaning that the CPS antigens of multiple clinically significant serotypes of a pathogen are added to its vaccine formulation to provide an effective vaccine with sufficient coverage and protection. However, the purified polysaccharide vaccines were found to have limited efficacy for the most vulnerable and immunosuppressed groups; children under 5 years of age and elderly people over 75 years of age,[86] due to the T-cell-independent immune activity of polysaccharide antigens, unlike T-cell antigens such as proteins, resulting in the lack of long-term immune memory.[85] These limitations led to the development of carbohydrate-protein (glycoconjugate) vaccines where the polysaccharide antigen is linked to a T-cell immune-active carrier protein.[87] Indeed, glycoconjugate vaccines have proved to be more effective than pure polysaccharide vaccines in providing protective immunisation against bacterial infections and can stimulate long-term immune memory and protection. The first successful glycoconjugate vaccine to be licensed was for *Haemophilus influenzae* type b (Hib) in 1987,[87] and subsequently, multivalent glycoconjugate vaccines have been licensed against other bacteria such as *Streptococcus pneumoniae* (up to 20 valency),[88] *Neisseria meningitis* (up to five valency),[89,90] with higher valency vaccines in development.[85,91]

1.10.2 *K. pneumoniae* potential vaccine development

The *K. pneumoniae* capsular polysaccharides have also been explored and investigated for the development of a multivalent vaccine.[92–94] These investigations also demonstrated that the CPSs are immunogenic potential vaccine antigens, and suggested that a vaccine formulation of about 25 of the most commonly isolated K-antigen serotypes would provide significantly broad coverage against clinically important strains.[95] Based on these findings, a 24-valent pure polysaccharide-based

vaccine was developed, which has been the most promising *K. pneumoniae* vaccine candidate to date and was proven to be safe and immunogenic in both children and adults through human clinical trials.[5,96,97] However, as has previously been observed in preliminary studies on purified polysaccharide vaccines,[98] this lacked long-lasting immune memory and never progressed to reach licensure. This was also partly due to the complexity of its production and to the highly variable diversity of *K. pneumoniae* capsular serotypes, making it considerably expensive to develop.[94]

O-antigens are generally masked by the CPS and thus not fully exposed for antibody binding in Gram-negative bacteria such as *K. pneumoniae*.[12,99] This is especially the case for hypervirulent *K. pneumoniae* strains as they produce a hypermucoviscous CPS which significantly masks the O-antigens.[99] As a result, there are currently no licensed O-antigen-based vaccines that have progressed beyond clinical trial investigations for *K. pneumoniae* (and other Gram-negative bacteria). However, since O-antigens have also proven to be safe and immunogenic in humans, they have been applied to develop candidate conjugate vaccines against Gram-negative bacteria such as *Shigella flexneri*, and extraintestinal pathogenic *Escherichia coli* (ExPEC).[100,101] Unlike other Gram-negative bacteria, *Shigella flexneri* is unencapsulated, therefore the LPS O-antigen component is an important antigen for the development of its conjugate vaccine.[101] For ExPEC, one of the candidate O-antigen conjugate vaccines referred to as ExPEC10V is now in phase 3 clinical trials.[102]

Since the O-antigens have low variability across *K. pneumoniae* serotypes, studies have shown that the inclusion of a minimum of four O-antigens (O1, O2, O3, and O5), from the eight O-antigens, to its multivalent vaccine formulation should provide broad coverage of more than 80% of global clinical isolates.[3,43,103] This would provide an advantage and convenience compared to the utilisation of the highly diverse capsular polysaccharide antigens. Indeed, a quadrivalent O-antigen conjugate vaccine known as Kleb4V has recently been investigated in phase I/II human clinical trials.[104] Recently, Wantuch et. al. reported a development and pre-clinical evaluation of a heptavalent conjugate vaccine, which is technically tetravalent but inclusive of two O1, O2, and O3 antigen subtypes (O1v1, O1v2, O2v1, O2v2, O3, O3b), and O5 antigen.[103,105] It was found that the antibody-binding and function of this vaccine was affected by the amount of capsular polysaccharide produced by the strains, with hyper-encapsulated strains exhibiting low antibody-binding compared to low capsular-producing strains.[104] This is also consistent with a similar study in which the same group investigated and compared the activity of monovalent conjugate vaccines against the hypervirulent K2 and the O1 *K. pneumoniae* serotypes.[104] It is believed that this is due to blockage of O-antigen antibodies by capsular polysaccharides of both hypervirulent and a variety of some classical *K. pneumoniae* strains, as it has been demonstrated by several other studies,[103,106,107] suggesting that candidate *K. pneumoniae* O-antigen vaccines may not provide protection to all clinical

strains. Therefore, to date, capsular polysaccharides represent the only proven effective and protective *K. pneumoniae* vaccine antigens despite the complexity of the development of such a multivalent conjugate vaccine.

In the development of these vaccines, the target CPS antigens must be obtained by bacterial fermentation of the desired strains, followed by separation and several purification steps of the polysaccharides, which are variable and dependent on factors such as the length, charge, and non-carbohydrate substitution patterns of the target polysaccharide antigens.[40] As previously noted, chemotyping of the targeted CPS antigens is one of the crucial initial steps in the process of vaccine manufacturing for identification and verification of serotype identity and purity testing of the purified polysaccharides.[84] Furthermore, during later production steps such as chemical conjugation of the polysaccharide antigens to a carrier protein, it is important to ensure that the structural integrity of the antigens is still maintained. This can be provided by NMR characterisation of these polysaccharide antigens.[108]

1.11 Scope of the research

This investigation was part of a global initiative which involved isolating, genotyping, and identification of new emerging *K. pneumoniae* strains that are clinically significant. Isolation and genotyping were performed by collaborators, and the chemotyping of the K-antigens of some of the identified clinically significant strains using NMR spectroscopy is presented here.

Aims

This study aimed to conduct an in-depth NMR analysis of the capsular polysaccharides of four emerging, clinically important strains of *Klebsiella pneumoniae*, that were identified as new serotypes based on genotyping by collaborators, namely, K102; K112; K122; and KL107, to characterise and elucidate their repeating unit structures (chemotypes) and to provide complete ^{1}H and ^{13}C NMR data, together with 1D ^{1}H, ^{13}C, and 2D ^{1}H-^{13}C diagnostic reference spectra. The diagnostic spectra and NMR data form part of a database for serotype identity, purity, and structural integrity testing of K-antigens to facilitate the development of *K. pneumoniae* capsular polysaccharide-based vaccine.

Objectives

1. Construct a database of known *K. pneumoniae* CPS structures by collecting and compiling data and proposed repeating unit structures reported in the literature to assist with the investigation.
2. For the CPSs of each of the four emerging *K. pneumoniae* strains, record the required 1D and 2D NMR experiments to acquire NMR data (spectra). The experiments include ^{1}H, ^{13}C, ^{1}H-^{1}H scalar and dipolar homonuclear spectra, and 2D ^{1}H-^{13}C heteronuclear spectra at 600 MHz and at a suitable temperature for each CPS with additional sample preparation as required such as sonication of some CPS samples.
3. Process the acquired data for the first CPS and establish a methodology to analyse and assign all protons and carbons to achieve full NMR characterisation and elucidation of the CPS repeating unit and apply this methodology to the remaining three CPSs.
4. Compare the NMR structural analysis results to those obtained using chemical analysis conducted by collaborators at the University of Trieste, Italy.
5. Compare the composition and proposed RU structure of each CPS to other *K. pneumoniae* CPSs using the compiled database to confirm whether each CPS is novel or not, and identify if there are significant structural similarities which can potentially provide serological cross-reactivities.
6. Present the established NMR data, including the 1D and 2D spectra to form part of the database for analysis of future batches of CPSs and potential vaccine development.

Chapter 2: Methods for NMR characterisation of polysaccharides

The structure of a polysaccharide repeating unit is determined by the types and number of monosaccharide residues, their linkage pattern and positions, whether they have α- or β-configurations, the presence or absence of non-sugar substituents, and their positions and amount if present (**Figure 1.4**). All these structural features can be identified using NMR analysis to allow the characterisation of polysaccharides. For *K. pneumoniae* capsular polysaccharides, the expected common monosaccharide residues are the hexose sugars: glucose, galactose, and mannose; the uronic acids: glucuronic acid and galacturonic acid; and the 6-deoxy sugars: rhamnose and fucose. All of these, apart from the 6-deoxy sugars, are usually found in the D-absolute configuration. The absolute configuration cannot be readily determined by NMR, but for some special cases, it can be confirmed based on the magnitude of the carbon glycosylation shifts of any disaccharide fragments in the RU and available configurational factors in the literature. Additionally, the CPSs of *K. pneumoniae* serotypes often contain pyruvate and acetyl group substituents which can be detected by NMR.

In this chapter, the interpretation and main parameters of the key NMR experiments commonly performed for the analysis of polysaccharides as applied in this investigation will be described.

2.1 NMR experiments and interpretation

2.1.1 1D ^{1}H NMR experiments

One-dimensional (1D) proton (^{1}H) and carbon (^{13}C) NMR experiments are the starting points in the structural characterisation of polysaccharides. Their 1D spectra can be divided into key distinctive regions, some of which can be quickly analysed for several structural features of the polysaccharide repeating unit.

The polysaccharide 1D ^{1}H NMR spectrum possesses three main distinctive regions: the anomeric region lies downfield between 4.4 - 5.6 ppm; followed by a ring region from 3.5 - 4.5 ppm; and up-field lies the methyl region between 1.1 - 2.8 ppm (**Figure 2.1**).[71,72,108,109] Generally, the proton NMR spectra of polysaccharides are complex with heavily crowded signals, particularly in the ring region, even for simple oligosaccharide sugars or polysaccharides composed of a disaccharide repeating unit (RU). This ring proton region cannot be immediately analysed from only the 1D proton spectrum and requires more informative correlation experiments for analysis; however, it is considered the “fingerprint region” for specific polysaccharides.[71] Therefore, the structural

information obtained from the analysis of 1D proton is derived from the most deshielded anomeric region and the most shielded methyl region.

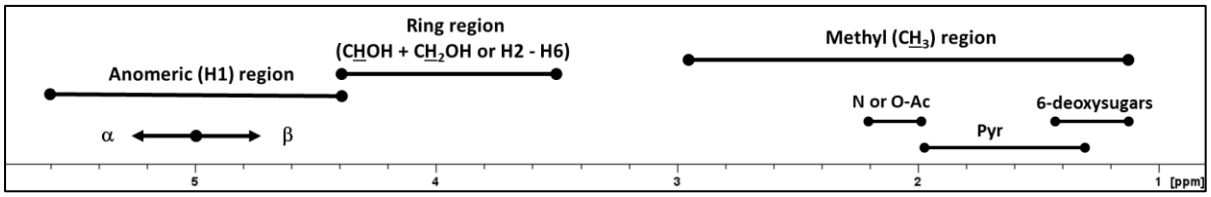


Figure 2.1: Summary of the key diagnostic regions of polysaccharide 1D ^{1}H NMR spectrum and some of the protons normally expected to resonate at each region.

The anomeric region is considered the most informative and diagnostic part of the proton spectrum as it provides well-resolved signals from the anomeric protons (H1) of aldose monosaccharide residues.[71,109] As a result, the number of signals in the anomeric region can reveal the number of constituent monosaccharide residues in the polysaccharide RU. However, some ring protons, such as H4 of Gal, H2 of ManNAc, H5 of α-GalA, and methylene protons (H6/6') of sugars with O-acetylation at carbon 6, may resonate within the anomeric region.[108] This may cause some ambiguity in determining the number of constituent sugar residues in the RU from 1D proton analysis alone. Furthermore, the chemical shifts of the anomeric protons can be used to distinguish α-pyranose sugars, which exhibit downfield anomeric proton signals (5.0 - 5.6 ppm), and β-pyranose sugars, which exhibit up-field anomeric proton signals (4.4 - 5.0 ppm) (**Figure 2.1**).[71,110] Moreover, the magnitudes of the observed $^{3}J_{H1,H2}$ coupling constants of pyranose sugars can be used to deduce their anomeric configurations. For example, a small $^{3}J_{H1,\,H2} \sim 3.7$ Hz is observed for α-pyranose sugars with axial H2, due to the equatorial-axial arrangement of H1 and H2, while for β-pyranose sugars with axial H2, a large $^{3}J_{H1,\,H2} \sim 7.8$ - 8.5 Hz is observed due to the trans-diaxial arrangement of H1 and H2.[111]

On the other end, the methyl region of the spectrum is considered the second-most informative part because it only exhibits signals of methyl protons (C<u>H</u>$_3$), which can indicate the presence of 6-deoxy sugars such as rhamnose or fucose (1.1 - 1.4 ppm); as well as non-carbohydrate substituents such as N- and O-acetyl groups (2.0 - 2.2 ppm) or pyruvate groups (~1.5 ppm or 1.3 - 2 ppm) (**Figure 2.1**).[71,72,110]

2.1.2 Water suppression experiments (Pre-sat and DOSY)

Deuterated water (D_2O) is the solvent commonly used in the preparation of polysaccharide samples for NMR analysis. Although this has the advantage of readily exchanging with the hydroxyl (O<u>H</u>) protons, it has a major disadvantage in proton-detected NMR experiments as it gives a significantly large hydrogen-deuterium oxide (HDO) residual peak. The chemical shift of this peak is highly temperature dependent and can be approximated by the equation, $\delta = 5.051 - 0.0111T$, where T is

the temperature in degrees celsius (°C).[112] This means that at temperatures between 70 °C (343 K) and 60 °C (333 K), the HDO signal resonates between 4.27 and 4.39 ppm which overlaps with the previously mentioned proton anomeric region (4.4 - 5.6 ppm), and, as a result this large residual signal can obscure some of the diagnostic proton anomeric signals. Since the HDO signal is temperature dependent, in some cases it can be simply overcome by varying and optimising the temperature at which the proton spectrum is recorded, to reveal the signals of interest.

Several solvent suppression experiments are available as an alternative to temperature variation, firstly, the most common method is pre-saturation (Pre-sat), which allows selective saturation of unwanted resonances before acquiring the spectrum.[113] Pre-sat is relatively simple to record and helps to improve the resolution of the spectra, however, it removes all signals at the same frequency as that of the solvent and can lead to distortion of nearby signals (polysaccharide proton anomeric signals), leading to a loss of information.[114] The second alternative method is diffusion-ordered spectroscopy (DOSY). This method takes advantage of the high diffusion rates of small molecules, including water and other solvents, compared to the low diffusivity of large solute molecules, including polysaccharides, which is related to the molecule's hydrodynamic radius.[83,115] Consequently, this is well-suited for residual water suppression when recording solution-proton NMR for analysis of polysaccharides. In this method, magnetic field gradient pulses are applied, which allow the suppression of all signals from small molecules that have high diffusion rates, while the signals of large molecules are unaffected.[83] Therefore, unlike pre-saturation, DOSY provides the desired result of selectively removing the HDO residual signal without affecting the nearby or obscured anomeric signals.

2.1.3 1D ^{13}C NMR experiments

The analysis of the diagnostic regions of the proton spectrum can be corroborated and complemented by the carbon NMR spectrum, which can also provide additional information. Compared to the 1D proton experiment, the 1D proton-decoupled ^{13}C NMR spectrum gives sharp and well-separated signals and is more diagnostic since the carbon signals resonate across a significantly larger range of 0 to 200 ppm (**Figure 2.2**).[116] However, a disadvantage is its significantly lower sensitivity (about 10,000 times lower) than the ^{1}H NMR, which requires longer experiment times and a larger sample amount to obtain good-quality spectra.[108]

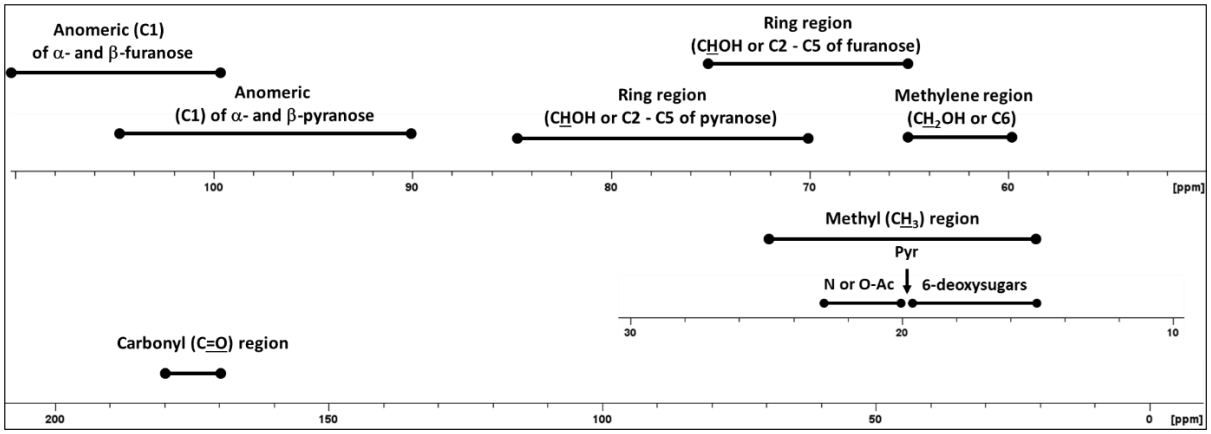


Figure 2.2: Summary of the key diagnostic regions of polysaccharide 1D ^{13}C NMR spectrum with some of the expected types of carbons to resonate in each region indicated.

In 1D ^{13}C NMR spectrum, the anomeric region, between 91 - 105 ppm, gives a more accurate indication of the number of constituent sugar residues in the RU as it only contains signals from C1 of aldopyranose sugars (and C2 of ketose sugars).[71,72,108–110] Moreover, the C2 of acetyl-linked pyruvic acid substituents, which forms the linkage to a monosaccharide, gives a signal within this region, resonating at ~100 ppm if the pyruvate is 4,6-linked to the monosaccharide residue, therefore a six-membered ring, and at ~110 ppm if the linkage is through the 2,3- or 3,4-positions, thus forming five-membered ring with the monosaccharide.[71,72] Additionally, the ^{13}C spectrum can discern pyranose sugars from furanose sugars as C1 of α- and β-furanose residues have more down-field signals (101 - 111 ppm) compared to C1 of their pyranose counterparts, which resonate between 91 - 101 ppm for α-pyranose residues, and at 95 - 105 ppm for β-pyranose residues (and C2 of keto sugars).[108,111]

The carbon ring region lies between 65 - 75 ppm for pyranose sugars and 70 - 85 ppm for furanose sugars.[111] Although the carbon ring signals are significantly well resolved compared to proton ring signals, they are still relatively crowded. Contrary to the proton spectrum, the carbon ring region only exhibits signals of secondary carbinolic carbons (C2 to C5) while the methylene group carbons (C6) resonate relatively up-field in a distinctive region between 60 – 65 ppm, and N-acetyl linked carbons are shifted up-field between 48 - 55 ppm.[72,108,110,117] The methyl carbons resonate further up-field (15 - 25 ppm) with diagnostic signals of methyl carbons ($\underline{C}H_3$) from acetyl substituents (20 - 23 ppm), pyruvate groups (~23 ppm), and 6-deoxysugars (15 - 19 ppm). Lastly, the carbonyl (C=O) signals from C6 of uronic acid residues, acetyl groups, and pyruvate groups give deshielded signals at 170 - 180 ppm.

Notably, the glycosylation (linkage) of monosaccharide residues has prominent effects on the chemical shifts of the linkage and nearby carbons, resulting in a significant deshielding effect (α effect) at the linkage carbon, which causes a large down-field shift of its resonance (5 - 10 ppm), while causing a shielding effect (β effect) of the neighbouring carbons, resulting in a small up-field shift of their resonances (1 - 2 ppm).[117] These are called glycosylation shifts and can be calculated by subtracting the carbon chemical shifts of the unsubstituted monosaccharides, which can be obtained from computer-assisted spectrum evaluation of regular polysaccharides (CASPER),[118] from the experimental chemical shifts of their respective constituent monosaccharide residues in the polysaccharide. This provides a simple method for identification and confirmation of linkage positions. Moreover, this means that the methylene carbon signals of C6-linked residues can be shifted downfield (65 - 68 ppm) into the ring carbon region.[108] The magnitudes of these glycosylation shifts are dependent on the types of monosaccharide units involved, their absolute and anomeric configurations, and the positions of the glycosidic linkages, these are known as stereochemical or configurational factors. In some cases, the application of these factors and rules, reported by Shashkov et al., not only predict the sizes of the glycosylation shifts but can be used to determine and confirm the absolute configurations of monosaccharide units, with the assumption the absolute configuration of one of the units is known.[119,120]

The O-acetyl substituent affects the carbon signals of the linked monosaccharide, resulting in a 0.6 - 3.5 ppm α-acylation effect on the geminal carbon, while causing β-acylation effects between 1.2 - 1.4 ppm on the vicinal carbons.[71] Similarly, the monosaccharide positions that are acetyl-linked by the pyruvic acid can be determined from the analysis of the carbon chemical shifts of the monosaccharide, as the carbons at the pyruvate-linked positions experience α-effects, whereas those adjacent experience β-effects.[71]

To aid the analysis of the carbon signals, the Distortionless Enhancement by Polarization Transfer experiment, known as DEPT-135, can be recorded. This inverts methylene carbon signals and eliminates quaternary carbons (carbonyls) and C2 of ketose signals, while the methyl and methine signals remain unchanged.[121] The DEPT-135 spectrum can thus assist with the assignment or identification of methylene carbon signals, especially if they are linkage positions and deshielded into the ring carbon region.

2.1.4 ^{1}H-^{1}H scalar homonuclear experiments (COSY and TOCSY)

Two-dimensional (2D) proton-proton (^{1}H-^{1}H) homonuclear experiments such as correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY) allow the deciphering of the ring proton signals of each sugar residue using the anomeric signals or other diagnostic signals such as methyl protons of 6-deoxy sugars as a starting point.[71,72]

A 2D COSY experiment establishes scalar (through-bond) correlations of protons on neighbouring carbons (vicinal coupling) and on the same carbon (germinal coupling, such as CH_2 protons) for each spin system, which means that starting from any diagnostic signals, for example, the anomeric protons, correlations such as H1 to H2; H2 to H3; and further if there is no overlap, can be observed (**Figure 2.3a**).[78] The complete elucidation of the ring protons from this process is, however, most often prevented by the extensive overlap of the ring signals. This can be further resolved by the 2D TOCSY experiment, which can give well-resolved, multiple-bond scalar correlations for each spin system through the transfer of magnetisation between its coupled protons.[122] The extent of the transfer of magnetisation is dependent on the magnitudes of the coupling constants between adjacent protons of each spin system ($^{3}J_{Ha,\ Hb}$), which are determined by the relative configurations of the ring protons, and mixing time. [122] This enables the TOCSY experiment to identify the types of aldopyranose residues and their ring protons by tracing correlations from each anomeric signal.

For glucopyranose sugars, TOCSY can give correlations from H1 to H2 and additional correlations from H1 to H3, H4, H5, and H6 since the $^{3}J_{Ha,\ Hb}$ coupling constants between the consecutive protons are large enough to allow the complete transfer of magnetisation across the spin system (**Figure 2.3b(i)**). This is because of the axial orientation of the ring protons in glucopyranose sugars. Similarly, for glucuronic acid, TOCSY can elucidate the full spin system, H1 to H5 (**Figure 2.3b(ii)**). However, for galactopyranose sugars, TOCSY gives no further correlations after H1 to H4 due to the small $^{3}J_{H4,\ H5}$ coupling constant because of the equatorial orientation of H4 (**Figure 2.3b(iii)**), and similarly applies for galacturonic acid. For mannopyranose sugars, TOCSY can give a correlation from H1 to H2, and sometimes H1 to H3 (with low intensity), due to a small $^{3}J_{H2,\ H3}$ coupling constant because of the equatorial orientation of H2 (**Figure 2.3b(iv)**). However, from H2 of a mannopyranose sugar, the complete spin system can be elucidated by TOCSY as the $^{3}J_{Ha,\ Hb}$ coupling constants for the remaining protons are large enough to allow the complete transfer of magnetisation till the last proton, and the H2 signal usually occurs downfield compared to other ring protons, thus making it diagnostic. Therefore, the use of TOCSY not only allows the deconvolution of the ring proton signals but can also provide information about the identity of aldopyranose residues, as summarised in **Figure 2.3**.

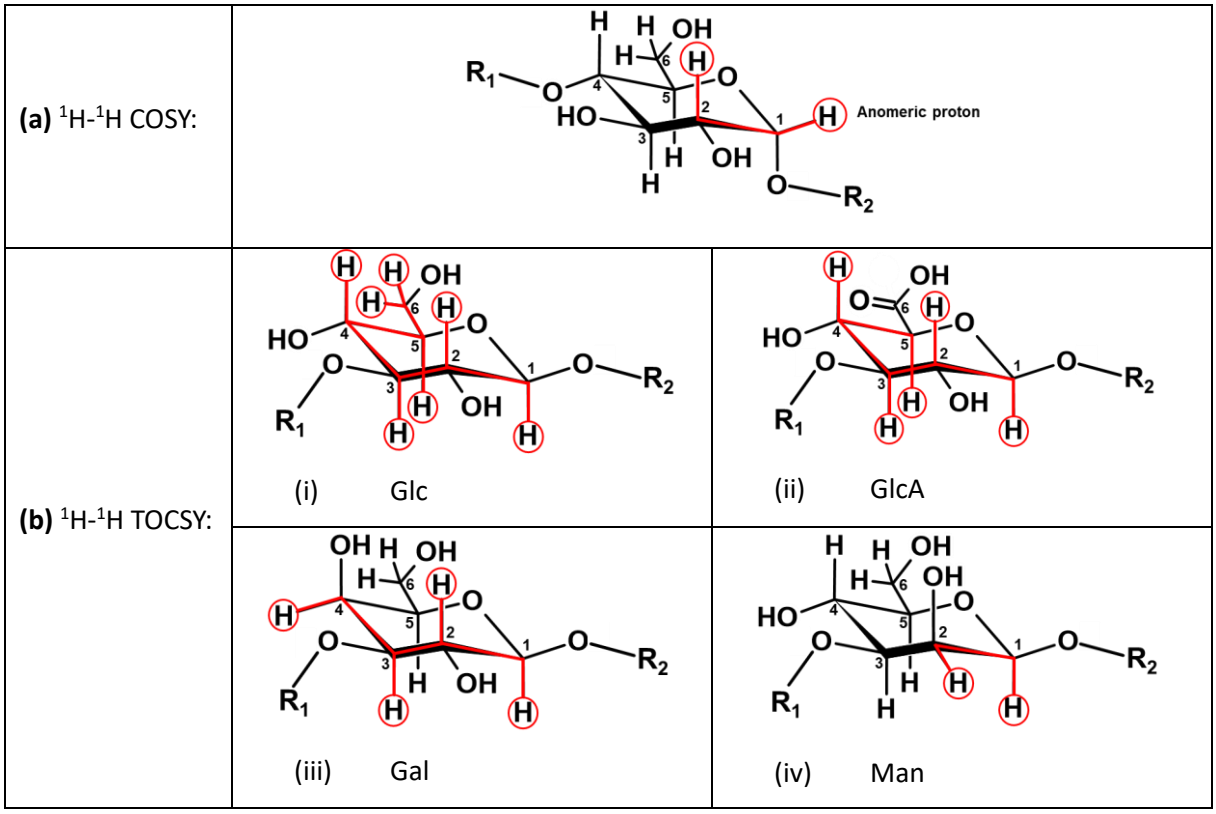


Figure 2.3: Summary of ^{1}H-^{1}H homonuclear scalar correlation NMR experiments, showing the assignments that can be observed from (a) COSY and (b) TOCSY dependent on the monosaccharide residue type (Glc = glucopyranose; Gal = galactopyranose; GlcA = glucuronic acid; Man = mannopyranose). R_1 and R_2 represent adjacent residues if the monosaccharide in question is 1,4- or 1,3-linked.

Information about the identity of the proton signals is not immediately apparent in the TOCSY experiment, thus 2D COSY and TOCSY are usually analysed together by overlaying the COSY spectrum onto the TOCSY spectrum. Since COSY only shows correlation cross-peaks of vicinal and geminal protons, while TOCSY shows the same cross-peaks and additional long-range correlation cross-peaks, the COSY cross-peaks appear on top of the identical cross-peaks on TOCSY, and the remaining TOCSY cross-peaks can reveal the additional proton signals of each spin system.

Additionally, selective 1D TOCSY experiments can be recorded by irradiating any diagnostic signal observed in the 1D proton spectrum, to provide a series of 1D projections of the 2D TOCSY correlations for each spin system.[123,124] These 1D variants are advantageous as they provide TOCSY spectra of higher resolution, reveal the characteristic shapes of the peaks and their multiplicity, permit the use of even longer mixing times thus allowing the magnetisation to flow to further protons, and can help with resolving ambiguities arising from the overlapping ring proton signals.[108,124] Furthermore, they can be used to identify the order of proton signals in each spin system by irradiating a single diagnostic proton with a step-wise increase of mixing times. This provides a series of 1D TOCSY spectra

which shows the order in which the magnetisation and coupling moves across the spin system as the mixing time increases.

As in the 1D proton NMR experiment, the HDO residual signal in the 2D TOCSY experiment can be suppressed using the previously described methods; pre-sat and DOSY. 2D DOSY-TOCSY was recorded for this investigation, which selectively suppresses the signals arising from low molecular weight compounds in the TOCSY spectrum without affecting the polysaccharide signals, like a 1D DOSY.

2.1.5 ^{1}H-^{1}H dipolar homonuclear experiment (NOESY)

The limitations of small scalar coupling constants encountered in sugars like galactopyranose can be overcome using 2D ^{1}H-^{1}H Nuclear Overhauser Effect Spectroscopy (NOESY) experiment, which gives dipolar (through-space) correlations dependent on the configuration and proximity of protons in space.[125,126] This means that the dipolar correlations can be from protons of the same sugar spin system (intra-residue correlations) and protons of neighbouring spin systems (inter-residue correlations). For α-pyranose sugars, only intra-residue correlations from H1 to H2 are expected, while additional correlations from H1 to H3, and H1 to H5 can be detected for β-pyranose sugars. Additionally, inter-residue correlations from H1 of each sugar residue to protons of the sugar it links to (and conversely), can be observed, which enables the determination of the sequence of residues in the RU, and in some cases the identification of the linkage positions.[111] However, there can be ambiguity in identifying the linkage positions if inter-residue correlations from H1 to several protons of the neighbouring residue appear. The information that can be acquired from the NOESY experiment is summarised using an example of a disaccharide unit in **Figure 2.4**. Overall, the NOESY interactions can give rise to even more overlapping cross-peaks and is best interpreted as an overlay with the TOCSY/COSY spectra. This not only aids the assignments of nearly or all the remaining protons that cannot be elucidated from the COSY and TOCSY experiments, but also assists with the identification of the sequence of the residues in the RU. Similarly, selective 1D NOESY experiments can also be recorded by irradiating the anomeric protons to obtain higher-resolution 1D spectra for each spin system.[127] The 1D spectra can also be overlaid with the key 1D TOCSY and 1D DOSY spectra to cross-validate the assignments.

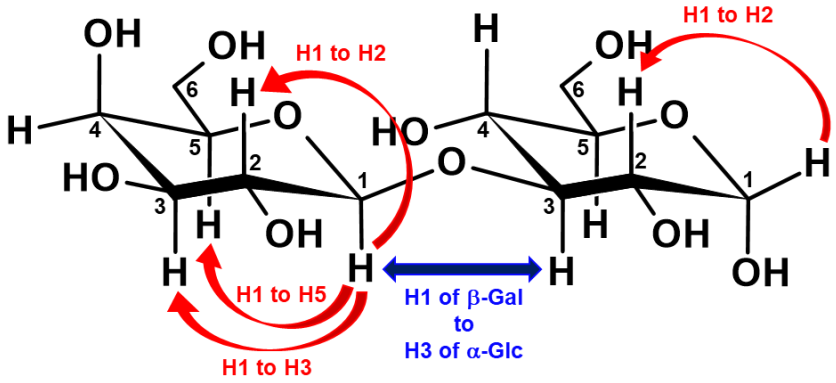


Figure 2.4: Example of a disaccharide unit, β-galactopyranose-(1→3)-α-glucose (β-Gal-(1→3)-α-Glc), showing intra-residue (in red) and inter-residue (in blue) correlations assignments that can be observed from ^{1}H-^{1}H NOESY experiments.

2.1.6 2D ^{1}H-^{13}C heteronuclear experiments (HSQC-DEPT, HSQC-TOCSY, HSQC-NOESY, and HMBC)

Once all or nearly all of the proton assignments have been made from the homonuclear experiments, they can be used in assigning the carbon signals of the spin systems using the heteronuclear single quantum coherence (HSQC) experiment.[128] HSQC provides scalar-correlations of protons to the carbons they are chemically bonded to (i.e., $^{1}J_{Ha,\ Ca}$), shown in red in **Figure 2.5**. The anomeric and methyl regions give well-resolved ^{1}H-^{13}C correlation cross-peaks which can usually be easily assigned. As with the 1D ^{1}H and ^{13}C spectra, the anomeric cross-peaks are highly diagnostic for the number of aldose sugar residues in the RU. However, yet again, the significant overlap of ring proton signals can present challenges in assigning and discerning their corresponding correlation cross-peaks. The 2D HSQC-DEPT experiment, which distinguishes ^{1}H-^{13}C correlation cross-peaks of methylene groups as inverted peaks, helps with the assignment of the H6-C6 signals which lie within the ring region.[108,129] The inverted cross-peaks can be displayed in a different colour from other cross-peaks for easy identification.

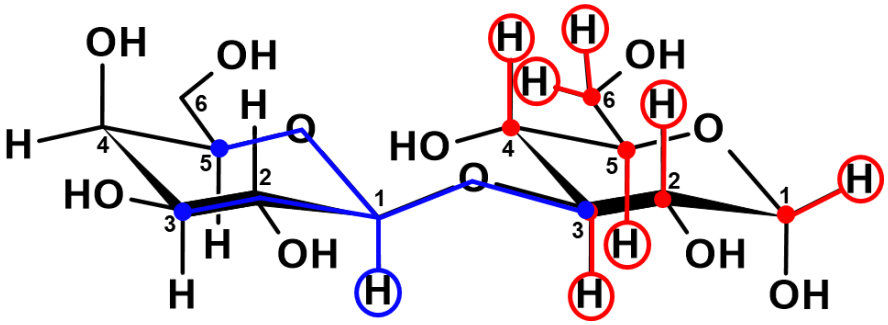


Figure 2.5: A β-galactopyranose-(1→3)-α-glucose (β-Gal-(1→3)-α-Glc) disaccharide unit, showing the key assignments obtained from the standard ^{1}H-^{13}C heteronuclear experiments, HSQC correlations in red and HMBC intra- and inter-residue correlations in blue.

2D HSQC-hybrid experiments; HSQC-TOCSY and HSQC-NOESY experiments, can be recorded to aid the assignment and resolution of ambiguity encountered for the ring cross-peaks.[130,131] These can also elucidate and confirm the tricky proton assignments that cannot be attained from the homonuclear experiments, including H5 and H6 of Gal sugars. 2D HSQC-TOCSY establishes multiple-bond ^{1}H-^{13}C scalar intra-residue correlations for each spin system from diagnostic cross-peaks, usually the anomeric signals. This is usually interpreted by overlaying it with the HSQC-DEPT spectrum (on top) and the selective 1D TOCSY experiments. Thus, aiding the cross-tracking of carbon signals of each spin system from their anomeric HSQC cross-peaks vertically along the well-resolved ^{13}C axis, whereas horizontally, the 1D TOCSY spectra allow the cross-validation of the ring proton signals to their HSQC cross-peaks. Similarly, HSQC-NOESY can be assigned to overcome the issue of small scalar coupling constants by providing ^{1}H-^{13}C dipolar correlations based on space coupling between protons. Similar to the standard NOESY experiment, HSQC-NOESY also provides inter-residue correlations, which can establish the sequence and linkage positions of the sugar residues.

The heteronuclear multiple bond correlation (HMBC) experiment, which can provide two- and three-bond ^{1}H-^{13}C scalar intra- and inter-residue correlations, is usually recorded and analysed as the final step in the structural elucidation of polysaccharides.[132] The intra-residue correlations are dependent on the sugar identity and coupling constants. Since the transfer of magnetisation is not limited to proton-carbon bonds, three-bond intra-residue correlations, H1 to C3 and H1 to C5 (through the ring oxygen bonds) can be observed for α-sugars, while H1 to C5 and sometimes C2 can be observed for β-sugars.[108] Notably, HMBC allows the assignment of carbonyl carbons (C6) of uronic acids from H4 or H5, and N- or O-acetyl groups. Three-bond inter-residue cross-peaks from the anomeric protons of residues to the linkage carbons of their neighbouring residue can also be assigned to establish linkage positions. These are corroborated by correlations from the linkage ring protons to the anomeric carbon of the preceding residue. Since the inter-residue correlations from HMBC are through scalar couplings, this unambiguously elucidates all the linkage positions and the sequence of the residues in the RU (as shown in blue in **Figure 2.5**).

2.2 Experimental methods for recording NMR experiments

Purified capsular polysaccharide samples of different strains were obtained from two collaborators. The first sample was acquired from the University of Trieste, Italy, and the remaining four samples were from the GVGH Vaccines Institute for Global Health, Siena. However, one of the capsular polysaccharide samples acquired from the GVGH Vaccines Institute for Global Health was found to be the same as that received from the University of Trieste. Each polysaccharide powder sample (3 - 5 mg) was prepared by exchanging thrice with D_2O, dissolved in the appropriate amount of D_2O (~0.7 mL), allowed to stand for about 15 min, and transferred to a 5 mm NMR tube.

1D (^{1}H, DOSY, ^{13}C and DEPT-135) and 2D (COSY, DOSY-TOCSY, NOESY, HSQC, HMBC and hybrid HSQC-TOCSY and HSQC-NOESY) NMR spectra were obtained using a Bruker Avance III 600 MHz NMR spectrometer equipped with a BBO Prodigy cryoprobe. The spectra were recorded and processed using the standard Bruker software (Topspin 4.1.4). Typically, 1D proton spectra were recorded using a 30-degree pulse and a D1 time of 2s, at temperatures between 318 - 343 K, dependent on what is found to be optimal for each CPS, as shown the quality (peak resolution) of their spectra. DOSY was recorded with a gradient strength (GPZ6) =70%, diffusion time (d20, “big delta”) = 0.1 s, and p30 (“little delta”) = 1000 µs. To reduce the acquisition time of the 2D experiments, a non-uniform sampling (NUS) of 20 - 50% was used. 2D DOSY-TOCSY experiments were performed using a mixing time of 150 - 180 ms, whereas 1D variants were recorded with a mixing time of 25 - 200 ms. 2D NOESY and 1D variants were recorded using a mixing time of 100 - 300 ms. The ^{1}H-^{13}C HSQC and HMBC experiments were optimized for J = 145 Hz and 6 Hz, respectively, and the HSQC-TOCSY and HSQC-NOESY experiments were recorded using mixing times of 120 and 300 ms, respectively. The spectra were referenced relative to acetone added to the samples later, which gave a ^{1}H signal at 2.225 ppm and a ^{13}C signal at 31.07 ppm. To obtain the required correlations, some experiments were repeated at higher resolution, using more scans; additionally, where peaks were broad and poorly resolved, further sonication was performed to improve the quality and resolution of the spectrum. The full list of recorded experiments and the corresponding Bruker pulse programs and coupling constants/mixing times (if applicable) are given in **Table 2.1**.

Table 2.1: NMR experiments performed, and corresponding Bruker NMR pulse programs used to record 1D and 2D NMR experiments for the structural characterization of *K. pneumoniae* capsular polysaccharides. The exact mixing times used in the data acquisition of the experiments for each CPS are provided in the respective chapters.

Experiment	Pulse program	Coupling constant/Mixing time
1D proton (^{1}H)	zg30	-
1D DOSY (Diffusion Ordered Spectroscopy)	ledbpgp2s1d	-
1D Selective TOCSY (Total Correlation Spectroscopy)	selmlgp	25 - 200 ms
1D selective NOESY (Nuclear Overhauser Effect Spectroscopy)	selnogp	100 - 300 ms
1D ^{13}C	zgpg30	-
1D ^{13}C-DEPT 135	deptsp135	-
2D COSY (Correlation Spectroscopy)	cosygpprqf	-
2D DOSY-TOCSY	ledbpgpml2s2d	150 - 180 ms
2D NOESY	noesyphpr	300 ms
2D HSQC-DEPT (Heteronuclear Single Quantum Coherence)	hsqcedetgpsisp2.3	145 Hz
2D HSQC-TOCSY	hsqcdietgpsisp	120 ms
2D HSQC-NOESY	hsqcetgpnosp	300 ms
2D HMBC (Heteronuclear Multiple Bond Correlation)	hmbcgplpndqf	6 Hz

Chapter 3: Structural elucidation of the K102 CPS RU

The K102 serotype of *K. pneumoniae* was identified as a novel serotype through genotyping and determined as KL102. This is one of the emerging clinically important strains, which are anti-microbial resistant and responsible for severe infections such as neonatal sepsis and related deaths. Our collaborators at the University of Trieste (Prof. Cescutti), received the strain from Prof. D'Andrea (University of Rome, Tor Vergata), cultivated it and performed the isolation and purification of the CPS (K-antigen). They provided the purified CPS with preliminary chemical composition analysis, which indicated the presence of galactose (Gal), glucose (Glc), and glucuronic acid (GlcA), and a sample (~15 mg) was provided for NMR characterisation while more detailed chemical studies were being performed at the University of Trieste. Full characterisation of the K102 CPS and elucidation of its RU structure was performed using a combination of 1D and 2D NMR experiments at a temperature of 318 K. The proposed RU structure was then compared to the results obtained for the chemical structural studies from our collaborators and to the database of known *K. pneumoniae* CPSs to confirm if the strain is novel.

3.1 1D ^{1}H and ^{13}C NMR experiments

As a starting point, the 1D ^{1}H NMR spectrum of K102 CPS gave well-resolved diagnostic signals in the anomeric region (4.4 - 5.6 ppm), however, there was a large solvent deuterium oxide or water (HOD) signal within this region at 4.59 ppm, which may obscure some signals (**Figure 3.1a**). The HOD signal was removed by recording a 1D DOSY experiment, which confirmed that there were no signals underneath it and showed the presence of six anomeric signals (5.45; 5.06; 4.81; 4.67; 4.65; and 4.50 ppm). This suggested that the CPS of K102 consists of six sugar residues, in other words, a hexasaccharide repeating unit (RU). The anomeric signals were labelled, H1 A to F, starting from the most downfield (5.45 ppm) to the most up-field (4.50 ppm) signal, with H1 D (4.67 ppm) and H1 E (4.65 ppm) overlapping, resulting in the broad, asymmetrical, peak observed (**Figure 3.1b**). Based on the chemical shifts and coupling constants of these anomeric signals, it was concluded that residue A (5.45 ppm) and residue B (5.06 ppm) were α-sugars, while the remaining residues C to F (4.81 - 4.50 ppm) were β-sugars.

The ring region of the 1D ^{1}H/1D DOSY spectra was heavily overlapped as expected, and there were no signals in the methyl region, which confirmed the absence of 6-deoxy sugar residues, or substituents such as O-acetyl groups and pyruvate groups in the CPS of K102.

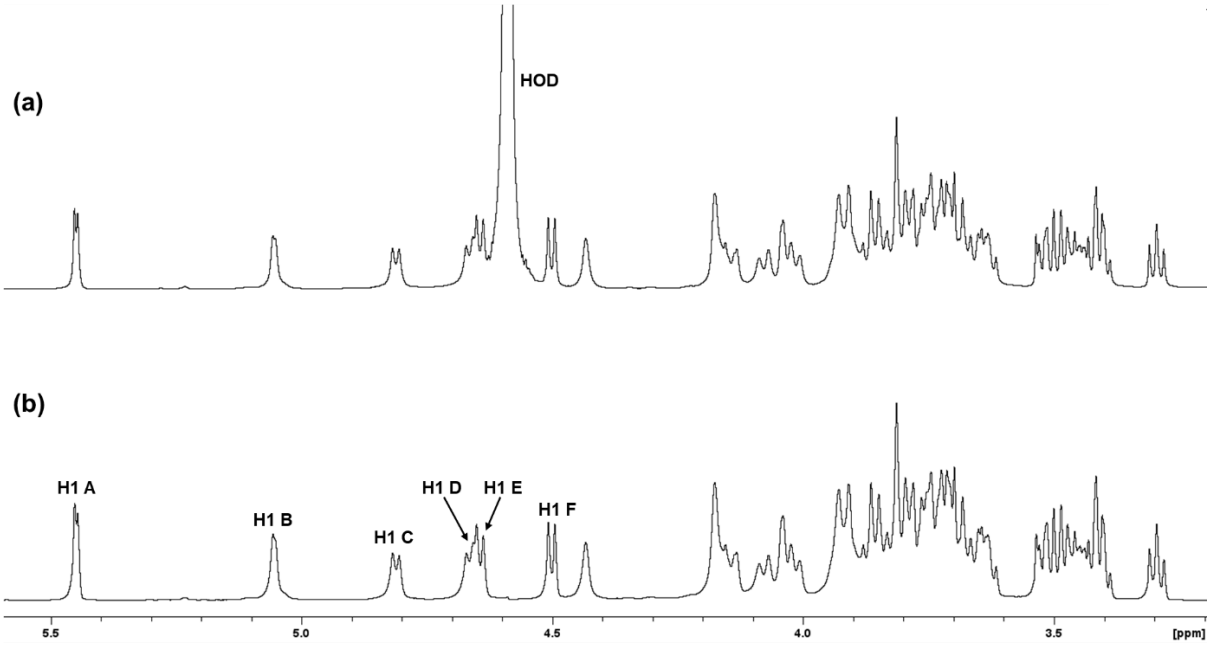


Figure 3.1: Overlay of the anomeric and ring regions of the **(a)** 1D ^{1}H NMR spectrum and **(b)** 1D ^{1}H-DOSY NMR spectrum of K102 CPS recorded at 318 K and 600 MHz with the diagnostic anomeric signals labelled from H1 A-F.

Similarly, the 1D ^{13}C NMR spectrum also revealed six signals in the anomeric region (**Figure 3.2**), thus confirming the hexasaccharide RU. Additionally, the 1D ^{13}C NMR spectrum gave a diagnostic signal at 175.7 ppm, characteristic of a carboxyl group (COOH) of a uronic acid residue, thus confirming the presence of GlcA as provided by composition analysis. Moreover, 1D ^{13}C-DEPT was recorded, which revealed five methylene (CH_2 or C6) signals as inverted peaks, consistent with a hexasaccharide RU that consists of five regular sugar residues and one uronic acid residue (**Figure 3.3a**). Two of the C6 signals were more deshielded (resonating at 67.0 ppm and 70.95 ppm), which indicated that two residues in the RU were C6 linked.

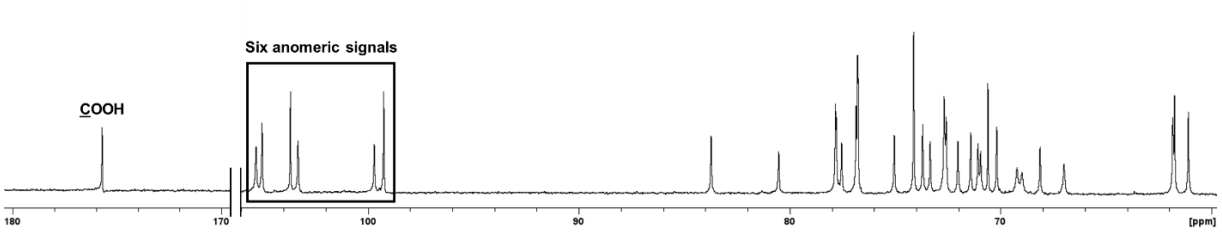


Figure 3.2: 1D ^{13}C NMR spectrum of K102 CPS recorded at 318 K and 150 MHz showing the diagnostic COOH signal of an uronic acid residue labelled, and the six anomeric signals highlighted by a rectangular border.

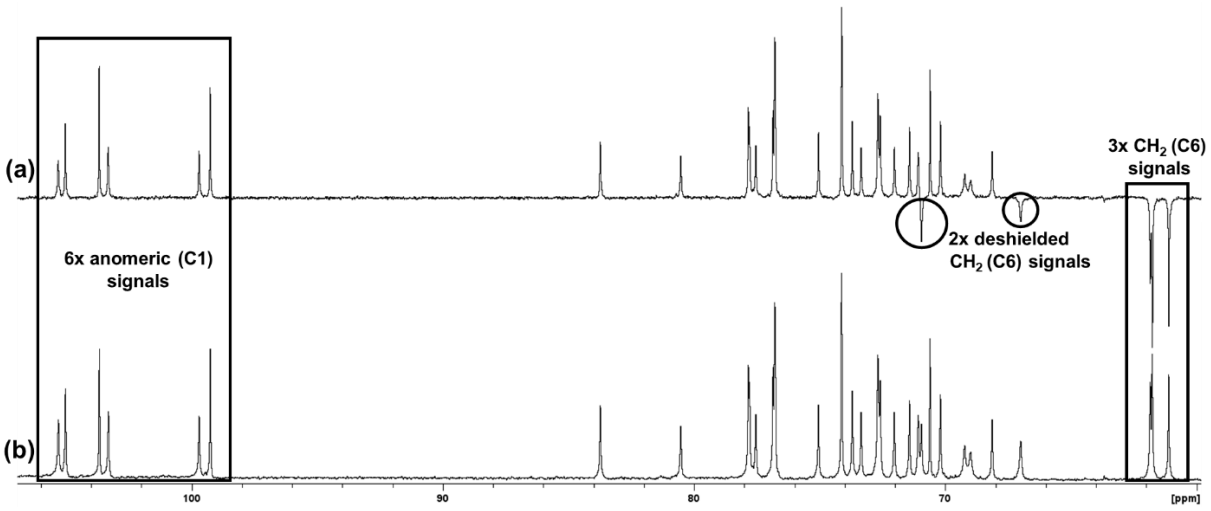


Figure 3.3: Overlay of **(a)** 1D ^{13}C-DEPT and **(b)** 1D ^{13}C NMR spectra of K102 CPS expanded in the anomeric and ring regions recorded at 318 K and 150 MHz. 1D ^{13}C-DEPT shows the methylene (CH_2) signals inverted.

The assignments of the carbon signals follow from the identification of the ring protons for each spin system; this was achieved using a combination of 1D and 2D proton-proton (^{1}H-^{1}H), homonuclear, correlation experiments with the diagnostic anomeric signals from 1D ^{1}H-DOSY used as a starting point.

3.2 Homonuclear (^{1}H-^{1}H) COSY, TOCSY, and NOESY experiments

Various proton assignments in the ring region can be established with 2D Correlation Spectroscopy (COSY), by following the scalar ^{1}H-^{1}H correlations starting from the anomeric signal of each residue. The 2D COSY spectrum of K102 CPS gave correlations from H1 to H2 for each residue, from H2 to H3 for residues A, B, C, and F, and from H3 to H4 for residue B, thereafter, the extensive overlap of ring signals hindered the elucidation of further correlations from COSY (**Figure 3.4**).

With the aid of the 2D Total Correlation Spectroscopy (TOCSY) experiment, which provides well-resolved multiple-bond scalar correlations from the anomeric signals, further assignments were elucidated dependent on the magnitude of the $J_{Ha,Hb}$ coupling constants between any neighbouring protons (Ha and Hb) for each sugar type. This led to the identification of each sugar residue type (Glc, GlcA, or Gal) in the hexasaccharide RU. 2D DOSY-TOCSY of K102 CPS was recorded (**Figure 3.5**), (with DOSY removal of cross-peaks from low molecular weight compounds including the HOD signal), and the assignment of this spectrum was facilitated by overlaying it with the 2D COSY, to clearly show the additional correlations established by TOCSY.

Using the most deshielded anomeric signal, H1 A at 5.45 ppm, 2D DOSY-TOCSY gave a correlation to H2 A at 3.53 ppm as in COSY, and revealed additional multiple-bond correlations, to H3 A at 3.70 ppm, H4 A at 3.42 ppm, H5 A at 3.72 ppm (overlapping with H3 A), and H6/6’ A at 3.83 and 3.78 ppm.

Therefore, DOSY-TOCSY elucidated the full assignments (H1 to H6) of residue A, and this is characteristic of an α-glucopyranose (α-Glc) residue.

From the anomeric signal of residue B, H1 B at 5.06 ppm, DOSY-TOCSY gave a correlation to H2 B at 4.14 ppm as in COSY, and additional correlations, to H3 B at 4.02 ppm, and H4 B at 4.43 ppm at which the scalar correlations cease. Thus, DOSY-TOCSY only elucidated H1 to H4 for residue B, this is characteristic of an α-galactopyranose (α-Gal) residue as the correlations to H4 and H5 are not usually seen, due to the small $J_{H4, H5}$ coupling constant between H4 and H5 for Gal sugars. Similarly, from H1 of residue C at 4.81 ppm, DOSY-TOCSY gave correlations to H2 C at 3.65 ppm, H3 C at 3.86 ppm, and H4 C at 4.18 ppm. For residue D H1 D at 4.67 ppm, DOSY-TOCSY gave correlations to H2 D at 3.63 ppm, H3 D at 3.72 ppm, and H4 D at 4.04 ppm, wherein the scalar coupling ceased. Therefore, the DOSY-TOCSY correlations from H1 of C and D elucidated H1 to H4 for each residue, characteristic of β-galactopyranose sugars (designated β-Gal^{I} and β-Gal^{II} for residue C and D, respectively).

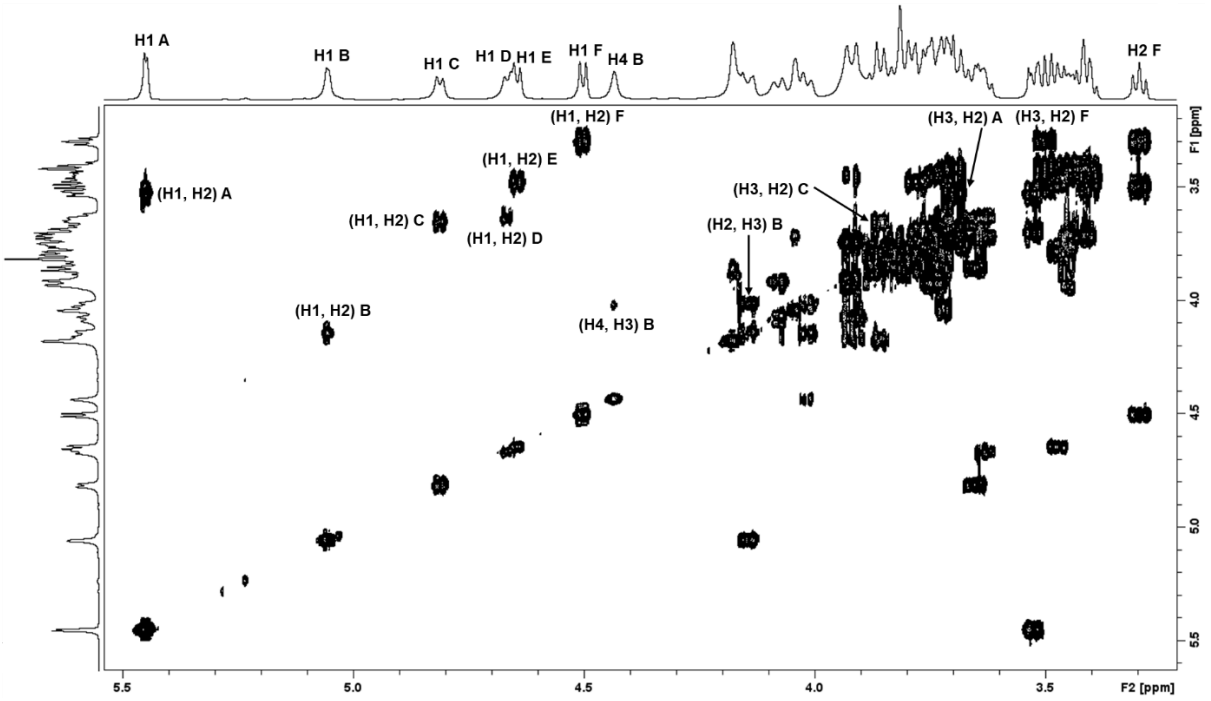


Figure 3.4: 2D ^{1}H-^{1}H COSY spectrum of K102 CPS recorded at 318 K and 600 MHz with diagnostic correlation cross-peaks labelled.

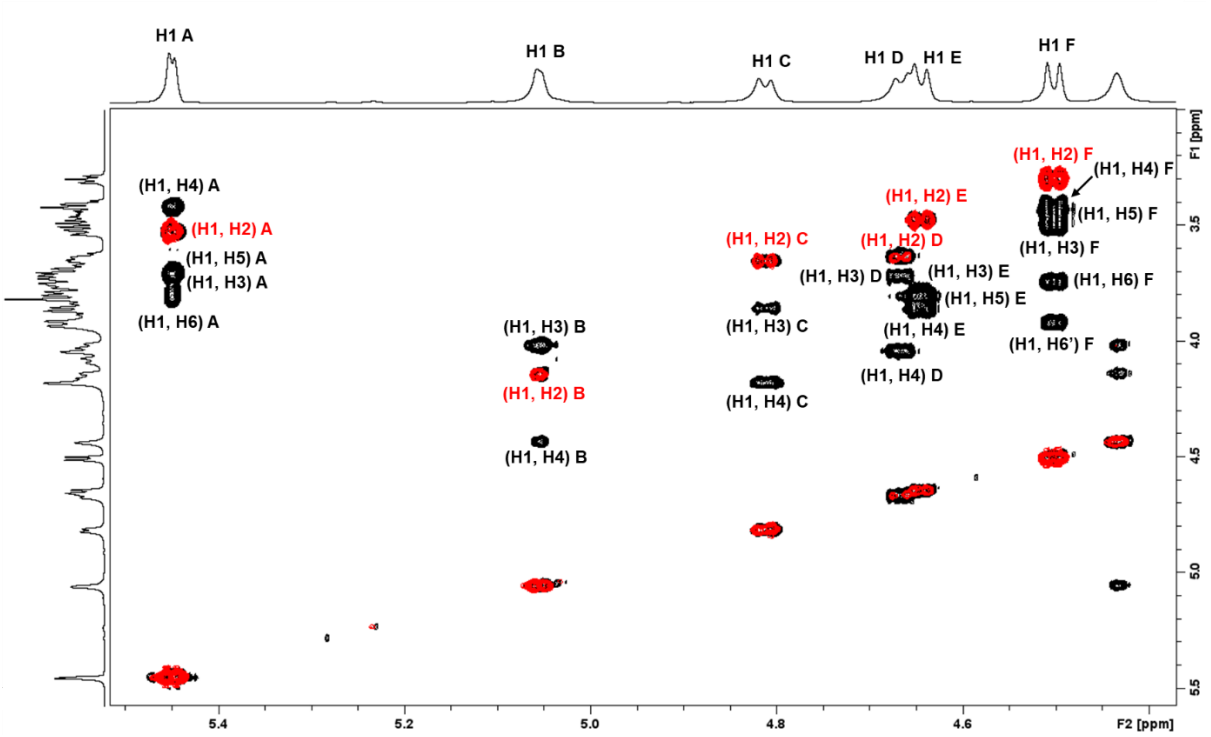


Figure 3.5: Anomeric region of the 2D COSY spectrum (red) overlaid on top of 2D DOSY-TOCSY spectrum (black) of K102 CPS recorded at 318 K and 600 MHz. A = α-Glc; B = α-Gal; C = β-Gal^{I}; D = β-Gal^{II}; E = β-GlcA; and F = β-Glc.

The anomeric signal of residue E at 4.65 ppm gave a correlation to H2 E at 3.47 ppm in DOSY-TOCSY and continued to H3 E at 3.77 ppm, H4 E at 3.86 ppm, and H5 E at 3.81 ppm, this is characteristic of a β-glucuronic acid (β-GlcA). Using the most up-field anomeric signal, H1 F at 4.50 ppm, DOSY-TOCSY gave correlations to H2 F at 3.30 ppm, H3 F at 3.50 ppm, H4 F at 3.41 ppm, H5 F at 3.45 ppm, and H6/6' F at 3.74 and 3.92 ppm. This is characteristic of a β-glucopyranose (β-Glc). The combination of 2D COSY and DOSY-TOCSY experiments elucidated the full proton assignments for the Glc (A and F) and GlcA (E) residues but elucidated only H1 to H4 assignments for the Gal residues (B, C, and D). These assignments were also confirmed by a series of 1D TOCSY experiments recorded by selectively irradiating the anomeric signal of each spin system; the 1D profiles also revealed the multiplicity of the signals.

An overlay of the 1D TOCSY spectra of each anomeric signal (**Figure 3.6a-c, e-g**) with the 1D DOSY spectra (**Figure 3.6d**) shows the assignments from H1 to H6 for residue A (α-Glc), H1 to H4 for residues B (α-Gal); C (β-Gal^{I}), and D (β-Gal^{II}), H1 to H5 for residue E (β-GlcA), and H1 to H6 for residue F (β-Glc). At this point, the remaining proton signals unassigned are the H5 and H6 of the Gal residues.

1D NOESY experiments were recorded by irradiating H1 of each residue to establish their intra- and inter-residue dipolar correlations. For residue A (α-Glc), 1D NOESY gave a single intra-residue

correlation to H2, as expected for an α-pyranose sugar, due to the equatorial configuration of H1, and gave an inter-residue correlation to H4 E (β-GlcA), thus suggesting that residue A (α-Glc) is linked to residue E (β-GlcA) in the RU (**Figure 3.7a**). Similarly, for residue B (α-Gal), 1D NOESY gave a single intra-residue correlation to H2, consistent with α-Gal, and gave inter-residue correlations to a preliminary unassigned signal at 3.93 ppm, and a pair of preliminary unassigned H6/6' signals at 3.89 and 3.82 ppm, (confirmed from ^{1}H-^{13}C HSQC-DEPT spectrum, which showed their correlation cross-peaks as inverted signals), (**Figure 3.7b**).

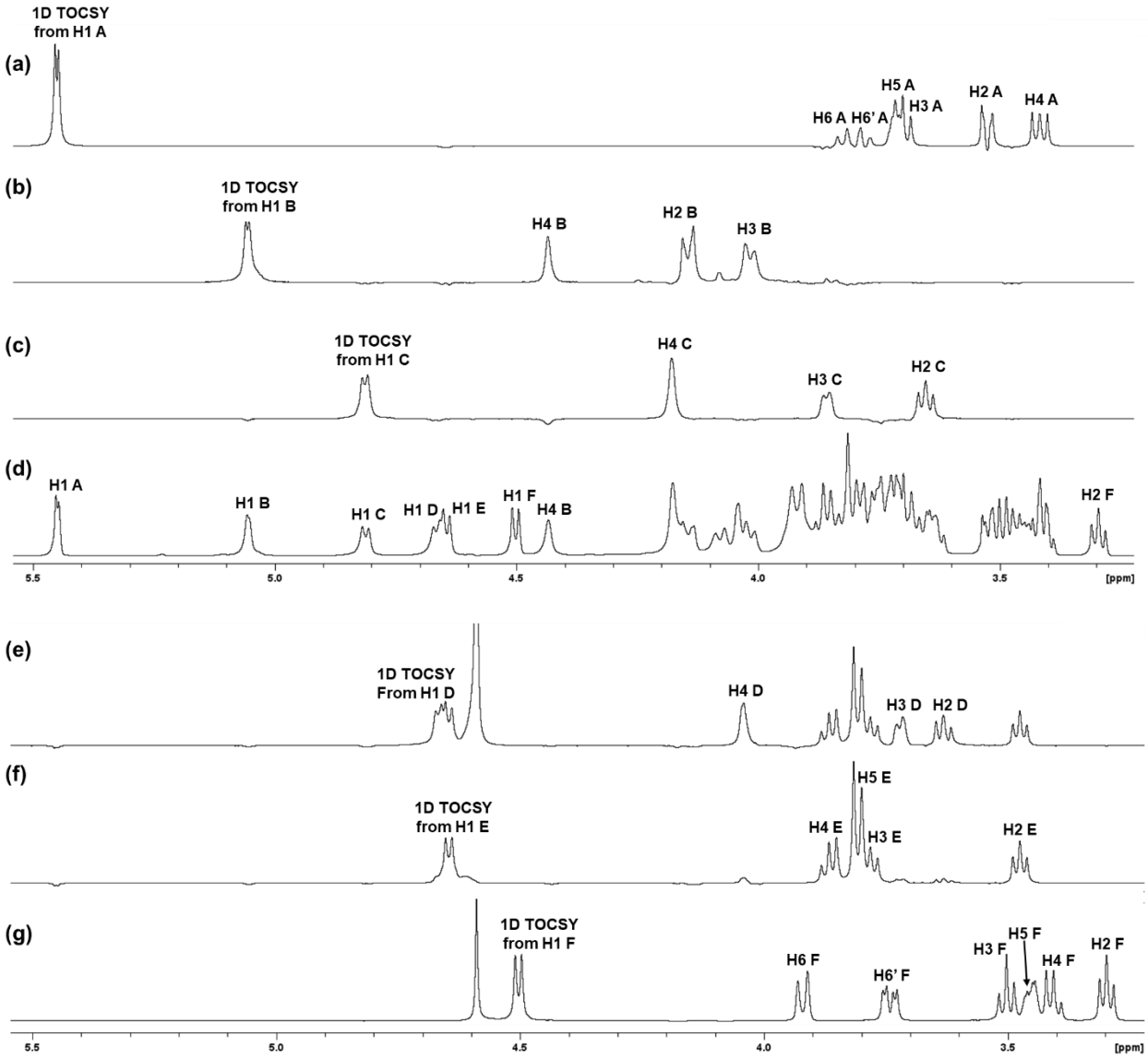


Figure 3.6: 1D TOCSY spectra for each spin system of K102 CPS with labels of the anomeric signals and their respective diagnostic intra-residue correlations given, obtained by selectively irradiating H1 of residues A (α-Glc) **(a)**, B (α-Gal) **(b)**, and C (β-Gal^{I}) **(c)**, overlaid onto 1D DOSY **(d)**, which shows assignments of the anomeric signals and some diagnostic ring signals, to which the assignments of the anomeric signals and intra-residue correlations of the TOCSY spectra are cross-validated. Similarly, 1D TOCSY spectra from H1 D (β-Gal^{II}) **(e)**, E (β-GlcA) **(f)**, and F (β-Glc) **(g)**, are presented.

The latter inter-residue correlations are characteristic of a C6 linkage, whereby inter-residue correlations from H1 of a residue to H5 and H6/6' of its neighbouring residue are typically observed. This is not a surprise as the analysis of the 1D ^{13}C-DEPT (**Figure 3.3**) suggested that two residues were C6 linked. The assignment of the signal at 3.93 ppm was elucidated as H5 D by the 1D NOESY from H1 D (**Figure 3.7e**), which gave intra-residue correlations to H2 D, and additional correlations, to H3 D, and H5 D at 3.93 ppm, as expected for a β-pyranose sugar due to the axial orientation of H1, H3 and H5. Consequently, this revealed that B (α-Gal) is linked to C6 D (β-Gal^{II}). Moreover, 1D NOESY from H1 of D gave an inter-residue correlation to H4 C, suggesting that D (β-Gal^{II}) is linked to C (β-Gal^{I}). For residue C (β-Gal^{I}), 1D NOESY gave intra-residue correlations to H2 C, and additional correlations, to H3 C and H5 C at 3.68 ppm, as expected for a β-pyranose sugar, and gave inter-residue correlations to H3 B (weak signal) and H4 B, thus suggesting that C (β-Gal^{I}) is linked to B (α-Gal) (**Figure 3.7c**). Similarly, for the rest of the β-sugars, E (β-GlcA) and F (β-Glc), 1D NOESY from H1 gave the expected intra-residue correlations for a β-pyranose sugar, to H2, H3 and H5 signals (**Figure 3.7f-g**). Finally, 1D NOESY from H1 E (β-GlcA) gave inter-residue correlations to H3 B and H4 B (weak signal), suggesting that E (β-GlcA) is linked to B (α-Gal), and 1D NOESY from H1 F (β-Glc) gave an inter-residue correlation to a signal at 4.16 ppm and a pair of preliminary unassigned H6/6' signals at 4.08 and 3.91 ppm. The only proton assignments that had not been elucidated were H5 and H6/6' of B, and H6/6' C; however, it can be postulated that the inter-residue correlations from H1 F are H5 B at 4.16 ppm and H6/6' B at 4.08 and 3.91 ppm, thus F (β-Glc) is linked to C6 B (α-Gal); this means that B (α-Gal) is the second C6 linked sugar residue. Therefore, the remaining unassigned protons at ~3.75 ppm must be H6/6' of C (β-Gal^{I}).

Thus, 1D NOESY elucidated the assignments of the H5 signals for the β-Gal residues (C and D), and consequently elucidated the assignments of H5 and H6/6' of B (α-Gal), and H6/6' of D (β-Gal^{II}), and gave the sequence of sugars in the RU (**Figure 3.8**), although, there may be some ambiguity regarding the linkage positions (this will be confirmed in later experiments). Thus, the full assignments for all the spin systems (except for H6/6' C) could be obtained from the combination of 1D and 2D homonuclear experiments of K102 CPS, and these will be used in the assignment of the attached carbons using 2D proton-carbon (^{1}H-^{13}C) experiments.

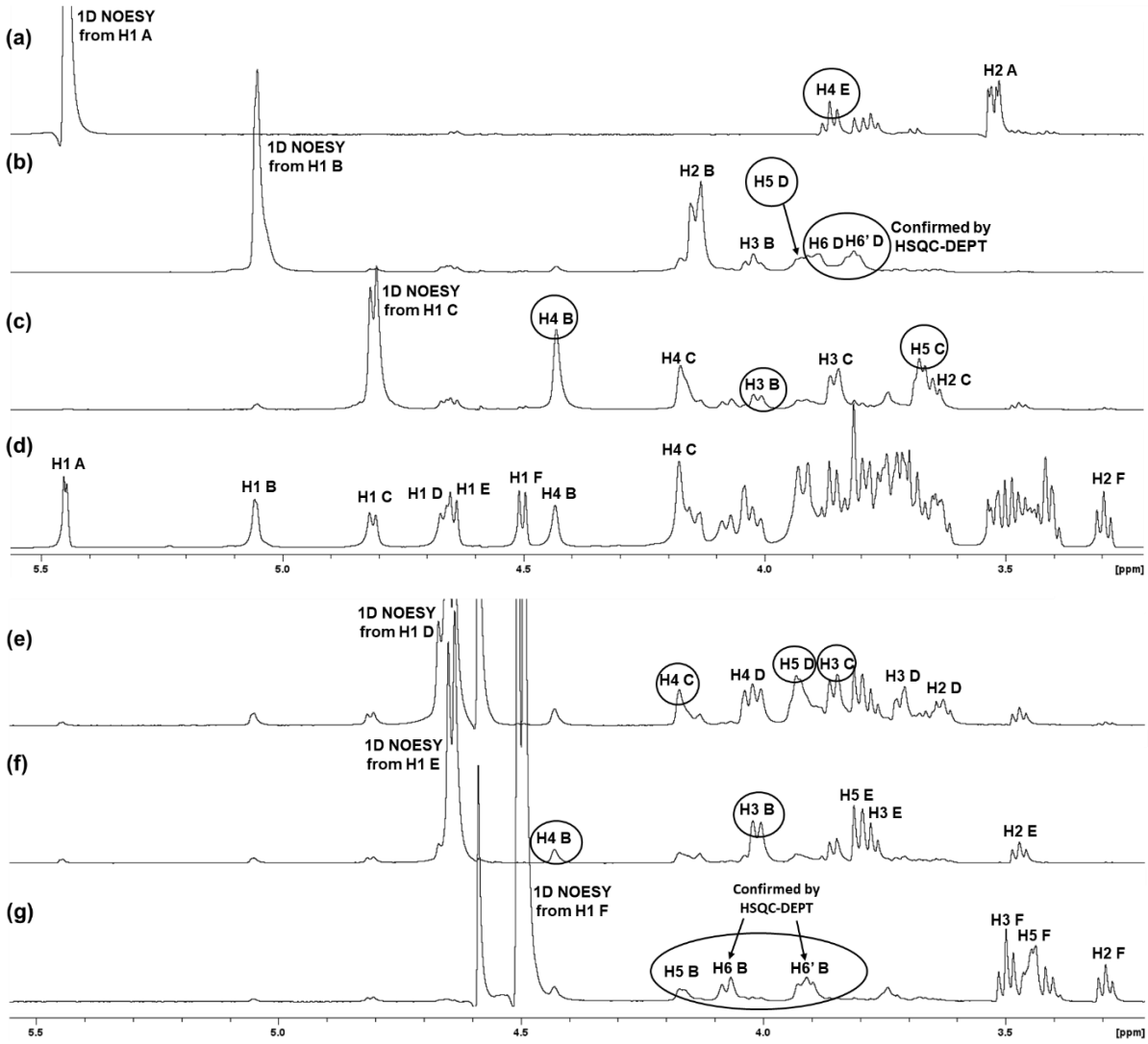


Figure 3.7: 1D NOESY spectra for each spin system of K102 CPS obtained by irradiating the anomeric signals of residues A (α-Glc) **(a)**, B (α-Gal) **(b)**, and C (β-Gal^{I}) **(c)**, overlaid onto 1D DOSY **(d)**, and from the anomeric signals of D (β-Gal^{II}) **(e)**, E (β-GlcA) **(f)**, and F (β-Glc) **(g)**. The anomeric signals are labelled on the NOESY spectra **(a-c & d-f)** and cross-validated onto the 1D DOSY overlays **(d)**, which also features labels for other diagnostic signals within the ring region, furthermore, the assignments of the key NOESY intra-residue correlations are given on each respective 1D NOESY spectrum, and the assignments of the inter-residue correlations represented within ovals.

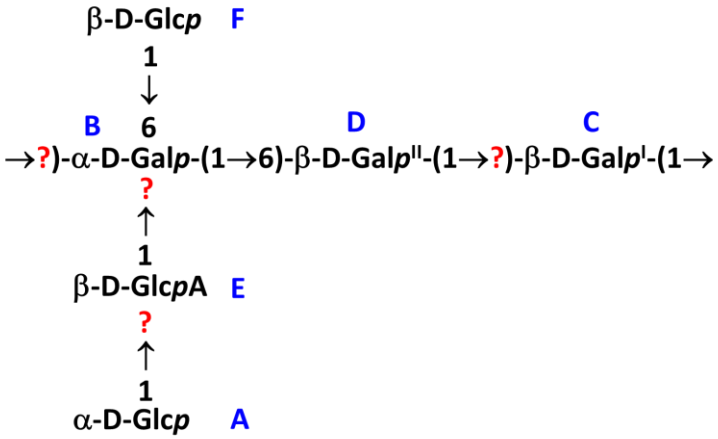


Figure 3.8: Proposed structure of K102 CPS showing the identity of sugar residues (A-F) as elucidated by scalar (2D COSY, 2D TOCSY, and 1D TOCSY) and dipolar (1D NOESY) homonuclear experiments.

3.3 2D Heteronuclear ^{1}H-^{13}C experiments (HSQC-DEPT, HSQC-TOCSY, HSQC-NOESY, and HMBC)

At this point, more information about the RU structure of K102 CPS was still required, including the assignments of the ^{13}C chemical shifts of every carbon atom of each sugar residue (i.e., from the 1D ^{13}C NMR spectrum), and confirmation of the sequence and linkage positions of the sugar residues in the RU. This information was obtained from a combination of 2D ^{1}H-^{13}C experiments using the preliminarily proton assignments as a starting point in their analysis. Moreover, the 2D heteronuclear experiments allowed further confirmation of the previous assignments and elucidated the remaining assignment for H6/6' of residue C (β-Gal^{I}).

The assignment of the 2D heteronuclear experiments began with the 2D HSQC spectrum, which gives correlations from every ^{1}H signal in the 1D ^{1}H NMR spectrum (shown along the horizontal axis of the 2D spectrum) to the ^{13}C signal of the carbon that it is directly attached to on the 1D ^{13}C NMR spectrum (shown along the vertical axis of the 2D spectrum). The 2D HSQC-DEPT spectrum (which provides the H6-C6 cross-peaks as inverted signals in red) was recorded for K102 CPS. The anomeric ^{1}H-^{13}C correlation cross-peaks and most of the ring cross-peaks could be assigned with ease, and the process was facilitated by overlaying the 2D HSQC-DEPT spectrum with the 1D TOCSY of each residue, as shown in **Figure 3.9** using residue A as an example. However, the overlap of certain ring proton signals resulted in ambiguity in the assignment of their ^{1}H-^{13}C correlation cross-peaks. These included proton signals such as H3 C (β-Gal^{I}) and H4 E (β-GlcA) resonating at 3.86 ppm, resulting in two ^{1}H-^{13}C cross-peaks vertically aligned to these overlapping proton signals, similarly, this was observed for H5 A (α-Glc) and H3 D (β-Gal^{II}) resonating at 3.72 ppm, H4 A (α-Glc) and H4 F (β-Glc) resonating at 3.41 ppm and 3.42 ppm, respectively.

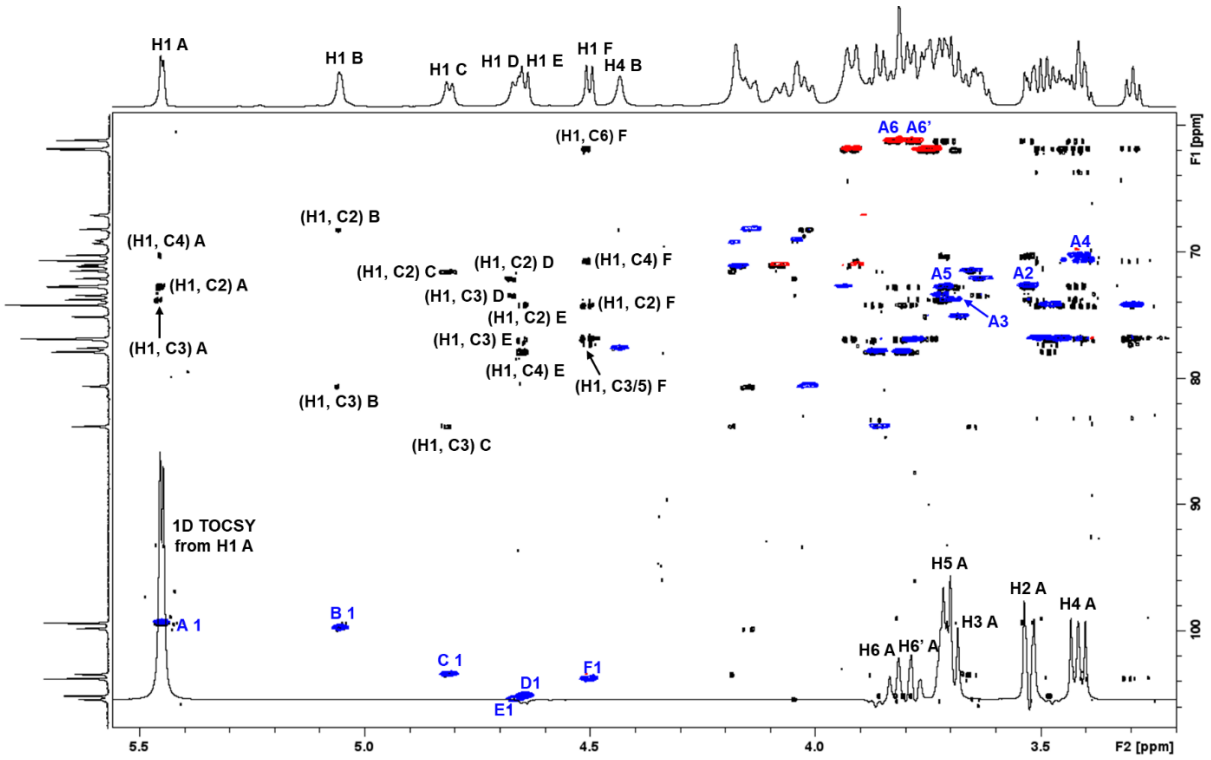


Figure 3.9: 2D ^{1}H-^{13}C HSQC-DEPT (blue and red) spectrum overlaid onto HSQC-TOCSY (black) and overlaid with 1D TOCSY from H1 A. The key anomeric assignments for the HSQC-DEPT and HSQC-TOCSY are given, and in the ring region, the labels of residue A are provided. A = α-Glc, B = α-Gal, and C = β-Gal^{I}, D = β-Gal^{II}, E = β-GlcA, and F =β-Glc.

This ambiguity was resolved with the aid of HSQC-TOCSY (**Figure 3.9**), which gives well-resolved scalar multiple-bond ^{1}H-^{13}C intra-residue correlation cross-peaks from the anomeric signal of each residue to its corresponding ^{13}C signals dependent on the sugar type. Some of these key intra-residue correlations given by the HSQC-TOCSY of K102 CPS included a correlation from H1 A (α-Glc) to C4 at 70.18 ppm, thus elucidating the assignment for H4-C4 A, and consequently elucidated H4-C4 F (β-Glc), which was confirmed by a correlation from H1 F to C4 at 70.60 ppm. From H1 of C (β-Gal^{I}), HSQC-TOCSY gave an intra-residue correlation to C3 at 83.72 ppm, thus elucidating H3-C3 C, and consequently elucidating H4-C4 E (β-GlcA), which was confirmed by a correlation from H1 E to C4 at 77.78 ppm, and gave a correlation from H1 D (β-Gal^{II}) to C3 at 73.34 ppm, thus elucidating the assignment of H3-C3 D, and consequently, H5-C5 A (α-Glc).

Further information was obtained from 2D HSQC-NOESY, which gives through-space ^{1}H-^{13}C intra- and inter-residue correlations. In the ring region, HSQC-NOESY gave intra-residue correlations from H5 to C6 D (β-Gal^{II}), H6 to C3 D, and H6 to C4 B (α-Gal), confirming the previous proton assignments postulated from 1D NOESY. These corresponded to the cross-peaks of H6-C6 of residue D and H6-C6 of residue B in the HSQC-DEPT, therefore, confirming the remaining methylene cross-peak as H6-C6

C. As in HSQC-TOCSY, some key intra-residue correlations were obtained from the anomeric protons, including H1 to C2 for residues A (α-Glc) and B (α-Gal), H1 to C5 for residue C (β-Gal^{I}), thus providing more confirmation for the assignment of H5-C5 C. From H1 of D (β-Gal^{II}), HSQC-TOCSY gave intra-residue correlations to C3 and C5, thus providing more confirmation for the assignment of H5-C5 D and gave H1 to C3 for F (β-Glc). Furthermore, HSQC-NOESY gave inter-residue correlations from H1 A to C4 E; H1 B to C6 D; H1 C to C4 B; H1 D to C3 C; H1 E to C3 B; and H1 F to C6 B. Thus, confirming the sequence of the residues in the RU that was elucidated by 1D NOESY, but in addition, provided the linkage positions of the residues without ambiguity as shown in **Figure 3.10**.

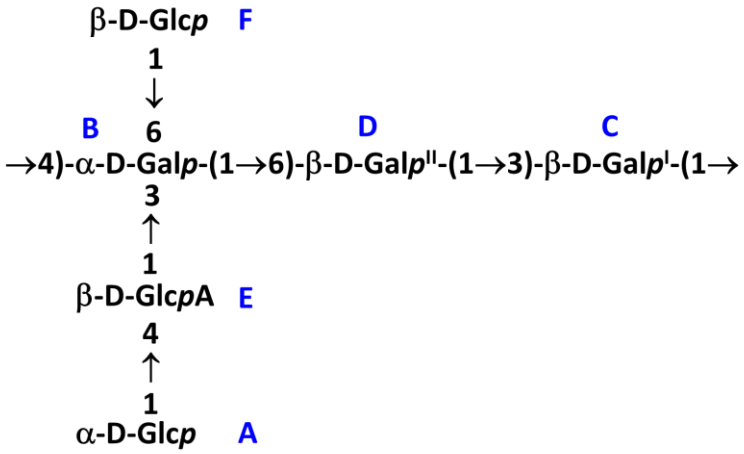


Figure 3.10: Proposed repeating structure of K102 CPS with the glycosylation pattern as provided by 2D HSQC-NOESY.

The characterisation and elucidation of the RU structure were finalised by the 2D ^{1}H-^{13}C HMBC experiment, which gives multiple-bond scalar ^{1}H-^{13}C intra- and inter-residue correlations. The HMBC intra-residue correlations are dependent on the sugar type, and since it can be set up to selectively give two or three-bond ^{1}H-^{13}C correlations, it is the main experiment for the elucidation or confirmation of the sequence and linkage positions of the residues. In the anomeric region, HMBC of K102 CPS (shown as an overlay with HSQC-DEPT in **Figure 3.11a**) gave intra-residue correlations from H1 to C3 and H1 to C5 for A (α-Glc); B (α-Gal); E (β-GlcA); and F (β-Glc), while no intra-residue correlations were observed for C (β-Gal^{I}) and D (β-Gal^{II}). Additionally, HMBC gave the following inter-residue correlations from H1 A (α-Glc) to C4 E (β-GlcA); H1 B (α-Gal) to C6 D (β-Gal^{II}); H1 C (β-Gal^{I}) to C4 B (α-Gal); H1 D (β-Gal^{II}) to C3 C (β-Gal^{I}); H1 E (β-GlcA) to C3 B (α-Gal); and H1 F (β-Glc) to C6 B (α-Gal), thus confirming the glycosylation pattern of the RU structure. Lastly, in the ring region HMBC gave a set of folded intra-residue correlations from H4 and H5 E to the COOH signal (**Figure 3.11b**), characteristic of an uronic acid residue, and thus gave further confirmation that E was the uronic acid residue, β-GlcA.

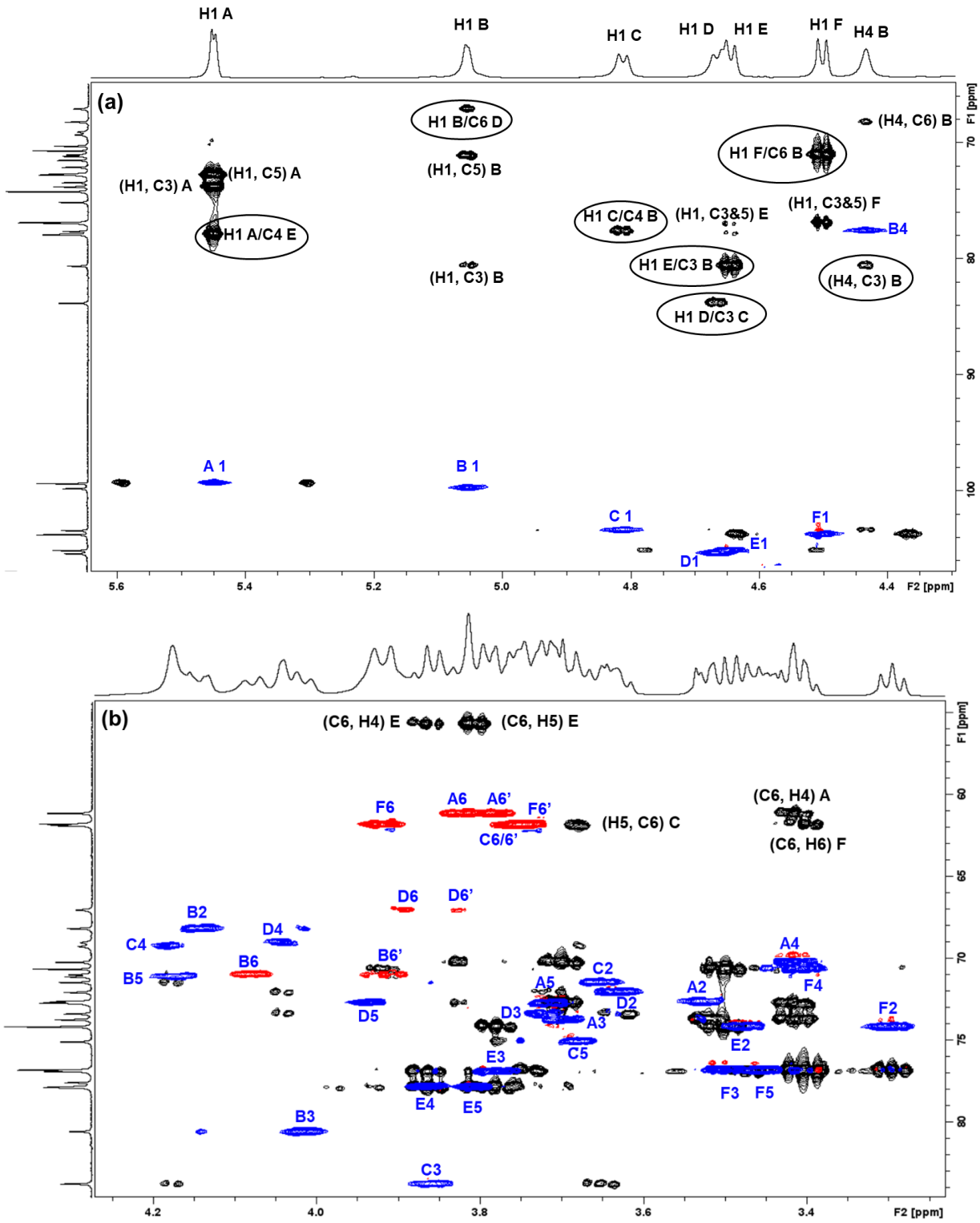


Figure 3.11: Anomeric region **(a)** and ring region **(b)** expansions of 2D ^{1}H-^{13}C HMBC overlaid with HSQC-DEPT of K102 at 318 K and 600 MHz, with labels of the key cross-peaks shown. A = α-Glc, B = α-Gal, C = β-Gal^{I}, D = β-Gal^{II}, E = β-GlcA, and F = β-Glc.

3.4 Summary of the full assignments and diagnostic spectra, and comparison of RU structural features to chemical analysis results

The structural characterisation and elucidation of the CPS of K102 began with the analysis of the anomeric region from the 1D ^{1}H/DOSY NMR spectra, which showed that K102 CPS was made up of a hexasaccharide RU. This was corroborated by 1D ^{13}C/^{13}C-DEPT NMR spectra, which confirmed the presence of an uronic acid residue in the chemical composition of the CPS. The assignments for nearly all the spin systems were established using a combination of 2D and 1D ^{1}H-^{1}H correlation experiments, COSY and TOCSY, which give scalar intra-residue correlations, and NOESY, which gives dipolar intra- and inter-residue correlations. Additionally, TOCSY provided the identification of the sugar type of each residue while NOESY indicated the sequence of the residues in the RU, although it was ambiguous in establishing the linkage positions.

The ^{1}H assignments were used to establish the assignments of the 1D ^{13}C NMR spectrum using 2D ^{1}H-^{13}C HSQC-DEPT experiment together with HSQC-hybrid experiments, HSQC-TOCSY and HSQC-NOESY, and the HMBC experiment. The elucidation of the proton-carbon pairs for each residue corroborated the proton assignments made using the homonuclear correlation experiments. The 2D HSQC-TOCSY gave ^{1}H-^{13}C intra-residue correlations which mainly resolved the ambiguity encountered in the assignment of some ^{1}H-^{13}C HSQC cross-peaks caused by overlapping ^{1}H signals in the ring region, while 2D HSQC-NOESY gave ^{1}H-^{13}C intra-residue correlations which provided confirmation of the tricky H5 and H6/6' assignments of the Gal sugar residues.

Additionally, HSQC-NOESY confirmed the sequence of residues in the RU and also elucidated the linkage positions through inter-residue correlations from each anomeric proton to the linkage carbon of the neighbour. The HMBC experiment provided confirmation of the assignment of the COOH of the GlcA residue by giving folded intra-residue correlations from H4 and H5 of β-GlcA to the COOH signal. Finally, HMBC provided confirmation of the linkage positions of the residues in the RU through its specialised three-bond scalar ^{1}H-^{13}C inter-residue correlations and thus completed the structural characterisation and elucidation of the CPS hexasaccharide RU of K102 (**Figure 3.12**).

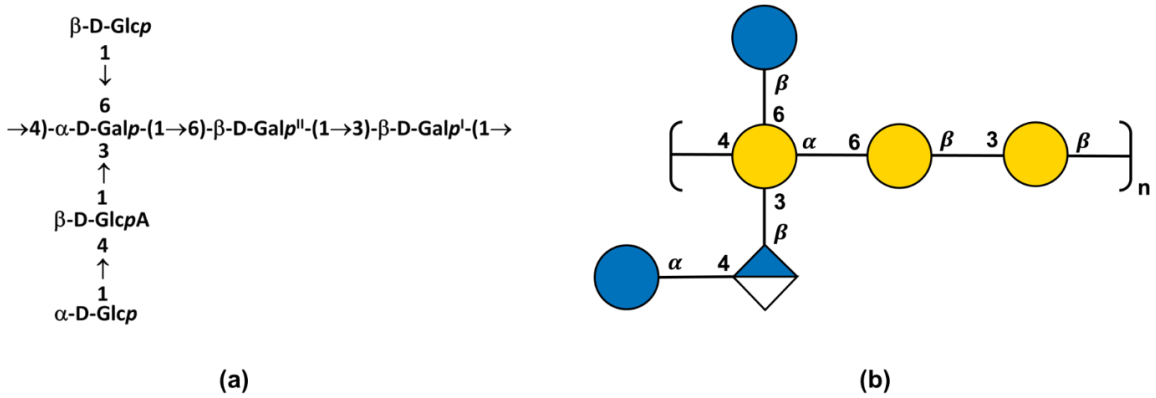


Figure 3.12: Elucidated repeating unit structure of the capsular polysaccharide of K102 with sugar residues represented in the International Union of Pure and Applied Chemistry (IUPAC) format (**a**) and the Symbol Nomenclature for Glycans (SNFG) format (**b**).

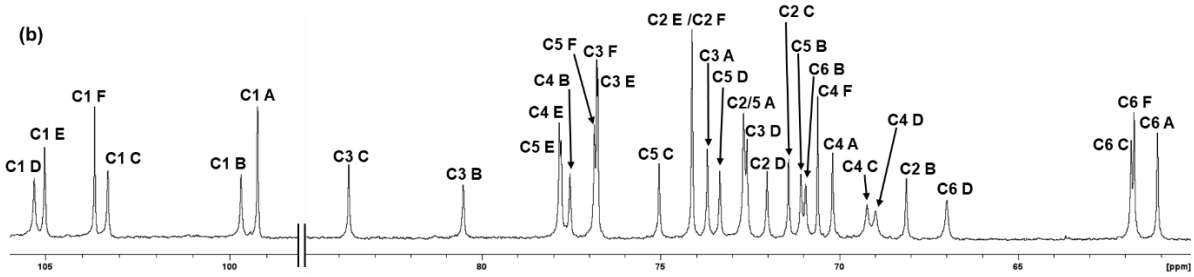


Figure 3.13: Fully labelled 1D ^{13}C NMR spectrum of K102 CPS expansion of the anomeric and ring regions recorded at 318 K and 150 MHz. A = α-Glc, B = α-Gal, C = β-Gal^{I}, D = β-Gal^{II}, E = β-GlcA, and F = β-Glc.

The fully labelled 1D ^{13}C NMR spectrum of K102 is provided in **Figure 3.13**, while the labelled 2D HSQC-DEPT spectrum is provided in **Figure 3.14**. The 1D ^{1}H and ^{13}C NMR spectra of K102 CPS serve as its identity "fingerprints" since the pattern or chemical shifts of their signals are diagnostic for its structure, and the 2D HSQC-DEPT spectrum serves as a 2D identity map of K102 CPS. These spectra can be used for identity testing of the CPS of K102.

The full ^{1}H and ^{13}C chemical shift data obtained from the 1D and 2D NMR experiments of K102 are summarised in **Table 3.1**. Using the CASPER chemical shifts prediction program, the chemical shifts of the unsubstituted monosaccharides present in K102 were obtained. These were used for confirmation of the identity of the sugar residues and to calculate the glycosylation shifts by subtracting the ^{13}C chemical shifts of the unsubstituted monosaccharides, obtained from CASPER, from the experimental ^{13}C chemical shifts of the residues in the CPS of K102. Large positive values of glycosylation shifts, which are shown in brackets below the underlined ^{13}C chemical shifts of the residues in **Table 3.1**, confirmed the linkage positions of the residues in the CPS RU.

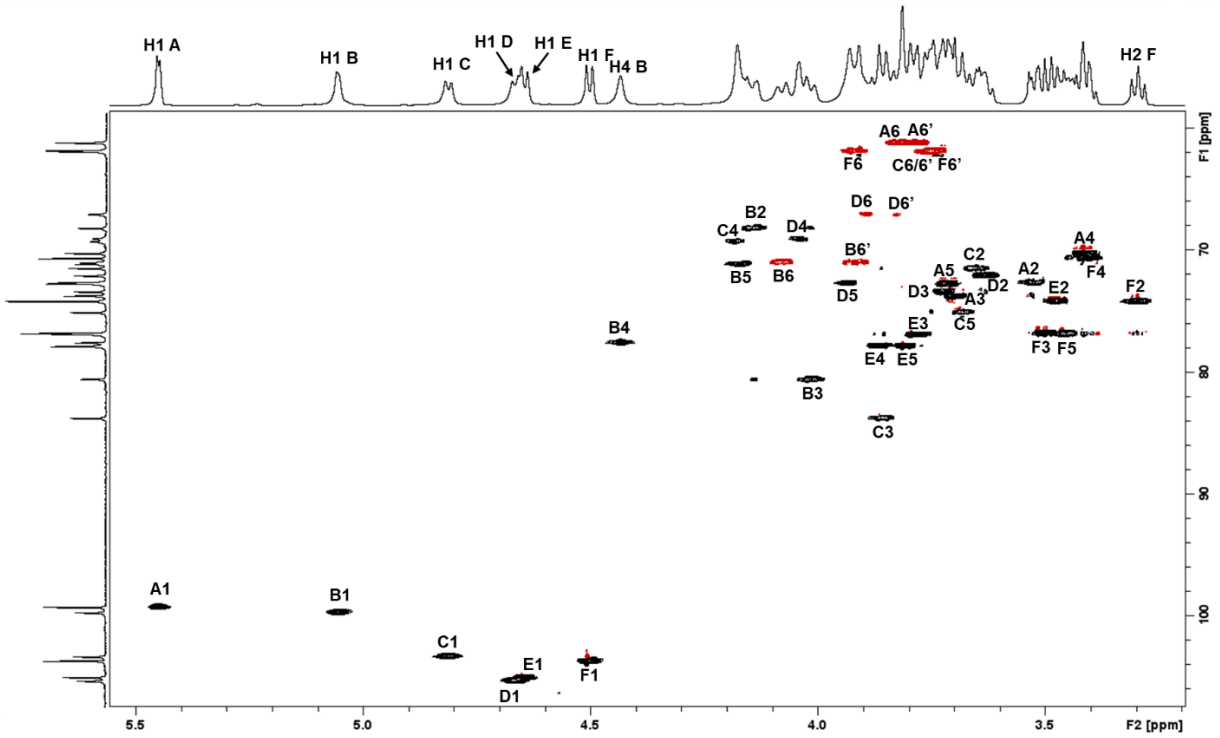


Figure 3.14: Fully labelled 2D ^{1}H-^{13}C HSQC-DEPT spectrum of K102 CPS recorded at 318 K and 600 MHz. Methylene cross-peaks are inverted (in red). A = α-Glc, B = α-Gal, C = β-Gal^{I}, D = β-Gal^{II}, E = β-GlcA, and F = β-Glc.

Further analysis of the magnitudes of the obtained glycosylation shifts and configurational factors reported by Shashkov[119] gave confirmation of the absolute configurations for nearly all the sugar residues of K102 CPS. It was found that the obtained glycosylation shift of 7.93 ppm for C1 of D (β-Gal*p*^{II}) is consistent with the reported shift range of +7.6 to +8.4 ppm, which is characteristic of a β-D-pyranose residue linked to C3 of a D-galactopyranose residue, thus confirmed the absolute configurations of D→3C as β-**D**-Gal*p*^{II}-(1→3)-β-**D**-Gal*p*^{I}. Similarly, the obtained glycosylation shift of 8.25 ppm for C1 of E (β-D-GlcA) is also consistent with the reported shift range of +7.6 to +8.4 ppm, corresponding to a β-D-pyranose residue linked to C3 of a D-galactopyranose residue, thus confirming the absolute configurations of E→3B as β-**D**-Glc*p*A-(1→3)-α-**D**-Gal*p*.

Structural features of the K102 RU were corroborated by chemical analysis performed at the University of Trieste. Methanolysis of the CPS followed by trimethylsilylation derivatisation and gas chromatography-mass spectrometric (GLC-MS) analysis showed the presence of galactose (Gal), glucose (Glc), and glucuronic acid (GlcA) in the molar ratio of 3.0:2.0:0.7, in agreement with the hexasaccharide sugar composition shown by NMR. Determination of the glycosidic linkages was achieved by GLC-MS analysis of the partially methylated alditol acetates which identified: terminal Glc*p*, 3-Gal*p*, 6-Gal*p*, and 3,4,6-Gal*p* in the molar ratio of 2.17:1.17:1.00:0.35. The methodology used

for linkage analysis does not detect uronic acids and is not quantitative, nevertheless, the data was consistent with a hexasaccharide repeating unit containing 3- and 6-linked Gal, a doubly branched Gal and two terminal Glc groups, as elucidated by NMR analysis.

Table 3.1: Complete ^{1}H-^{13}C NMR chemical shift data of the sugar residues in K102 CPS with the glycosylation shifts shown in brackets; the ^{13}C chemical shifts of the linkage positions are underlined.

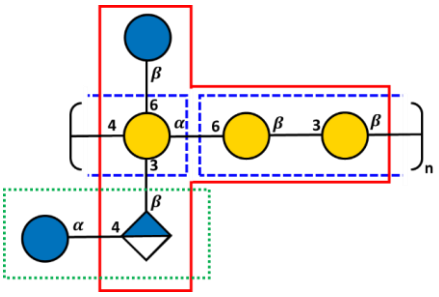
Sugar residues	H1 C1	H2 C2	H3 C3	H4 C4	H5 C5	H6/6' C6
A **α-D-Glc*p*-(1→**	5.45 99.25 (6.26)	3.53 72.68 (0.21)	3.70 73.70 (-0.08)	3.42 70.18 (-0.53)	3.72 72.68 (0.31)	3.83/3.78 61.10 (-0.74)
B **→3,4,6)-α-D-Gal*p*-(1→**	5.06 99.70 (6.52)	4.14 68.12 (-1.23)	4.02 <u>80.52</u> (10.39)	4.43 <u>77.53</u> (7.25)	4.16 71.07 (-0.23)	4.08/3.91 <u>70.95</u> (8.91)
C **→3)-β-D-Gal*p*^{I}-(1→**	4.81 103.31 (5.94)	3.65 71.41 (-1.55)	3.86 <u>83.72</u> (9.94)	4.18 69.22 (-0.47)	3.68 75.04 (-0.89)	3.75 61.83 (-0.01)
D **→6)-β-D-Gal*p*^{II}-(1→**	4.67 105.30 (7.93)	3.63 72.02 (-0.94)	3.72 73.34 (-0.44)	4.04 68.98 (-0.71)	3.93 72.69 (-3.24)	3.89/3.82 <u>67.00</u> (5.16)
E **→4)-β-D-Glc*p*A-(1→**	4.65 105.02 (8.25)	3.47 74.12 (-0.88)	3.77 76.56 (0.03)	3.86 <u>77.78</u> (5.09)	3.81 77.54 (0.61)	- 175.70 (-0.77)
F **β-D-Glc*p*-(1→**	4.50 103.67 (6.83)	3.30 74.12 (-1.08)	3.50 76.79 (0.03)	3.41 70.60 (-0.11)	3.45 76.76 (0.00)	3.74/3.92 61.75 (-0.09)

3.5 Comparison of the composition and structure of the K102 RU with other *Klebsiella* serotypes and potential cross-reactivity

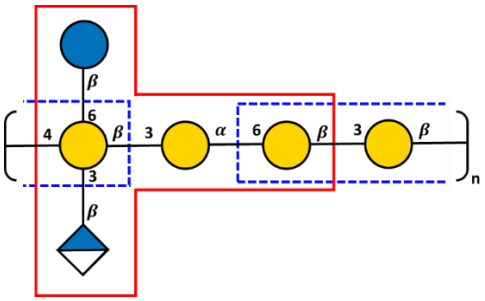
The RU structure of K102 CPS was compared to the CPS RU structures of other *K. pneumoniae* serotypes in published chemical structural studies using the compiled database. Through this comparison, it was confirmed that the RU structure of K102 CPS was novel among the *K. pneumoniae* K-antigens and that it possesses two general unique structural features. Firstly, compared to other K-antigens, K102 has an unusual hexasaccharide RU type that is comprised of a trisaccharide main chain, a disaccharide terminal, and a monosaccharide terminal, designated as 3 + 2 + 1 RU type. Additionally, the homopolysaccharide main chain that consists of only Gal sugar residues is another unusual structural feature compared to other *K. pneumoniae* K-antigens.

The K-antigen of K102 also shares structural features with some *K. pneumoniae* K-antigen serotypes. Firstly, it contains the same sugar composition (Gal; Glc; GlcA) as the K-antigens of serotypes K8; K11; K15; K22; K25; K27; K37; K51; and K82. Moreover, K102 CPS possesses similar epitopes (potential antibody-binding sites) to K-antigens of K15 and K18 as shown in **Figure 3.15**. These common epitopes may potentially provide serological cross-reactivity between the K-antigen of K102 and the K-antigens of K15 and K18, and these will be investigated in-depth by our collaborators, GVGH.

(a) K102



(b) K15



(c) K18

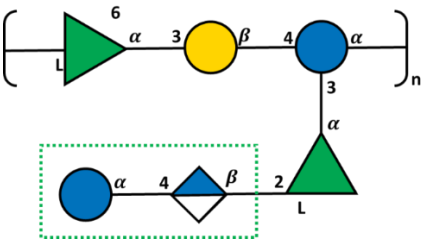


Figure 3.15: Capsular polysaccharide repeating units of *K. pneumoniae* (**a**) K102; (**b**) K15; and (**c**) K18, showing two structural features shared by K102 and K15: a pentasaccharide unit formed by the same types of monosaccharide residues with the same sequence (as shown by the red borders in (**a**) and (**b**)), and a trisaccharide unit formed by the same types of monosaccharide residues with the same sequence and linkage positions (as shown by the blue dashed borders in (**a**) and (**b**)), and a structural feature shared by K102 and K18: a disaccharide terminal branch formed by the same monosaccharide residues with the same linkages and sequence (as shown by the green dotted borders in (**a**) and (**c**)).

3.6 Conclusions

The CPS of K102 was successfully characterised, and the RU structure elucidated using NMR spectroscopic analysis. The NMR analysis agreed with the chemical analysis performed at the University of Trieste, however, it was demonstrated that NMR analysis was more reliable for accurately determining the polysaccharide structural features such as sugar composition, and linkages, additionally, it gave the monosaccharide sequence in the RU. Although it is not always possible for some polysaccharides, the NMR experimental chemical shifts data was able to confirm the D-absolute configurations for all sugar residues, apart from residue A (α-D-Glc*p*). This was based on the calculated glycosylation shifts and the configurational factors provided by Shashkov.[119] The proposed RU structure was found to be novel, having a unique 3 + 2 + 1 RU type among other *K. pneumoniae* capsular polysaccharides. This corroborates the identification of this strain, corresponding to K102, as a novel serotype through genotyping by our collaborators. Furthermore, K102 CPS possesses similar epitopes to the K-antigens of K15 and K18, which may potentially provide cross-reactivity. The established procedure for the NMR analysis of K102 CPS was implemented in the next two chapters to characterise and elucidate the structures of the 3 remaining *K. pneumoniae* K-antigens.

Chapter 4: Structural elucidation of K112 CPS

K. pneumoniae serotype K112 is among the emerging clinically significant strains of this bacterium that have been identified as novel using genotyping by our collaborators, this was determined as KL112. A capsular polysaccharide sample (~15 mg) of this strain was received from our collaborators at the GVGH Vaccines Institute for Global Health, Siena, for in-depth NMR analysis. Preliminary chemical composition analysis of the CPS was conducted by the GVGH group which revealed the presence of mannose (Man); galactose (Gal); and glucuronic acid (GlcA). In this chapter, the full NMR characterisation and structural elucidation of the RU structure of this K-antigen is reported and later compared to comprehensive chemical studies performed at the University of Trieste (Prof. Cescutti). The optimum temperature for conducting the NMR experiments on K112 CPS was 333 K.

4.1 1D ^{1}H, ^{13}C, and 2D/1D Homonuclear NMR experiments

As before, the DOSY experiment (**Figure 4.1b**) was used to remove the water signal in the anomeric region of the 1D ^{1}H NMR spectrum (**Figure 4.1a**) of K122 CPS to reveal five anomeric signals, suggesting a pentasaccharide RU. The signals were designated labels as H1 A - E, from the most downfield to the most up-field signal. The chemical shifts of the anomeric signals and coupling constants for residue A (5.38 ppm) and B (5.18 ppm) were characteristic of α-sugar residues, while those for residue C (4.71 ppm), D (4.63 ppm), and E (4.54 ppm) were characteristic of β-sugar residues. The absence of signals in the methyl region confirmed that no acetylation, 6-deoxy sugars, or pyruvate groups were present. The analysis of the 1D ^{13}C/^{13}C-DEPT spectra corroborated the 1D DOSY-^{1}H NMR spectrum and revealed five anomeric signals, four methylene group signals, which confirmed the presence of a GlcA sugar residue, and no methyl group signals. The assignments of the ring protons for each spin system and identification of their sugar types were accomplished from 2D ^{1}H-^{1}H COSY and DOSY-TOCSY spectra, using the proton anomeric signals as a starting point (**Figure 4.2**).

From H1 of residue A, 2D COSY gave a correlation to H2 at 4.00 ppm, and applying 2D DOSY-TOCSY gave further multiple-bond correlations to H3, overlapping with H2 at 4.00 ppm, and stopped with H4 at 4.23 ppm, which is a characteristic pattern of an α-Gal residue. It was determined that H3 overlaps with H2 because, for Gal sugars, H4 is always the most deshielded signal in TOCSY. Notably, from H1 of B, the 2D COSY experiment gave a correlation to H2 at 4.31 ppm, and DOSY-TOCSY did not give any further multiple-bond correlations.

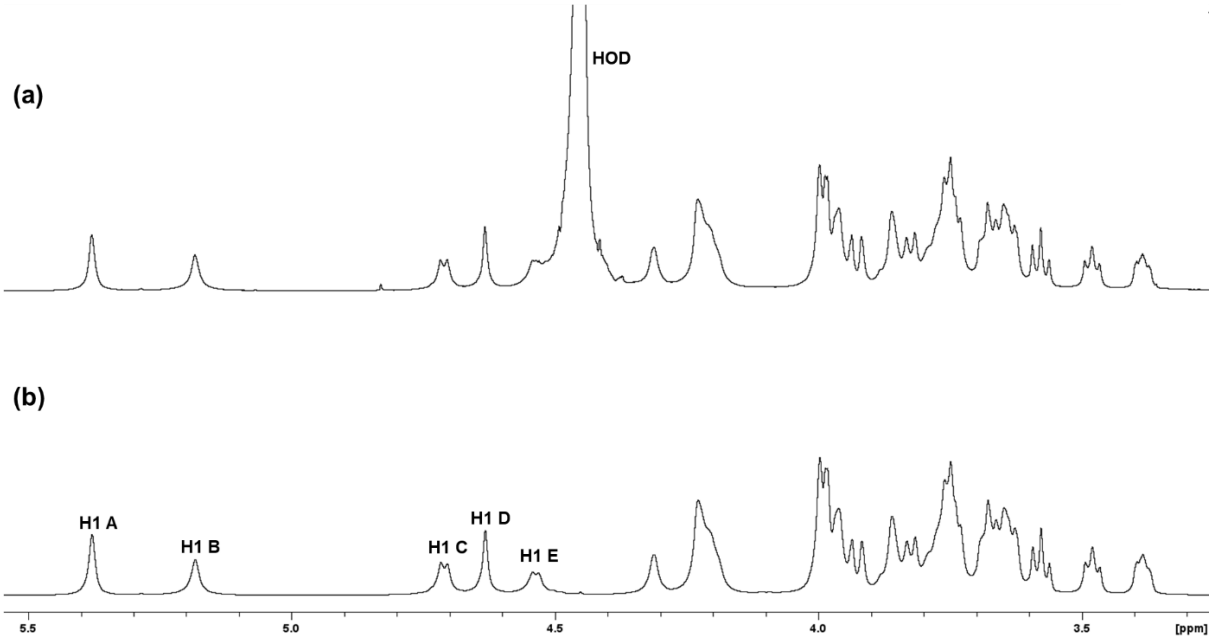


Figure 4.1: Overlay of **(a)** 1D ^{1}H NMR spectrum and **(b)** 1D ^{1}H-DOSY NMR spectrum of K112 CPS recorded at 318 K and 600 MHz with the diagnostic anomeric signals labelled from H1 A-E.

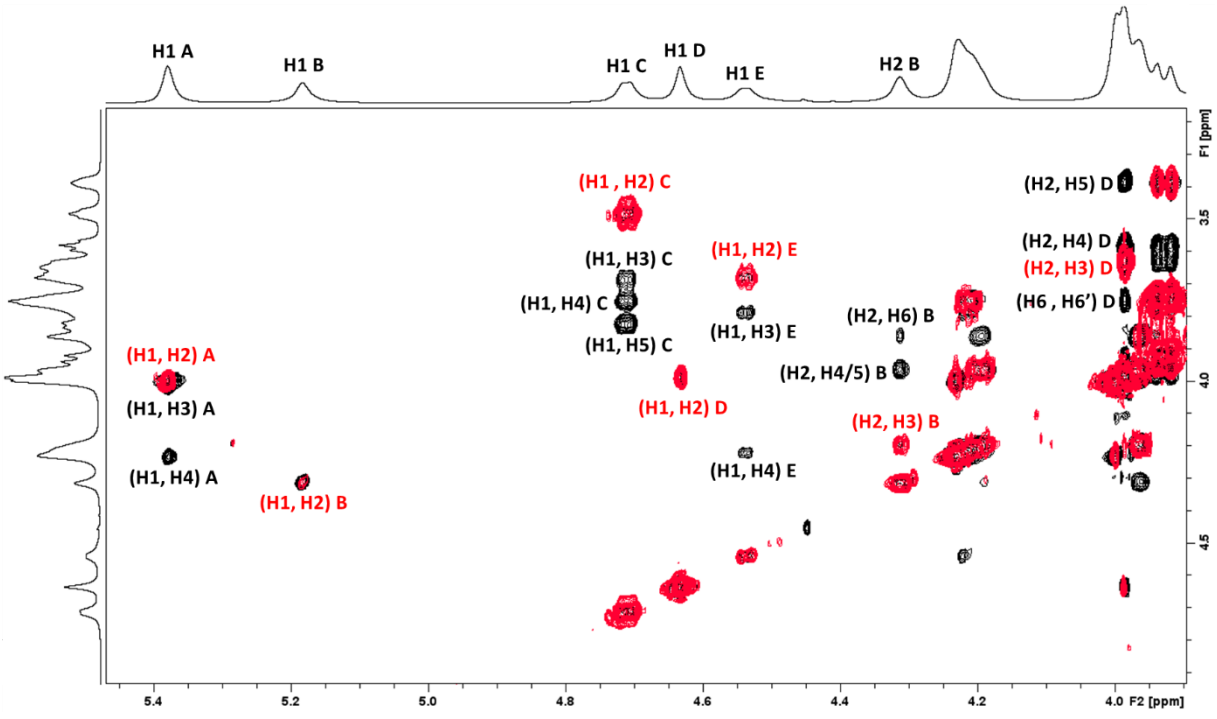


Figure 4.2: 2D ^{1}H-^{1}H COSY (red) overlaid upon DOSY-TOCSY (black) expansion spectra of K112 CPS with diagnostic signals labelled, recorded at 333 K and 600 MHz. A = α-Gal; B = α-Man; C = β-GlcA; D = β-Man; E = β-Gal.

This pattern was not observed in the TOCSY spectrum of the previously described K102 CPS and is characteristic of a mannose sugar type, due to their small $J_{H2, H3}$ coupling constant. However, from H2 of B, COSY gave a well-resolved correlation to H3 at 4.20 ppm, and DOSY-TOCSY gave further long-range correlations to H4 at 3.97 ppm, H5 at 3.96 ppm (overlapping with H4), and H6 at 3.86 ppm.

Therefore, residue B was identified as an α-Man sugar residue. From H1 of C, 2D COSY elucidated H2 at 3.48 ppm, and DOSY-TOCSY revealed H3, H4, and H5 (3.68; 3.75; and 3.82 ppm respectively), characteristic of a β-GlcA residue. 2D COSY and DOSY-TOCSY from H1 and H2 of residue D revealed the same pattern observed for residue B and elucidated H2 at 3.99 ppm, H3 at 3.63 ppm, H4 at 3.85 ppm; H5 at 3.38 ppm; and H6 at 3.93 ppm, characteristic of a β-Man sugar residue. Lastly, 2D COSY and DOSY-TOCSY from H1 of residue E gave the same pattern as for residue A and elucidated H2 at 3.68 ppm, H3 at 3.78 ppm; and H4 at 4.22 ppm, characteristic of a β-Gal residue. These assignments were confirmed by a series of 1D TOCSY experiments recorded by selectively irradiating H1 of each of the Gal sugars (A and E) and the GlcA sugar (C), and from H2/H5 for each of the Man sugars (B and D).

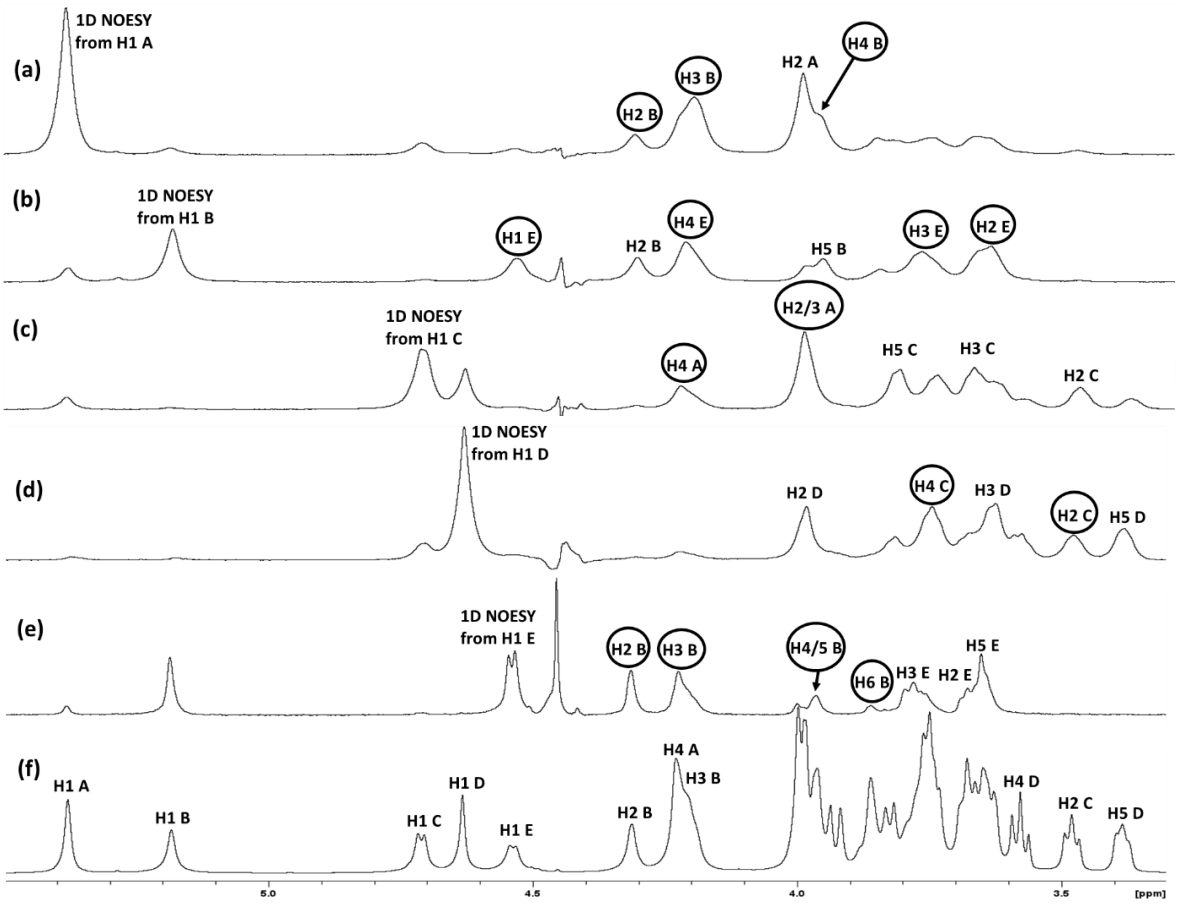


Figure 4.3: 1D NOESY spectra of K112 CPS recorded at 300 ms by irradiating H1 of residues A (α-Gal) **(a)**; B (α-Man) **(b)**; C (β-GlcA) **(c)**; D (β-Man) **(d)**; and E (β-Gal) **(e)**, overlaid onto 1D DOSY spectrum **(f)**. The labels of the anomeric signals on each NOESY spectrum are provided and the assignments of the respective intra- and inter-residue correlations given with inter-residue correlations highlighted within ovals, and the assignments are cross validated on the 1D DOSY spectrum **(f)**.

More information was obtained from a series of 1D NOESY spectra recorded by irradiating H1 of each residue (**Figure 4.3a-f**). The 1D NOESY profiles confirmed some of the preliminary assignments, and particularly, the 1D NOESY from H1 of E (β-Gal) gave the expected intra-residue correlations for a β-pyranose sugar (H1 to H3 and H1 to H5), thus, elucidating H5 of E at 3.65 ppm (**Figure 4.3e**). Additionally, the series of 1D NOESY spectra elucidated the sequence of the residues in the RU by providing inter-residue correlations from H1 of A (α-Gal) to multiple protons of B (α-Man); H1 of B (α-Man) to multiple protons of E (β-Gal); H1 of C (β-GlcA) to multiple protons of A (α-Gal); H1 of D (β-Man) to multiple protons of C (β-GlcA); and H1 of E (β-Gal) to multiple protons of B (α-Man). This suggests a repeating unit structure with the sequence:

B (α-Man) → E (β-Gal) as the main chain; and **D (β-Man) → C (β-GlcA) → A (α-Gal) → B (α-Man**) as the side chain.

However, the NOESY spectra could not be used to assign the linkage positions and thus these, and the remaining assignments (H5 and H6 of A (α-Gal), and H6 of E (β-Gal)) were assigned using the 2D ^{1}H-^{13}C heteronuclear experiments.

4.2 2D ^{1}H-^{13}C NMR experiments

With nearly all the proton assignments of the spin systems obtained from the homonuclear experiments (except for H5 and H6 of A, and H6 of E), the 2D ^{1}H-^{13}C heteronuclear experiments were acquired to assign the ^{13}C chemical shifts, confirm the sequence, and elucidate the linkage positions for each residue. Beginning with the 2D HSQC-DEPT spectrum, the cross-peaks were assigned using the proton assignments and confirmed by overlaying the 1D TOCSY profile of each residue. Additionally, the 2D HSQC-TOCSY confirmed the HSQC-DEPT assignments and resolved the ambiguities which arose from overlapping proton signals (**Figure 4.4**). Furthermore, the 2D HSQC-NOESY provided further confirmation and elucidated the linkage positions for residues A (α-Gal); B (α-Man); and C (β-GlcA), through inter-residue correlations from H1 A (α-Gal) to C3 B (α-Man); H1 C (β-GlcA) to C3 A; and H1 D (β-Man) to C4 C (β-GlcA), and thus, elucidated the side chain as:

D (β-Man) → 4C (β-GlcA) → 3A (α-Gal) → 3B (α-Man).

At this point, apart from the methylene cross-peaks for H6-C6 of A and E (appearing as inverted peaks in red in **Figure 4.4**), the cross-peak for H5-C5 of A could be assigned as it was the only remaining cross-peak in the ring region (**Figure 4.4**).

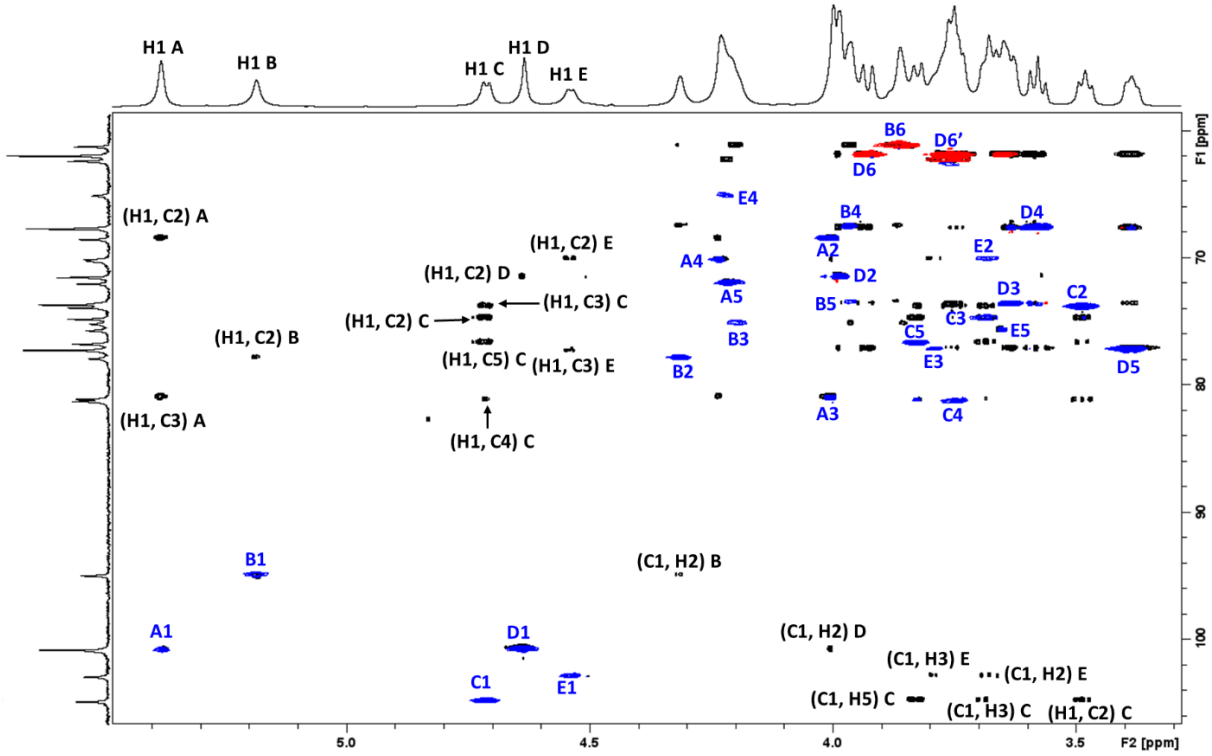


Figure 4.4: Labelled 2D ^{1}H-^{13}C HSQC-DEPT spectrum (blue and methylene peaks, red) overlaid onto 2D HSQC-TOCSY spectrum (black). A = α-Gal; B = α-Man; C = β-GlcA; D = β-Man; E = β-Gal.

Finally, the 2D HMBC spectrum was assigned, giving intra-residue correlations from H6 to C5 of A and a corresponding correlation from C5 to H6 of A (α-Gal) in the ring region, confirming the assignments of H5-C5 and H6-C6 of A (α-Gal), consequently confirming the assignment of H6-C6 of E (β-Gal) as it was the remaining methylene cross-peak in the HSQC-DEPT. The HMBC spectrum also gave folded correlation cross-peaks from H5 and H4 of C (α-Gal) to the COOH signal, confirming that residue C is the β-GlcA residue. In the anomeric region, HMBC gave the diagnostic intra-residue correlations which provided further confirmation for some of the preliminary assignments: from H1 of A to C3, to C5, and C2 (with a low intensity); from H1 of B to C2, to C3, and C5; from H1 of D to C2, and C5, and no intra-residue correlations from H1 of C and E (**Figure 4.5**). Furthermore, HMBC gave the key inter-residue correlations: from H1 A (α-Gal) to C3 B (α-Man); H1 B (α-Man) to C3 E (β-Gal); H1 C (β-GlcA) to C3 A (α-Gal); H1 D (β-Man) to C4 C (β-GlcA); and H1 E (β-Gal) to C2 B (α-Man) (**Figure 4.5**), confirming the sequence of the residues, elucidating the linkage positions, and thus confirming the repeating unit structure shown in **Figure 4.6**.

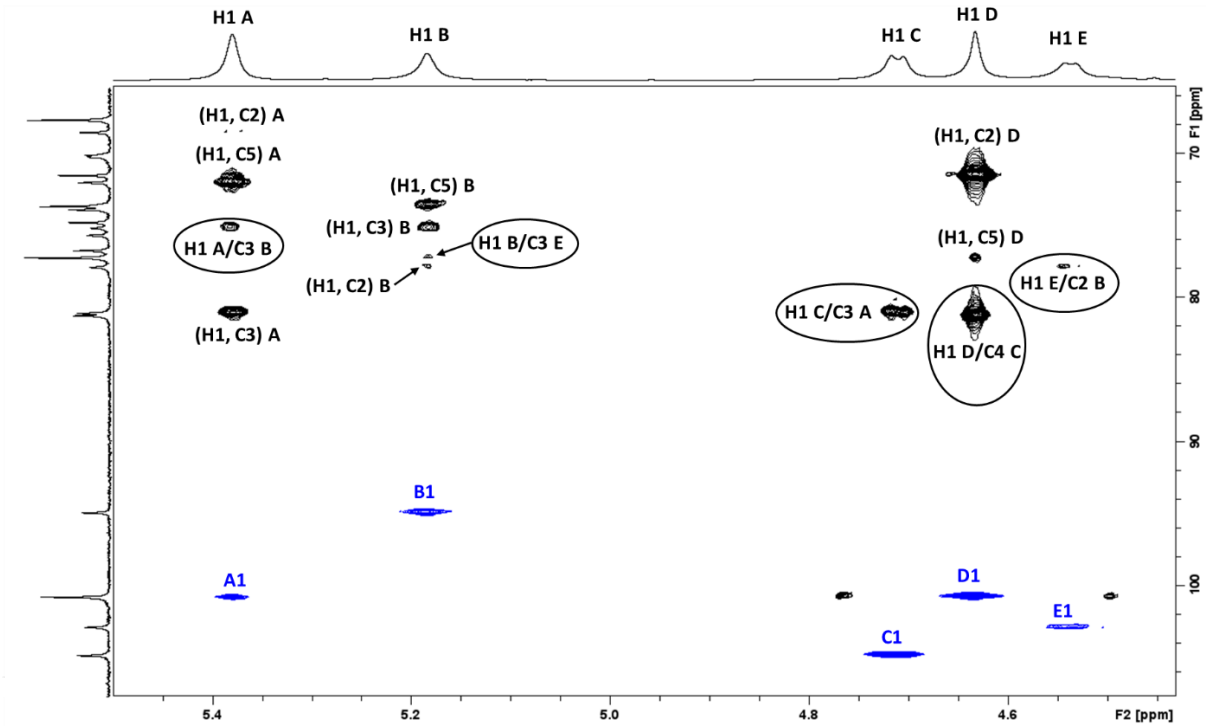


Figure 4.5: Expansion of the anomeric region for an overlay of 2D ^{1}H-^{13}C HMBC spectrum (black) and HSQC-DEPT (blue) of K112 CPS with the diagnostic intra- and inter-residue correlations labelled. The inter-residue correlations are highlighted by the ovals. A = α-Gal; B = α-Man; C = β-GlcA; D = β-Man; E = β-Gal.

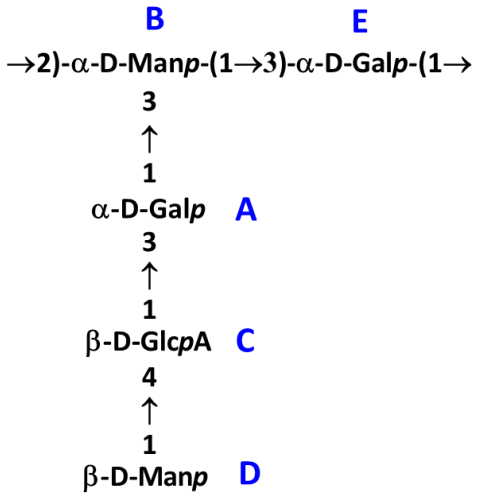


Figure 4.6: The proposed repeating unit structure of *K. pneumoniae* K112 capsular polysaccharide.

4.3 Summary of the full assignments, presentation of diagnostic spectra, and comparison of RU structural features to chemical analysis

The 1D DOSY-^{1}H NMR spectrum of K112 CPS gave five anomeric signals, consistent with a pentasaccharide RU. This was corroborated by 1D ^{13}C/^{13}C-DEPT NMR spectra that also confirmed the presence of a GlcA residue by revealing four methylene group signals, in agreement with the provided chemical compositional analysis. Using 2D COSY and TOCSY spectra, further assignments of the spin systems were elucidated dependent on the sugar residue type. The TOCSY patterns were consistent with A = α-Gal; B = β-Man; C = β-GlcA; D = β-Man; and E = β-Gal. A series of 1D NOESY spectra from H1 of each residue elucidated the sequence of the residues in the RU, and the assignment for H5 of E was elucidated by 1D NOESY from H1 of E, leaving only H5 and H6 of A, and H6 of E unassigned.

Using the obtained proton assignments of the spin systems, the 2D HSQC-DEPT spectrum was successfully assigned with the aid of 2D heteronuclear-hybrid experiments, such as the HSQC-TOCSY and HSQC-NOESY, and the 2D HMBC experiment. Notably, the HMBC experiment allowed the elucidation of the assignments for H5-C5 of A (α-Gal); and H6-C6 of A (α-Gal) and E (β-Gal), which could not be elucidated from preliminary experiments, and ultimately, HMBC confirmed the sequence of sugar residues in the RU and clarified the linkage positions. Therefore, the 2D HSQC-DEPT spectrum was fully assigned and labelled as shown in **Figure 4.7a**, and consequently, the 1D ^{13}C NMR spectrum was also labelled as shown in **Figure 4.7b**. The obtained ^{1}H and ^{13}C chemical shifts for each residue were recorded as shown in **Table 4.1** and the glycosylation shifts were calculated. As expected, the obtained glycosylation shifts were large and positive at the linkage positions, thus confirming the glycosylation pattern. However, for K112 K-antigen, the D-absolute configurations of the residues could not be confirmed from the values of the glycosylation shifts, as no sequence met the rules described by Shashkov.[119]

As for K102 CPS, structural features of K112 CPS RU acquired from NMR analysis were corroborated by chemical analysis performed at the University of Trieste. Composition analysis involved methanolysis of the CPS followed by trimethylsilylation derivatisation and GLC-MS analysis, which showed the presence of mannose (Man), galactose (Gal), and glucuronic acid (GlcA) in the molar ratio of 1.5:1.0:0.4. Glycosidic linkage analysis was performed by GLC-MS analysis of the partially methylated alditol acetates which indicated: terminal Man*p*; 3-Gal*p*; 2,3-Man*p* in the molar ratio of 0.9:1.0:0.7. These results were reasonably in agreement with the sugar composition and glycosylation pattern shown by NMR. However, the techniques gave a low recovery of Gal in sugar composition and linkage analysis because of the linkage to GlcA. Additionally, the methodology for linkage analysis does

not detect uronic acids such as GlcA and is not quantitative, nevertheless, the data was consistent with the pentasaccharide repeating unit elucidated by NMR analysis.

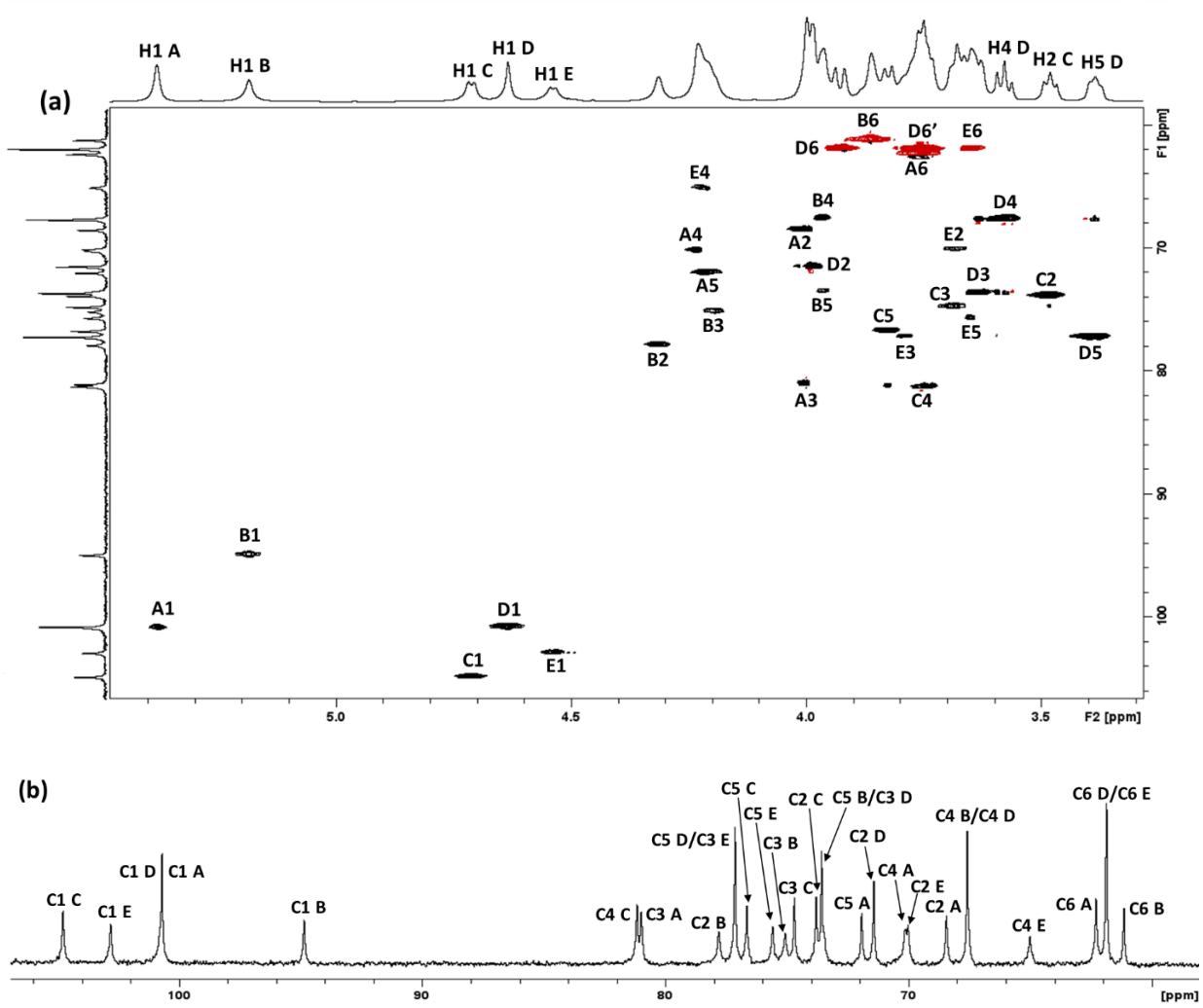


Figure 4.7: Fully labelled **(a)** 2D ^{1}H-^{13}C HSQC-DEPT identity map and **(b)** 1D ^{13}C NMR spectra of K112 CPS recorded at 333 K, and at 400 and 150 MHz, respectively. A = α-Gal; B = α-Man; C = β-GlcA; D = β-Man; E = β-Gal.

Table 4.1: Complete ^{1}H-^{13}C NMR chemical shifts data of the sugar residues in K112 CPS with the glycosylation shifts shown in brackets; the carbon chemical shifts of the linkage position are underlined.

Sugar residues	H1 C1	H2 C2	H3 C3	H4 C4	H5 C5	H6 C6
A **→3)-α-D-Gal*p*-(1→**	5.38 100.71 (7.53)	4.00 68.45 (-0.90)	4.00 <u>80.99</u> (10.86)	4.23 70.15 (-0.13)	4.22 71.93 (-0.63)	3.77 62.30 (0.26)
B **→2,3)-α-D-Man*p*-(1→**	5.18 94.87 (-0.07)	4.31 <u>77.80</u> (6.11)	4.20 <u>75.08</u> (3.83)	3.97 67.59 (-0.35)	3.96 73.58 (0.24)	3.86 61.15 (-0.84)
C **→4)-β-D-Glc*p*A-(1→**	4.71 104.79 (8.02)	3.48 73.80 (-1.20)	3.68 74.69 (-1.84)	3.75 <u>81.17</u> (8.48)	3.82 76.66 (-0.27)	- 175.91 (-0.56)
D **β-D-Man*p*-(1→**	4.63 100.72 (6.17)	3.99 71.43 (-0.70)	3.63 73.60 (-0.43)	3.58 67.59 (-0.10)	3.38 77.14 (0.14)	3.93; 3.75 61.86 (-0.13)
E **→3)-β-D-Gal*p*-(1→**	4.54 102.83 (5.46)	3.68 70.04 (-2.92)	3.78 <u>77.14</u> (3.36)	4.22 65.02 (-4.67)	3.65 75.59 (-0.34)	3.65 61.86 (0.02)

4.4 Comparison of the composition and structure of the K112 RU with other *Klebsiella* serotypes and potential cross-reactivity

Comparison of the proposed repeating unit structure of K112 CPS (**Figure 4.8**) with other *K. pneumoniae* K-antigens revealed that it was a novel structure. This repeating unit comprised of a disaccharide main chain and a trisaccharide terminal, designated as 2 + 3 RU type, is unusual and the first among currently known *K. pneumoniae* K-antigens with pentasaccharide repeating units. Nevertheless, it was found that K112 has some common structural features with various *K. pneumoniae* K-antigens.

Its monosaccharide composition (Man; Gal; GlcA) is similar to that of serotypes K3, K20, K21, K43, K66, and K74. Interestingly, it was found that the RU structure of K112 CPS is similar to that of K20 with respect to every structural feature apart from K20 comprising a disaccharide side chain lacking the terminal β-Man*p* residue (**Figure 4.9**). This suggests that K112 may be closely related and serologically cross-reactive with K20. Furthermore, the GlcA → Gal disaccharide unit that forms part of the terminal branch of K112 CPS RU is similar to the terminal branches of the K-antigens of *K. pneumoniae* serotypes K24, K31, and K43 (**Figure 4.10a-c**). As previously described for K102 CPS (**Section 3.5**), such similar epitopes may potentially provide some cross-reactivity.

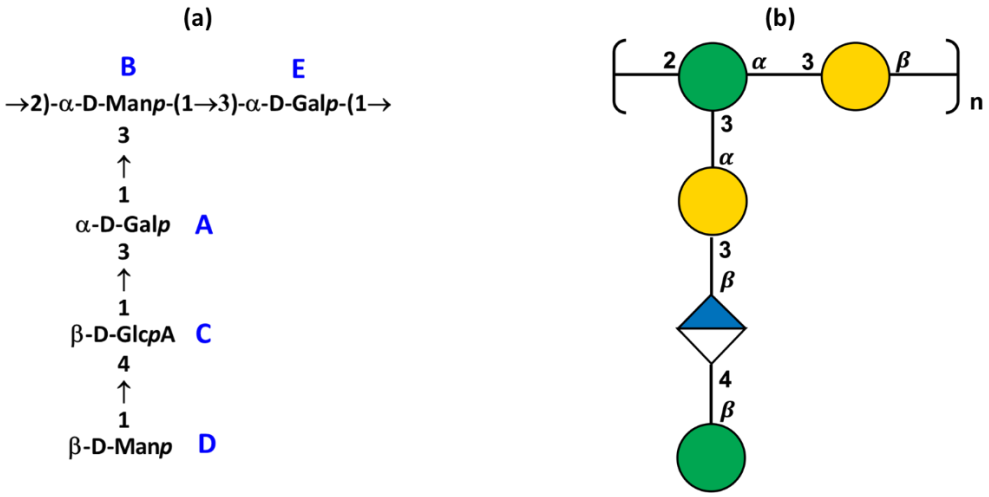


Figure 4.8: Proposed repeating unit structure of the capsular polysaccharide of K112 with sugar residues represented in the **(a)** IUPAC format and the **(b)** SNFG format.

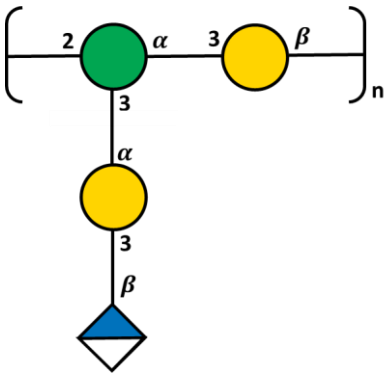
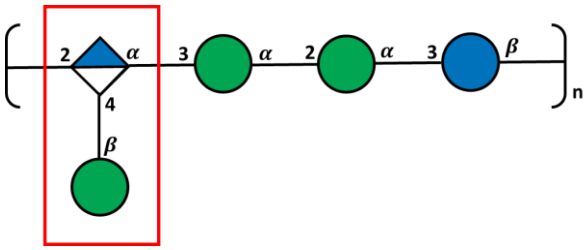
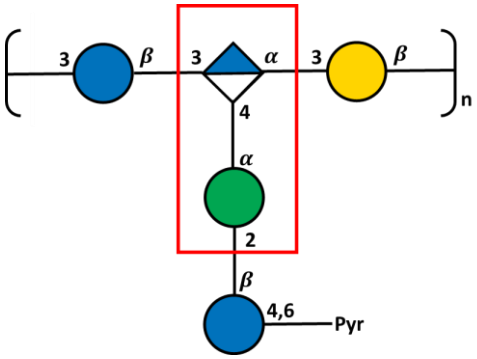


Figure 4.9: The capsular polysaccharide repeating unit of *K. pneumoniae* serotype K20 having the same structure as serotype K112 besides lacking the terminal β-Manp residue.

(a) K24



(b) K31



(c) K43

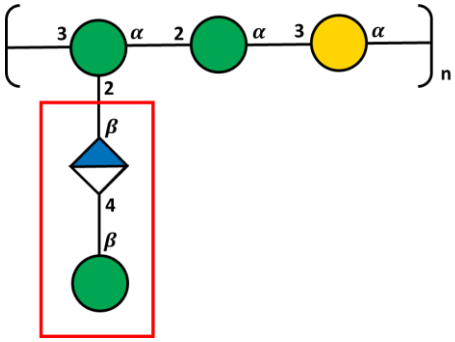


Figure 4.10: *K. pneumoniae* K-antigens of **(a)** K24; **(b)** K31; and **(c)** K43, having the same Man → 4GlcA disaccharide unit as the K-antigen of K112 (highlighted by the red borders), an epitope that may potentially provide cross reactivity between K112 and these K-antigens.

4.5 Conclusions

In conclusion, the NMR experimental data demonstrated that the K-antigen of K112 is made up of a novel repeating unit structure, with a RU type designated as a 2 + 3 (**Figure 4.8**). This confirms the identification of K112 as a novel serotype through genotyping by collaborators. Chemical composition analysis and linkage analysis performed at the University of Trieste corroborated the structural features elucidated by NMR. However, the chemical analysis demonstrated some limitations in accurately determining the molar ratio of sugars, due to the presence of GlcA, and could not identify the linkages of GlcA residues, thus demonstrating the advantages and the power of NMR spectroscopy for elucidation of CPS RU structural features.

Interestingly, although K112 CPS was found to be a novel among *K. pneumoniae* K-antigens, it appeared that it only differs with the addition of the terminal β-Man residue forming part of the side chain, in comparison to the repeating unit structure of *K. pneumoniae* serotype K20. This suggested that these serotypes may be closely related to each other and most likely to be serologically cross-reactive. Additionally, K112 CPS contains some key epitopes present in *K. pneumoniae* K-antigens of serotypes K21; K31; and K43, which may potentially provide cross-reactivity. As previously noted, the cross-reactivity studies will be performed by collaborators, GVGH.

Chapter 5: Structural characterisation of K122 and KL107

The capsular polysaccharides of *K. pneumoniae* serotype K122 and KL107 were the final polysaccharides to be characterised, using NMR spectroscopy, during this investigation. These were also genotypically identified as novel strains by our collaborators; however, it was later found in the investigation that KL107 is not novel, hence it is being referred to with its KL designation. The CPSs of these two strains were also received from GVGH for in-depth NMR analysis. Preliminary chemical analysis performed by the GVGH group demonstrated that they both have identical sugar compositions: glucose (Glc), mannose (Man), and glucuronic acid (GlcA) and are thus reported together in this chapter. As with the previous capsular polysaccharides, a detailed chemical analysis was performed at the University of Trieste, and compared to the structural results obtained from NMR. The NMR experiments were recorded at 343 K, which appeared to give high-quality spectra for these samples, using the 600 MHz instrument. Since the chemical composition of these two K-antigens was the same and comprised the sugar residues previously encountered in the structural elucidation of K102 and K112 CPS repeating units, in-depth explanations of the NMR experiments will not be detailed here.

5.1 K122 CPS characterisation

5.1.1 1D ^{1}H, ^{13}C, and 2D/1D homonuclear experiments

For the K122 CPS, the 1D DOSY-^{1}H NMR experiment (**Figure 5.1**) revealed six anomeric signals (5.33; 5.32; 5.12; 4.93; 4.65; and 4.46 ppm), designated as H1 A - F from the most down-field to the most up-field signal, respectively, with H1 A and H1 B (5.33 and 5.32 ppm) signals overlapping, thus suggesting a hexasaccharide RU. The chemical shifts and coupling constants of the anomeric signals suggested the presence of three α-sugars (residues A - C) and three β-sugars (residues D - F). 1D ^{13} C and ^{13}C-DEPT experiments (**Figure 5.2**) confirmed the proton analysis by showing six anomeric carbon signals, three methylene group signals and three COOH group signals, suggesting the presence of three uronic acid residues, and revealed two diagnostic up-field signals (near 67 ppm) of C4 of mannose residues.

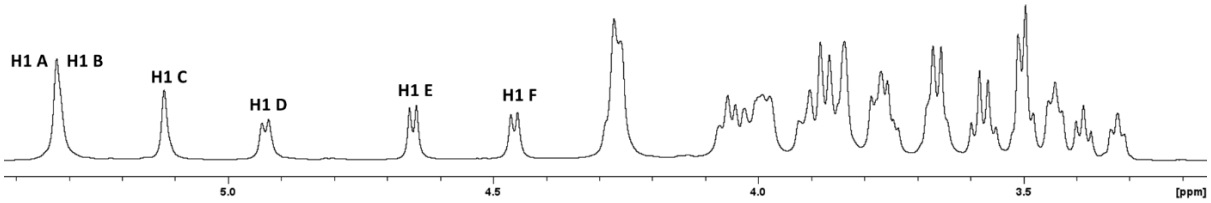


Figure 5.1: 1D DOSY-^{1}H NMR spectrum of K122 CPS recorded at 343 K and 600 MHz, showing six diagnostic anomeric signals labelled from H1 A - F.

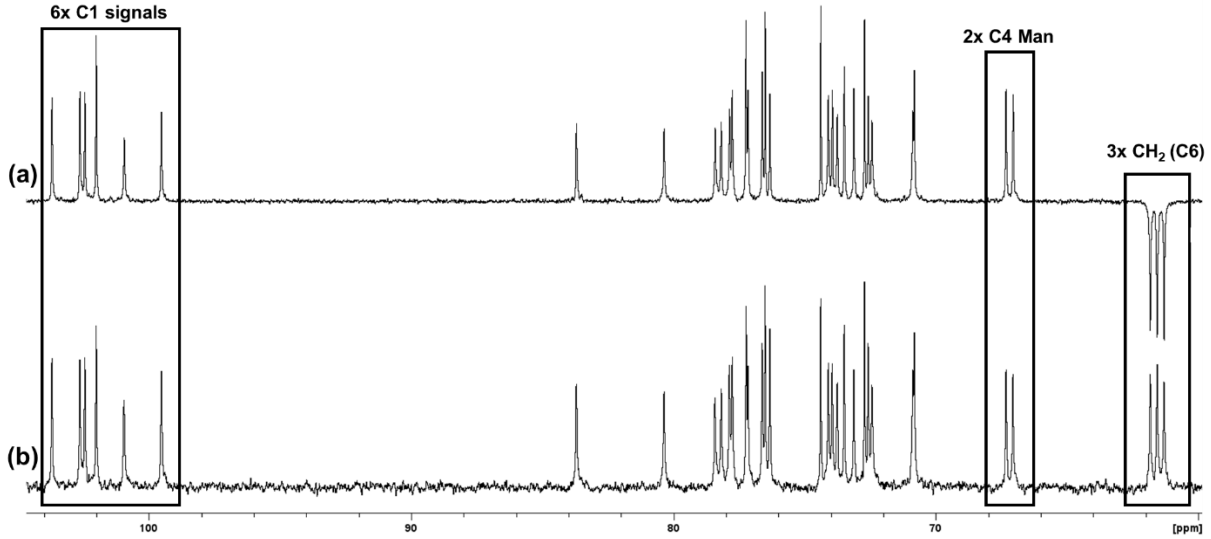


Figure 5.2: Overlay of **(a)** 1D ^{13}C-DEPT and **(b)** ^{13}C NMR spectra of K122 CPS recorded at 343 K and 150 MHz, showing six anomeric signals, three methylene (CH_2) as inverted signals, and two C4 mannose signals. 1D ^{13}C gave three carbonyl signals downfield (not shown) and were removed by the ^{13}C-DEPT experiment.

As previously described, 2D proton scalar correlation homonuclear experiments (COSY and TOCSY) were used to assign ring proton signals and identify each sugar residue (**Figure 5.3**). 2D COSY gave correlations from H1 of residue A to H2, and H2 to H3 (with low intensity), 2D TOCSY revealed no further correlations for residue A but confirmed the assignment of H3 with a correlation from H1 to H3 (with low intensity), characteristic of an α-Man sugar residue. As a result of the overlapping anomeric signals of residues A and B, only a correlation from H1 to H2 was identified for residue B in the 2D COSY and TOCSY experiments, which prevented the identification of residue B in these experiments. As for residue A, 2D COSY of residue C showed correlations from H1 to H2, and H2 to H3 (with low intensity). 2D TOCSY gave a correlation from H1 to H3 (with low intensity) for residue C, which was characteristic of a second α-Man sugar residue. Notably, the H2 proton signals of residues A and C overlapped at 4.27 and 4.26 ppm, respectively, this resulted in obstruction of the assignment of further TOCSY correlations from these signals, which would have elucidated the rest of the spin

systems for these mannose residues. For residue D, 2D COSY and TOCSY experiments revealed all proton assignments (H1 to H6), characteristic of a β-Glc sugar residue; and revealed H1 to H5 for residues E and F, characteristic of two β-GlcA sugars, respectively. Therefore, the analysis of the 2D COSY and TOCSY spectra revealed the presence of two mannose residues (A and C), consistent with the analysis of the 1D ^{13}C NMR spectrum (i.e., the two C4 signals characteristic of mannose), one glucose residue (D), and two glucuronic acid residues (E and F). However, 1D ^{13}C NMR analysis suggested the presence of three GlcA residues, which meant that the unidentified spin system of residue B must be a third GlcA residue. This was confirmed by subsequent experiments (**Section 5.1.2**).

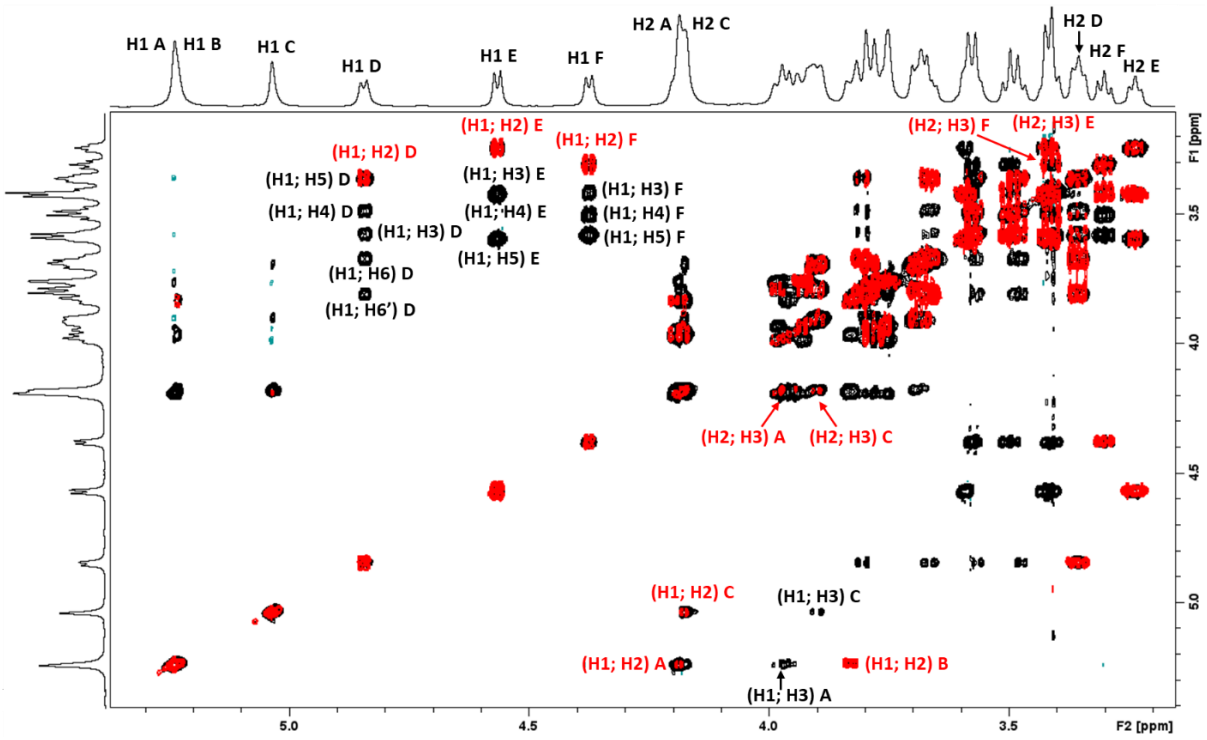


Figure 5.3: 2D ^{1}H-^{1}H COSY (red) overlaid onto 2D TOCSY (black) of K122 CPS recorded at 343 K and 600 MHz, with labels of the diagnostic signals given (A = α-Man^{I}; B = α-GlcA; C = α-Man^{II}; D = β-Glc; E = β-GlcA^{I}; F = β-GlcA^{II}). The cross-peaks of A-C are labelled along F2 (horizontal axis).

A series of 1D TOCSY experiments recorded by irradiating the anomeric signals and other diagnostic signals of each residue confirmed the assignments from 2D COSY and 2D TOCSY and were later used to aid the assignment of ^{1}H-^{13}C experiments. Additionally, 2D NOESY and 1D NOESY (from H1 of each residue) experiments were recorded and analysed, however, none of the remaining protons (H4 to H6 of residues A and C, and H3 to H5 of residue B) could be assigned from these experiments. Furthermore, due to several unknown proton assignments at this point, the NOESY inter-residue correlations could not be completely assigned immediately. Nevertheless, the following inter-residue correlations were obtained: H1 A (α-Man^{I}) to H3 D (β-Glc); H1 B (α-GlcA) to H3 C (α-Man^{II}); and H1 C

(α-Man^{II}) to H3 A (α-Man^{I}), suggestive of the following sequential arrangements of the residues in the RU:

A $\rightarrow$ 3-D; B $\rightarrow$ 3-C; C $\rightarrow$ 3-A

α-Man^{I} $\rightarrow$ 3-β-Glc; α-GlcA $\rightarrow$ 3-α-Man^{II}; α-Man^{II} $\rightarrow$ 3-α-Man^{I}

The remaining assignments and repeating unit structural features were obtained from the 2D ^{1}H-^{13}C heteronuclear experiments.

5.1.2 2D ^{1}H-^{13}C heteronuclear experiments

The 2D HSQC-DEPT spectrum was assigned based on the preliminary proton assignments, and 2D HSQC-TOCSY and HSQC-NOESY experiments which were performed to confirm the assignments and resolve ambiguities arising from overlapping proton signals. However, even the HSQC-hybrid experiments failed to elucidate the assignments of H4 to H6 for residues A and C; and H3 to H5 for residue B, and their corresponding carbons. These assignments were obtained from 2D ^{1}H-^{13}C HMBC experiment, which revealed the diagnostic intra-residue correlations in the anomeric region: H1 to C2, H1 to C3, and H1 to C5 for residues A (α-Man^{I}) and C (α-Man^{II}); thus elucidating H5-C5 of these residues (**Figure 5.4**). Moreover, key intra-residue correlations from C3 to H4 of residues A and C were obtained in the ring region of HMBC (**Figure 5.5**), which gave the assignment of H4-C4 of these β-Man residues, thus leaving only H6-C6 as their remaining assignments. Importantly, from the ring region of the HMBC spectrum, the following intra-residue correlations were obtained for residue B: C2 to H3; H3 to C4; H4 to C5; and a correlation from the folded signal of C6 (COOH) to H4 and H5, thus revealing the full H-C assignments of residue B and confirming it as an α-GlcA sugar. Similarly, HMBC gave correlations from two other folded C6 (COOH) signals to the H4 and H5 signals of residues E and F, thus confirming these residues as β-GlcA^{I} and β-GlcA^{II} sugars, respectively (**Figure 5.5**).

At this point, the only remaining assignments were H6-C6 of residues A (α-Man^{I}) and C (α-Man^{II}), which were assigned from the HSQC-TOCSY experiment, and gave correlations from H3 to C6, and H5 to C6 for residue A; and H5 to C6 for residue C. Therefore, the HSQC-DEPT spectrum was fully assigned, as shown in **Figure 5.4** and **Figure 5.5**. Lastly, the key HMBC inter-residue correlations were assigned to identify the glycosylation pattern of the residues as shown in **Figure 5.4**: H1 A (α-Man^{I}) to C3 D (β-Glc); H1 B (α-GlcA) to C3 C (α-Man^{II}); H1 C (α-Man^{II}) to C3 of A (α-Man^{I}); H1 of D (β-Glc) to C3 of B (α-GlcA); H1 of E (β-GlcA^{I}) to C4 of B (α-GlcA); and H1 of F (β-GlcA^{II}) to C2 of A (α-Man^{I}). These assignments were corroborated by further inter-residue correlations from the ring protons (H2 and H3 of A, H3 and H4 of B, H3 of C, and H3 of D) to C1 of the preceding residues, therefore, the repeating unit structure shown in **Figure 5.6** was obtained.

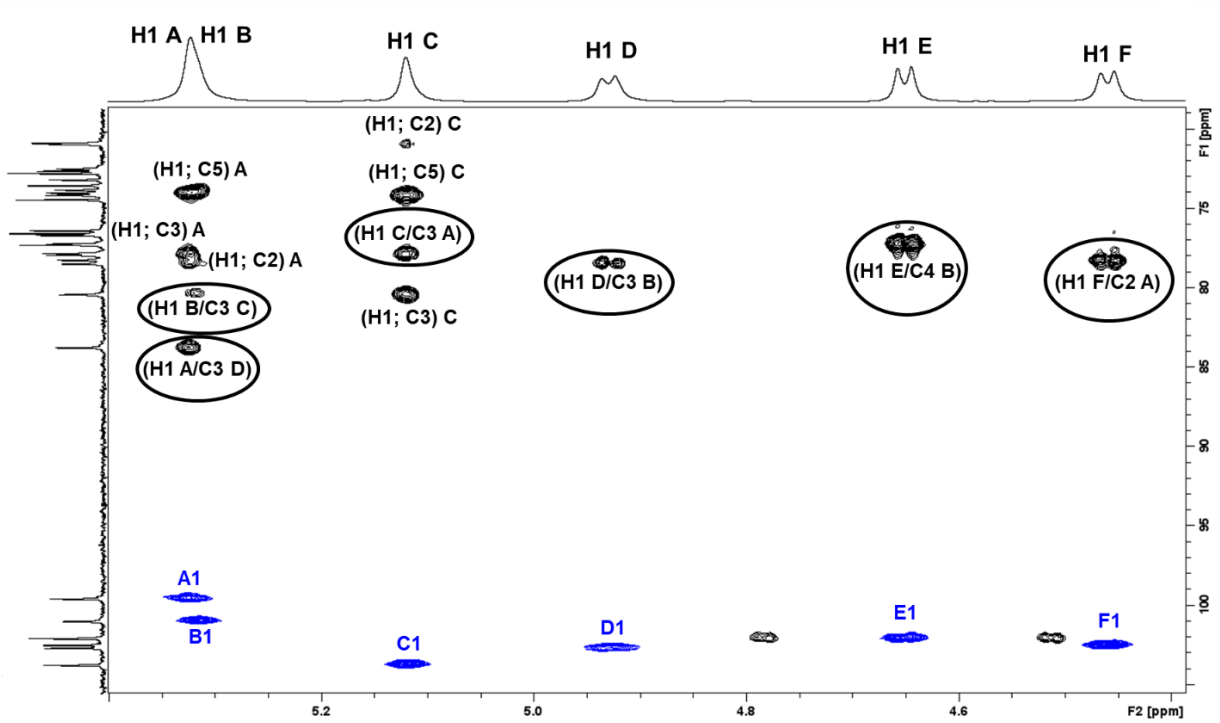


Figure 5.4: Fully assigned expansion of the anomeric region of 2D ^{1}H-^{13}C HMBC (black) overlaid with HSQC-DEPT (blue) of K122 CPS recorded at 343 K and 600 MHz. The key inter-residue correlations providing the glycosylation pattern are shown within ovals. A = α-Man; B = α-GlcA; C = α-Man; D = β-Glc; E = β-GlcA^{I}; F = β-GlcA^{II}.

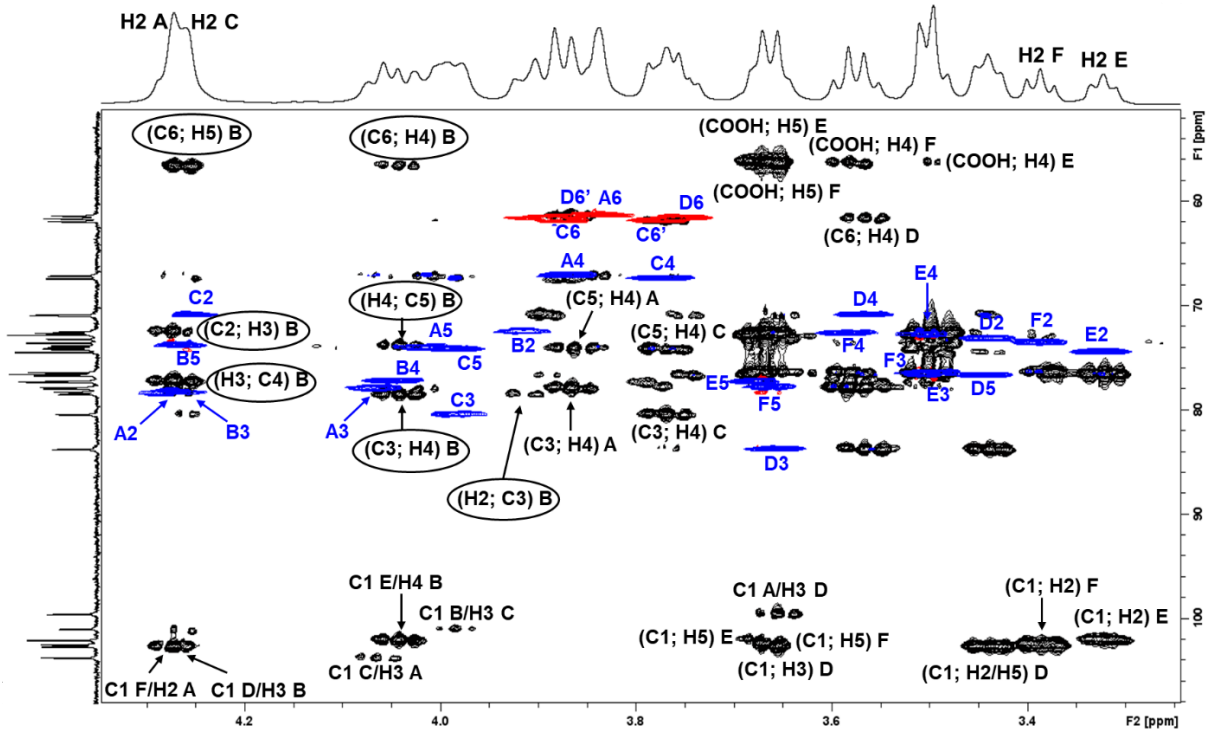


Figure 5.5: Labelled expansion of 2D ^{1}H-^{13}C HMBC spectrum (black) overlaid with HSQC-DEPT (blue, methylene cross-peaks inverted (red)) of K122 CPS recorded at 343 K and 600 MHz. The HMBC assignments which allowed the complete elucidation of ^{1}H-^{13}C signals of residue B are shown within ovals. A = α-Man; B = α-GlcA; C = α-Man; D = β-Glc; E = β-GlcA^{I}; F = β-GlcA^{II}.

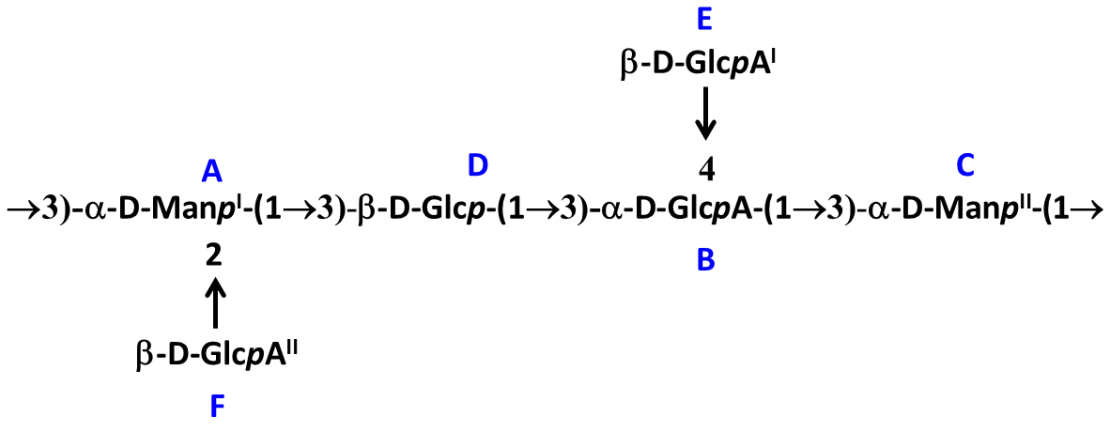


Figure 5.6: Proposed repeating unit structure of *K. pneumoniae* K122 capsular polysaccharide.

5.1.3 Summary of the full assignments and key diagnostic spectra, and comparison to chemical analysis results

The 1D proton (DOSY) and carbon NMR spectra of K122 CPS gave six anomeric signals, suggesting a hexasaccharide RU. 1D carbon showed three carboxyl signals and the DEPT experiment revealed three inverted methylene signals, thus indicating the presence of three GlcA residues and three hexoaldose residues in the hexasaccharide RU. Using the combination of 1D and 2D homonuclear experiments (COSY, TOCSY, and NOESY) and the anomeric protons as a starting point, assignments for some of the ring protons and the identities of the sugar residues were determined: A (α-Man^{I}); B (α-GlcA); C (α-Man^{II}); D (β-Glc); E (β-GlcA$_{I}$); and F (β-GlcA^{II}). However, there was extensive overlap of several diagnostic proton signals, such as H1 of residues A (α-Man^{I}) and B (α-GlcA); and H2 of residues A (α-Man^{I}) and C (β-Man^{II}), which prevented the complete elucidation of the proton assignments of these spin systems from the TOCSY experiments. Nevertheless, the remaining proton assignments of these residues were successfully obtained from 2D HSQC-TOCSY and HMBC experiments, and HMBC elucidated the complete glycosylation pattern of the sugar residues in the RU. Finally, the 2D ^{1}H-^{13}C HSQC-DEPT identity map and the diagnostic 1D ^{13}C NMR spectra were fully assigned as shown in **Figure 5.7a, b.**

The acquired ^{1}H and ^{13}C chemical shifts of the residues of the K122 CPS repeating unit are given in **Table 5.1**, and the calculated glycosylation shifts of the residues are shown in brackets below the carbon chemical shifts. The linkage positions gave positive and large glycosylation shift values (+4.2 to + 9.2 ppm), thus confirming the identified glycosylation pattern and the proposed repeating unit structure shown in **Figure 5.8**. However, the expected D-absolute configurations of the residues in this repeating unit could not be confirmed using the NMR data and glycosylation shift values. This is

because none of the disaccharide fragments within the unit matched the conditions of the configurational factors reported in the literature.[119,120]

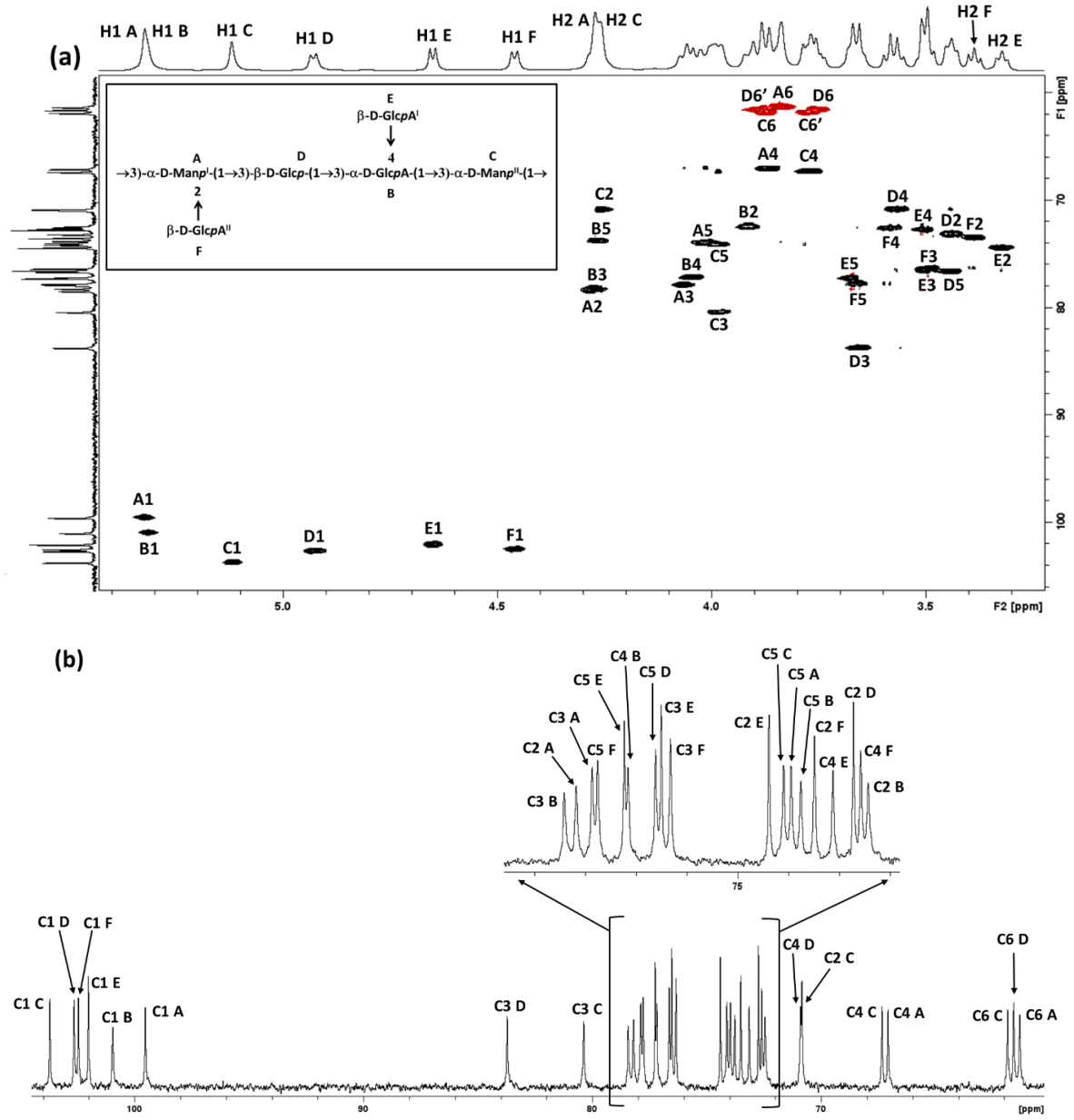


Figure 5.7: Labelled **(a)** 2D HSQC-DEPT identity map spectrum, with methylene groups cross-peaks shown in red, and **(b)** labelled 1D ^{13}C NMR spectrum of K122 CPS recorded at 343 K , and 600 and 150 MHz, respectively. A = α-Man; B = α-GlcA; C = α-Man; D = β-Glc; E = β-GlcA^{I}; F = β-GlcA^{II}.

Table 5.1: Complete ^{1}H and ^{13}C NMR chemical shifts data of the sugar residues in K122 CPS with glycosylation shifts shown in brackets below; carbon chemical shifts of the linkage carbons are underlined.

Residues	H1 C1	H2 C2	H3 C3	H4 C4	H5 C5	H6 C6
A **→2,3)-α-D-Man*p*^{I}-(1→**	5.33 99.53 (4.59)	4.27 <u>78.20</u> (6.51)	4.07 <u>77.87</u> (6.62)	3.87 67.06 (-0.88)	4.02 73.95 (0.61)	3.84 61.30 (-0.69)
B **→3,4)-α-D-Glc*p*A-(1→**	5.32 100.95 (7.99)	3.91 72.44 (0.18)	4.27 <u>78.42</u> (4.90)	4.04 <u>77.16</u> (4.25)	4.27 73.76 (1.29)	- 176.59 (-0.83)
C **→3)-α-D-Man*p*^{II}-(1→**	5.12 103.71 (8.77)	4.26 70.82 (-0.87)	3.98 <u>80.36</u> (9.11)	3.77 67.32 (-0.62)	3.99 74.10 (0.76)	3.77; 3.88 61.83 (-0.16)
D **→3)-β-D-Glc*p*-(1→**	4.93 102.64 (5.80)	3.44 73.13 (-2.07)	3.66 <u>83.70</u> (6.94)	3.57 70.88 (0.17)	3.44 76.62 (-0.14)	3.75; 3.89 61.56 (-0.28)
E **β-D-Glc*p*A^{I}-(1→**	4.65 102.02 (5.25)	3.32 74.39 (-0.61)	3.50 76.51 (-0.02)	3.50 72.72 (0.03)	3.67 77.23 (0.3)	- 176.34 (-0.13)
F **β-D-Glc*p*A^{II}-(1→**	4.46 102.45 (5.68)	3.39 73.49 (-1.51)	3.50 76.33 (-0.20)	3.58 72.58 (-0.11)	3.66 77.76 (0.83)	- 176.20 (-0.27)

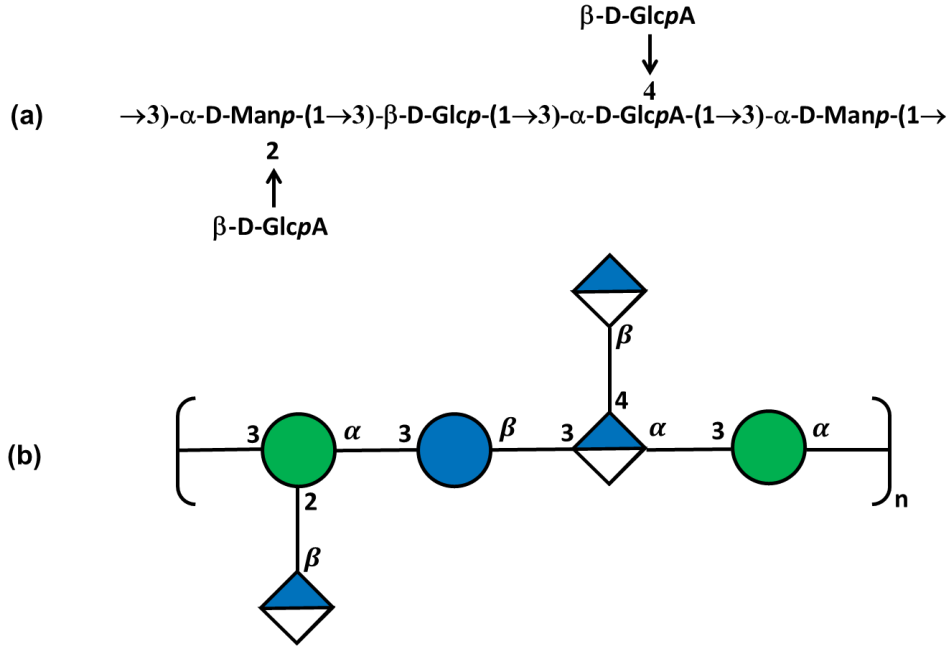


Figure 5.8: Proposed repeating unit structure of K122 capsular polysaccharide represented in (**a**) IUPAC and (**b**) SNFG format.

Chemical analysis performed at the University of Trieste corroborated the structural features of the repeating unit of K122 CPS. This followed the same methodology as for K102 (Section 3.4) and K112 (**Section 4.3**). The composition analysis showed the presence of mannose (Man); glucose (Glc); and glucuronic acid (GlcA) in the molar ratio of 3.1:1.0:3.9, respectively, which was not in good agreement with the hexasaccharide sugar composition revealed by NMR. Linkage analysis showed: 3-Glc*p*; 3-Man*p*; 2,3-Man*p*; and 2,3-Glc*p* (which may have resulted from acid interference and undermethylation of the 3-Glc), in molar ratio of 2.2:1.0:1.1:2.9. These results can be attributed to the presence of the three GlcA residues, which form strong bonds to attached neighbouring sugars resulting in reduced recovery of those sugar residues. Furthermore, the methodology used for linkage analysis does not detect uronic acids and is not quantitative, however, the data was consistent with the hexasaccharide repeating unit elucidated by NMR analysis.

5.2 KL107 CPS characterisation

5.2.1 1D ^{1}H, ^{13}C, and 2D/1D Homonuclear experiments

1D ^{1}H-DOSY NMR spectrum of KL107 (**Figure 5.9**) revealed six signals in the anomeric region, however, two of these signals at 4.66 and 4.31 ppm were shown to be deshielded methylene protons (H6 and H6') for one of the residues by performing experiments such as 2D ^{1}H-^{13}C HSQC-DEPT, which showed the cross-peaks corresponding to these two signals as inverted peaks. This was characteristic of O-acetylation at C6 in one of the residues in the RU. Indeed, the presence of O-acetylation was revealed by the proton signal of an O-acetyl group in the methyl region of the 1D ^{1}H/DOSY spectrum. This confirmed that one of the sugar residues was O-acetylated at C6, resulting in deshielded methylene group (H6 and H6') proton signals appearing in the anomeric region. Thus, only four anomeric proton signals (5.28; 5.17; 4.78; and 4.47 ppm) were present (designated H1 A - D), suggesting a tetrasaccharide RU. This was corroborated by 1D ^{13}C and ^{13}C-DEPT NMR experiments, which gave four anomeric signals and three methylene group signals, as shown in **Figure 5.10**, and also one methyl group signal from an O-acetate and two carbonyl (C=O/COOH) signals (not shown), suggesting the presence of three hexoaldose residues and one uronic acid residue.

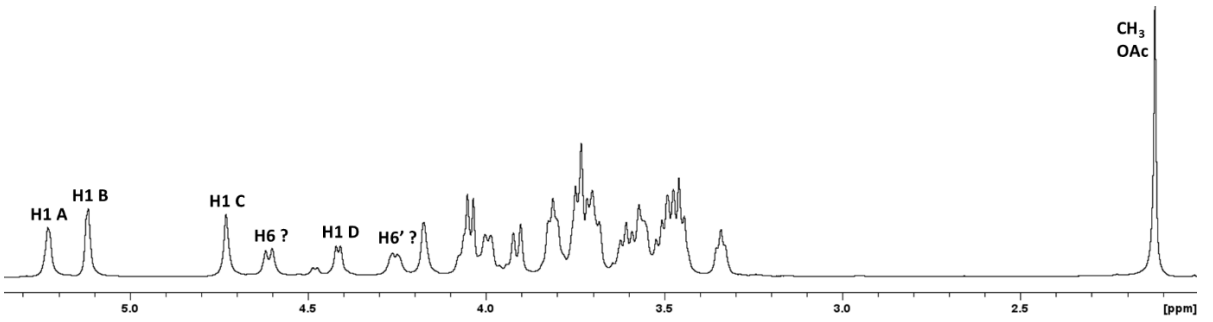


Figure 5.9: 1D DOSY-^{1}H NMR spectrum of KL107 CPS recorded at 343 K and 600 MHz, showing four diagnostic anomeric signals labelled from H1 A – D and two signals from H6 and H6', and a methyl group signal of OAc up-field.

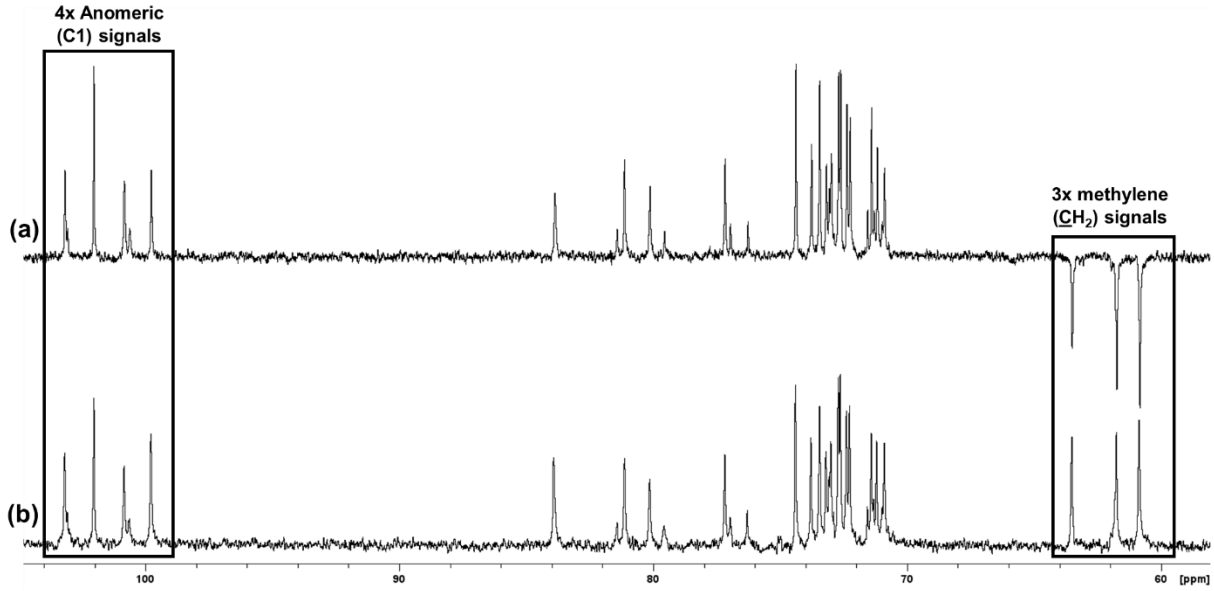


Figure 5.10: Expansion of an overlay of **(a)** 1D ^{13}C-DEPT and **(b)** ^{13}C NMR spectra of KL107 CPS recorded at 343 K and 150 MHz, showing anomeric region with 4 signals, ring region, and methylene region with 3 signals. One methyl group signal and two carbonyl group signals (not shown) are present in the methyl and carbonyl regions, respectively.

Using 2D COSY, 2D TOCSY, and a series of 1D TOCSY experiments recorded by irradiating the anomeric protons for residues A and B, H6 for residue C, and H5 for residue D, the full assignments of each spin system were obtained, as shown by the labelled 1D TOCSY spectrum of each residue in **Figure 5.11a-d**, overlaid on the 1D DOSY spectrum (**Figure 5.11e**). The TOCSY patterns, anomeric proton signals, and their coupling constants revealed the identities of residues A - D as α-Glc (**Figure 5.11a**); α-GlcA (**Figure 5.11b**); β-Man (**Figure 5.11c**); and β-Glc (**Figure 5.11d**), respectively. Thereafter, the 2D NOESY spectrum and 1D NOESY spectra from H1 of each residue were assigned to confirm some of the previous assignments and to elucidate the sequence of residues in the RU. Only intra-residue correlations from H1 to H2 were observed for residues A and B (α-sugars), whereas additional intra-

residue correlations from H1 to H3 and H1 to H5 were obtained for residues C and D (β-sugars) (**Figure 5.12**).

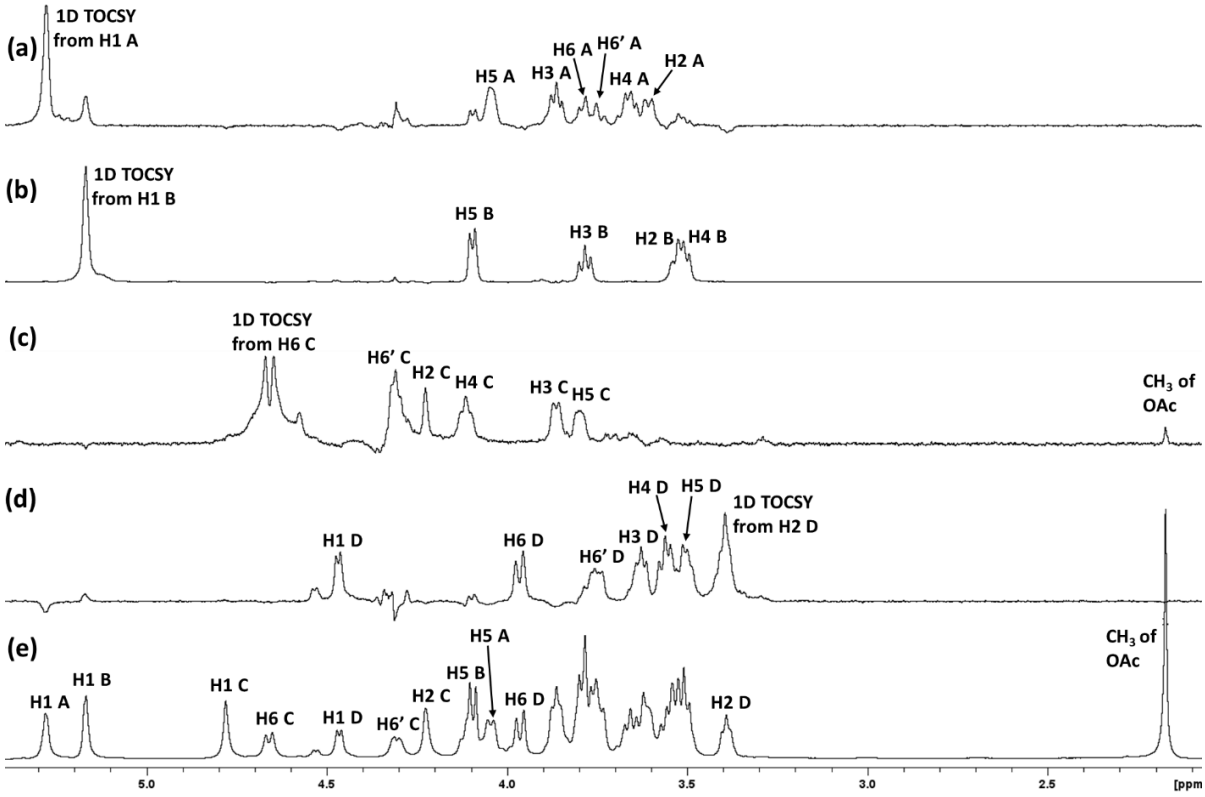


Figure 5.11: 1D TOCSY spectra of KL107 CPS recorded by irradiating H1 of residue A **(a)**, H1 of residue B **(b)**, H6 of residue C **(c)**, and H1 of residue D **(d)**, overlaid on the 1D DOSY spectrum (**e**). The diagnostic protons are labelled on each TOCSY spectrum and confirmed by cross-validation on the 1D DOSY spectrum. A = α-Glc; B = α-GlcA; C = β-Man6Ac; D = β-Glc. In some cases, there are small peaks from protons in close proximity to the irradiated proton.

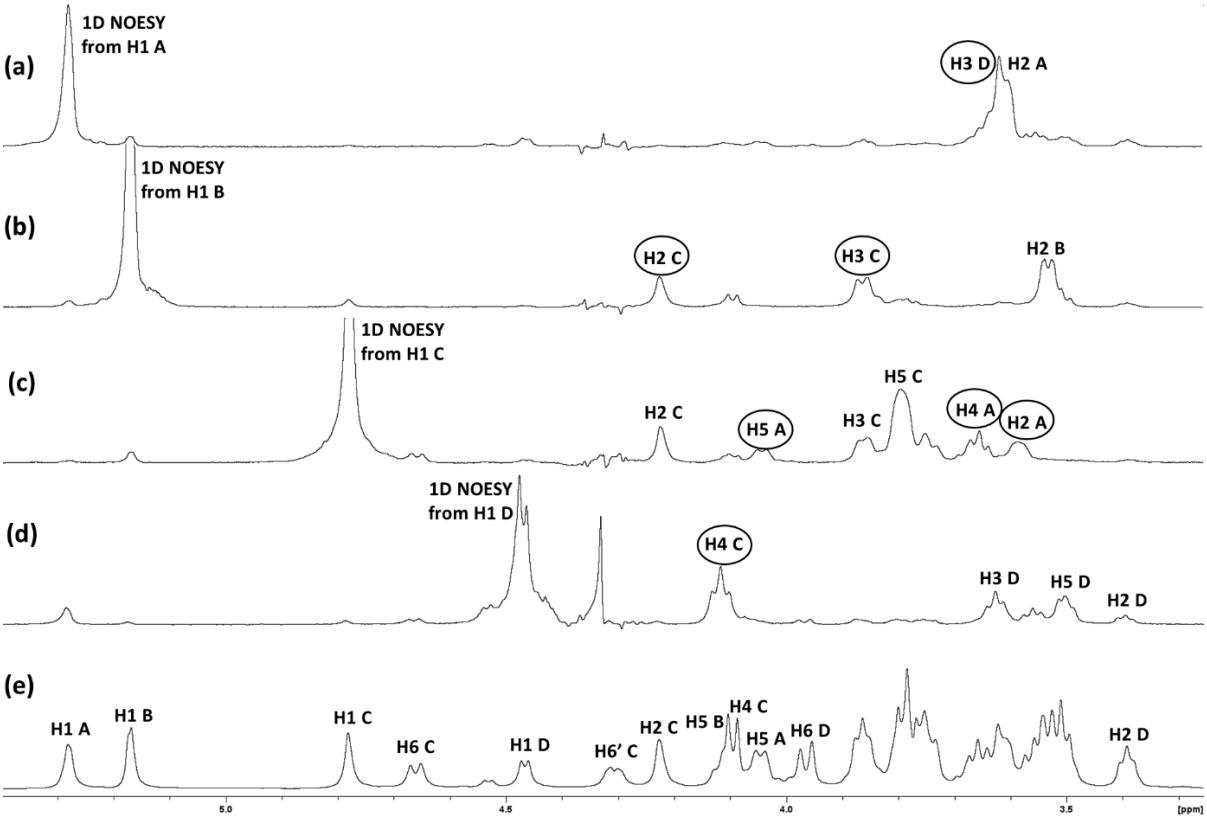


Figure 5.12: Labelled 1D NOESY spectra of KL107 CPS recorded by irradiating the anomeric protons of residues A (α-Glc) **(a)**, B (α-GlcA) **(b)**, C (β-Man6Ac) **(c)**, and D (β-Glc) **(d)**, overlaid upon 1D DOSY spectrum **(e)** for cross-validation of the signals. In some cases, there are small peaks from protons in close proximity to the irradiated proton.

5.2.2 2D ^{1}H-^{13}C heteronuclear experiments

Using the proton assignments obtained from the 2D and 1D homonuclear experiments, the 2D HSQC-DEPT spectrum was assigned, and the 2D HMBC spectrum provided additional confirmation of the assignments and resolved the ambiguities arising from overlapping proton signals (**Figure 5.13**). Additionally, HMBC gave diagnostic intra-residue correlations from H5 and H4 of residue B to the folded COOH signal, confirming residue B as α-GlcA, and gave a correlation from H6 and H6' of residue C to the C=O signal of OAc, confirming that residue C was O-acetylated at C6. Finally, the glycosylation pattern of the residues was determined from the key HMBC inter-residue correlations: H1 A (α-Glc) to C3 D (β-Glc); H1 B (α-GlcA) to C3 C (β-Man6Ac); H1 C (β-Man6Ac) to C4 A (α-Glc); and H1 D (β-Glc) to C4 C (β-Man6Ac), and corroborated by additional inter-residue correlations from the ring protons to C1 of the preceding residues, thus revealing the repeating unit structure shown in **Figure 5.14**.

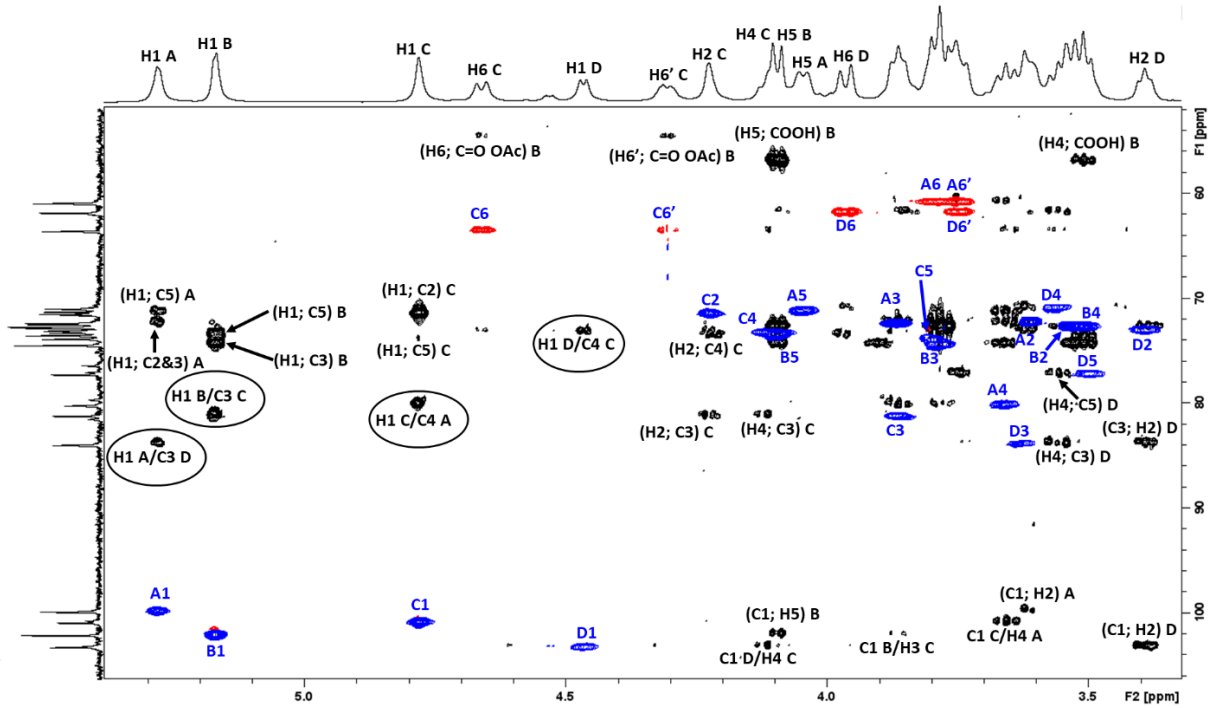


Figure 5.13: Labelled 2D HSQC-DEPT of KL107 CPS (blue, methylene group cross-peaks inverted (red)) overlaid onto HMBC (black), and the inter-residue correlations are highlighted by the ovals. Experiments were recorded at 343 K and 600 MHz. A = α-Glc; B = α-GlcA; C = β-Man6Ac; D = β-Glc.

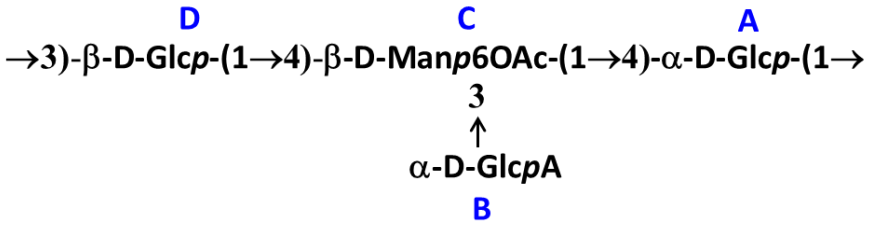


Figure 5.14: Proposed repeating unit structure of *K. pneumoniae* KL107 capsular polysaccharide.

5.2.3 Summary of the full assignments and key diagnostic spectra and comparison to chemical analysis results

For KL107 CPS, four anomeric signals were revealed by 1D proton and 1D carbon NMR spectra, suggesting a tetrasaccharide RU. Notably, two additional signals from H6 and H6' were observed in the anomeric region of the proton spectrum, and a methyl group signal of an O-acetyl group was observed in the methyl regions of both the proton and carbon spectra. This demonstrated the presence of a sugar residue with O-acetylation at C6. Additionally, the carbon spectrum gave two carbonyl signals and the DEPT experiment showed only three inverted methylene signals (one deshielded because of an OAc substituent), thus revealing the presence of one GlcA residue and three hexoses in this tetrasaccharide RU. Full proton assignments for all the spin systems were successfully obtained just from the homonuclear experiments (COSY, TOCSY, and NOESY). Finally, the 2D ^{1}H-^{13}C HSQC-DEPT identity map and the diagnostic 1D ^{13}C NMR spectra were fully assigned, as shown in **Figure 5.15a, b**. 2D HMBC was sufficient to aid and confirm the HSQC-DEPT assignments and elucidate the linkage positions of the residues.

The obtained ^{1}H and ^{13}C chemical shifts of the residues of KL107 CPS are given in **Table 5.2**, and the calculated carbon glycosylation shifts are shown in brackets below the carbon chemical shifts. The glycosylation shift values obtained are positive and large at the linkage positions of the residues (+5.5 to +9.5 ppm), consistent with the proposed repeating unit structure of KL107 CPS shown in **Figure 5.16**. As for K122 CPS, it was not possible to confirm the expected D-absolute configurations of the residues using the glycosylation shift values and the configurational factors reported in the literature.[119,120]

As for K122 CPS, chemical analysis of KL107 CPS was conducted to complement the structural features acquired from NMR analysis. The composition analysis revealed the presence of mannose, glucose, and glucuronic acid (as for K122 CPS) in the molar ratio of 1.0:1.3:1.0, in good agreement with the tetrasaccharide sugar composition shown by NMR. Linkage analysis identified: 3-Glc*p*, 4-Glc*p*, and 3,4-Man*p* in the molar ratio of 1.2:1.3:1.0. Besides the lack of detection of uronic acid residues and quantitative data, this was consistent with the tetrasaccharide repeating unit elucidated by NMR analysis.

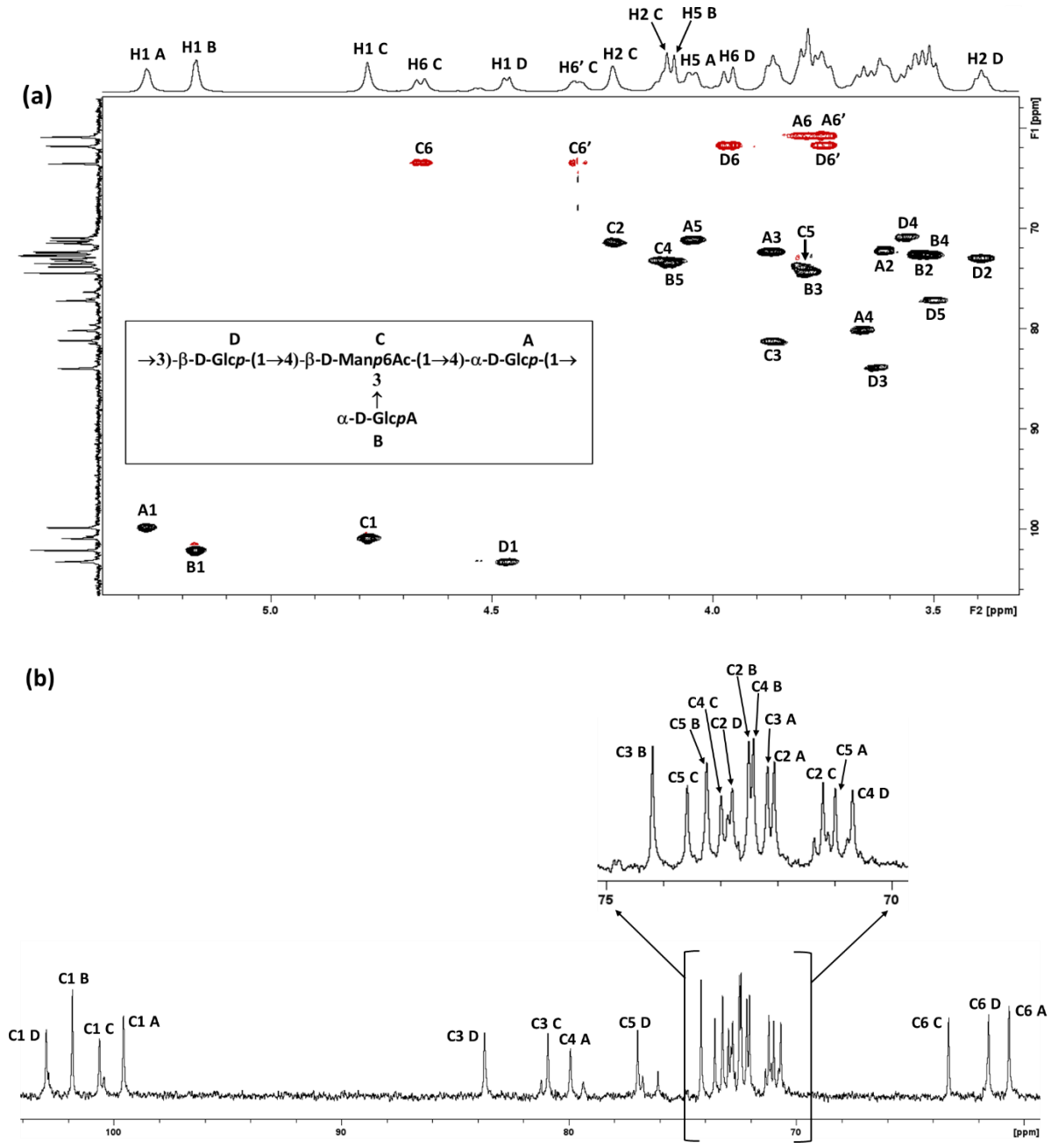


Figure 5.15: Labelled **(a)** 2D HSQC-DEPT identity map spectrum and **(b)** 1D ^{13}C NMR spectrum of KL107 CPS recorded at 343 K and 600 MHz. A = α-Glc; B = α-GlcA; C = β-Man6Ac; D = β-Glc.

Table 5.2: Complete ^{1}H and ^{13}C NMR chemical shifts data of the sugar residues in KL107 CPS with glycosylation shifts shown in brackets below the carbon chemical shifts and the linkage carbons are underlined.

Residues	H1 C1	H2 C2	H3 C3	H4 C4	H5 C5	H6 C6	OAc
A **→4)-α-D-Glc*p*-(1→**	5.28 99.79 (6.80)	3.61 72.27 (-0.20)	3.86 72.39 (-1.39)	3.66 <u>80.13</u> (9.42)	4.05 71.21 (-1.16)	3.75; 3.79 60.87 (-0.97)	-
B **α-D-Glc*p*A-(1→**	5.17 102.04 (9.08)	3.53 72.65 (0.39)	3.79 74.40 (0.88)	3.51 72.72 (-0.19)	4.10 73.45 (0.98)	- 176.89 (-0.53)	-
C **→3,4)-β-D-Man*p*6Ac-(1→**	4.78 100.86 (6.31)	4.23 71.42 (-0.71)	3.86 <u>81.14</u> (7.11)	4.12 <u>73.21</u> (5.52)	3.80 73.79 (-3.21)	4.66; 4.31 63.51 (1.52)	2.17 21.13 174.61
D **→3)-β-D-Glc*p*-(1→**	4.47 103.20 (6.36)	3.39 73.03 (-2.17)	3.63 <u>83.92</u> (7.16)	3.56 70.91 (0.20)	3.50 77.18 (0.42)	3.75; 3.97 61.76 (-0.08)	-

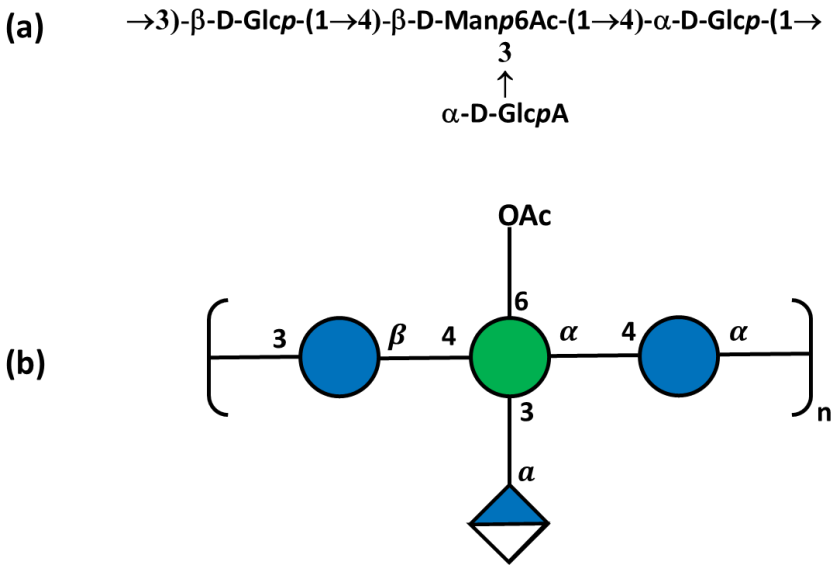


Figure 5.16: Proposed repeating unit structures of KL107 capsular polysaccharide represented in (**a**) IUPAC and (**b**) SNFG format.

5.3 Comparison of the composition and structure of the KL107 and K122 RU with other *Klebsiella* serotypes and potential cross-reactivity

The repeating unit structures and sugar composition of K122 CPS and KL107 CPS were compared to other *K. pneumoniae* K-serotypes reported in published chemical studies (i.e., the database in the appendix (**Table 1A**)). It was found that the monosaccharide composition of capsular polysaccharides of K122 and KL107 (GlcA; Glc; and Man) is common among K2; K4; K5; K24 and K39. A thorough comparison of these K-antigens revealed that the proposed RU structure of K122 is novel among *K. pneumoniae* K-antigens and has a 4 + 1 + 1 RU type as do the repeating units of K15 and K27 (**Table 1.2**). Conversely, it was found that the proposed RU structure of KL107 is not novel but identical to serotype K2 with respect to every structural feature.

The repeating unit of K2 was originally reported by Corsaro et al. but without the O-acetylation at C6 of the β-Man residue.[133] However, a recent investigation of K2 by Ovchinnikova et al.[48] revealed the presence of O-acetylation and demonstrated that a deacetylated K2 CPS gave chemical shifts in agreement with the NMR data of K2 initially published by Corsaro et al. The polysaccharide isolation procedures performed by Corsaro et al. may have removed the O-acetylation. Additionally, the NMR chemical shifts of KL107 CPS obtained in this study were also in agreement with the recently published NMR data of K2 CPS by Ovchinnikova et al. This means that KL107 is a strain of *K. pneumoniae* serotype K2. Although significant serological cross-reactivity would be expected between the O-acetylated and O-deacetylated versions of this K-antigen, the labile O-acetyl group could be an important epitope for immunogenicity. Furthermore, Corsaro et al. performed absolute configuration analysis on the K2 CPS and determined the sugar residues to have D-configurations. Therefore, the chemical shifts of KL107 CPS compared to K2 confirm the D-configurations for the sugar residues of KL107.

The KL107 (K2) tetrasaccharide repeating unit structure is similar to the repeating units of K13; K30; and K33, with the main difference being the presence of a 3,4-(S)-pyruvate substituted β-Gal sugar in their chemical compositions, giving them a pentasaccharide RU (**Figure 5.17**). The K-antigen of K13 has the 3,4-(S)-pyruvate substituted β-Gal as a terminal residue, linking to the branching α-GlcA residue, and additionally, the mannose residue in the backbone chain has a β-configuration with no O-acetylation, as shown in **Figure 5.17a**. In the RU structure of K30 (**Figure 5.17b**), the 3,4-(S)-pyruvate substituted β-Gal forms a terminal branch at the β-Man residue, replacing the O-acetyl substituent in the corresponding α-Man residue of KL107 (**Figure 5.16**). The only other differences are the β-configurations of the mannose residue and the glucose residue it attaches to, and this β-glucose contains 33% of O-acetylation in the capsular polysaccharide. Furthermore, the K-antigen of K30 is

identical to K33, which unlike K30 is completely (100%) O-acetylated at the β-Glc residue. As a result, it is unsurprising that cross-reactivities have been reported between the serotypes K2; K13; and K30,[55] which can be explained by their structural similarities. Lastly, K2 shares a trisaccharide fragment with K69, as highlighted in **Figure 5.17d**. On the other end, K122 has a tetrasaccharide fragment common to K35, which includes the terminal β-GlcA residue, as highlighted in **Figure 5.17c**. These common epitopes may potentially result in cross-reactions, which will be further investigated by our vaccine partner, GVGH.

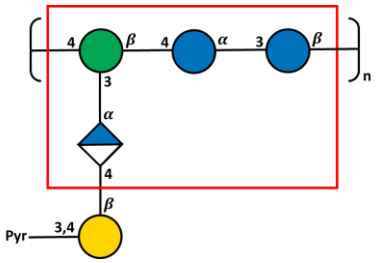
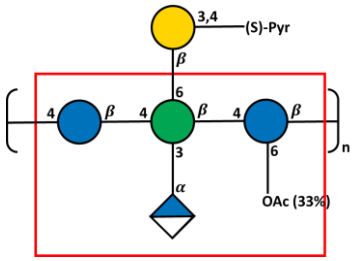
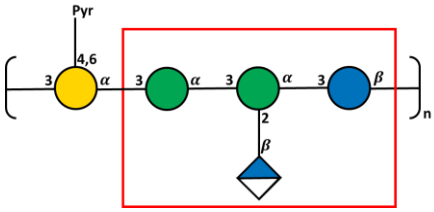
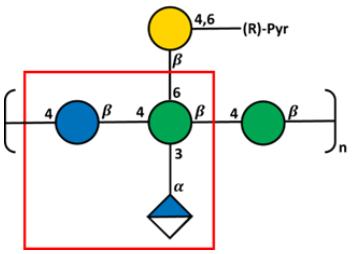
(a)	K13	
(b)	K30 (K33 with 100% OAc)	
(c)	K35	
(d)	K69	

Figure 5.17: *K. pneumoniae* K-antigens sharing epitopes with KL107: **(a)** K13, **(b)** K30/K33, and **(d)** K69, and sharing epitopes with K122: **(c)** K35, which may provide serological cross-reactivity, shown by the red borders.

5.4 Conclusions

Both K122 and KL107 capsular polysaccharides were successfully characterised, and their repeating unit structures elucidated by NMR. Structural features of the obtained repeating units were corroborated by chemical analysis performed by collaborators at the University of Trieste. As with the previously analysed structures (K102 and K112), the advantages of NMR analysis over chemical techniques for accurately determining the sugar compositions, glycosidic linkages and elucidating the sequence of residues were demonstrated. The proposed repeating unit structure of *K. pneumoniae* serotype K122 was found to be novel amongst the *K. pneumoniae* K-antigens, with a 4 + 1 + 1 hexasaccharide RU type as the CPSs of K15 and K27, while KL107 CPS was found to be the same as the K-antigen of *K. pneumoniae* serotype K2, suggesting that KL107 is serotype K2. This demonstrates the limitations of genotyping as it may not be 100% accurate in identifying some serotypes, although, it is more viable, accurate and advantageous to the traditional serotyping approach. K2 (KL107 CPS RU) shares structural similarities with small differences with *K. pneumoniae* K-antigens such as K13; K30 (K33); K35; and K69, while K35 possesses a tetrasaccharide fragment which is found in the hexasaccharide RU of K122 CPS. In fact, K2; K13 and K30 have been reported to be cross-reactive serotypes, which can be explained by the epitopes they share.[55] The cross-reactivities of KL107 and K122 will be investigated by our collaborator, the GVGH group.

Chapter 6: General Conclusions

This investigation aimed to utilise a combination of 1D and 2D NMR experiments to fully characterise and elucidate the capsular polysaccharide repeating unit structures of four emerging clinically important strains of *K. pneumoniae* that were identified using genotyping as novel KL serotypes: KL102, KL112, KL122, and KL107; these are referred to with their corresponding K-serotypes names: K102, K112, K122, respectively, except for KL107 (revealed to be a known serotype during this investigation). The purified capsular polysaccharide of K102 was obtained from our collaborators (Prof. Cescutti) at the University of Trieste, and those of K112, K122, and KL107 from our collaborators at the GVGH Vaccines Institute for Global Health in Siena.

NMR characterisation of each CPS was successfully conducted, and the complete proton and carbon assignments were obtained and presented with fully labelled 2D ^{1}H-^{13}C HSQC-DEPT and 1D ^{13}C NMR spectra, as well as the chemical shift data of each CPS recorded in respective tables. CASPER was used to obtain the chemical shifts of the unsubstituted monosaccharides found in each CPS, which were utilised to calculate the glycosylation shifts of the carbon resonances of each CPS.[118] Glycosylation shifts are the differences between the ^{13}C chemical shifts of the unsubstituted monosaccharides and the experimental ^{13}C chemical shifts of their counterparts in the CPS. The chemical shift of a linkage carbon is shifted down-field, resulting in positive and large glycosylation shift values (5 - 10 ppm), while those of the neighbouring carbons are shifted up-field, resulting in negative glycosylation shift values (1 - 2 ppm).[119] Therefore, this confirms the linkage positions of the residues in the CPS.

The structural features of the repeating units acquired using NMR spectroscopy were later compared to in-depth chemical sugar composition and linkage analysis conducted by our collaborators (Paola Cescutti) at the University of Trieste. Although the chemical analysis corroborated and complemented the NMR analysis, there were some discrepancies in the chemical analysis results, which could be rationalised. These were mainly due to the presence of glucuronic acid residues, which form a stable linkage to the attached residue, resulting in low recovery and detected quantity of that residue. Furthermore, the standard chemical linkage analysis method did not detect uronic acids and was not quantitative. As a result, this demonstrated the advantages of NMR analysis for characterisation and structural elucidation of polysaccharides. As noted, NMR spectroscopy required a relatively low amount of sample (5 - 10 mg) per capsular polysaccharide and is non-destructive, additionally, it accurately determined the sugar composition, elucidated the linkage positions, and also provided the sequence of residues. Furthermore, it identified the presence and position of the O-acetyl substituent in KL107 CPS. However, NMR also has limitations, as it does not identify the absolute configurations

(D or L) of sugars, although in some cases, these can be successfully confirmed by comparing the glycosylation shifts to configurational factors of specific disaccharide units within a repeating unit, which are available in the literature, as was achieved for K102 CPS (**Section 3.4**).

The proposed CPS repeating unit structures of the strains targeted in this investigation are shown in **Figure 6.1a-d**. These were compared to the CPS RU structures of other *K. pneumoniae* serotypes to confirm whether these repeating units are novel or not. Furthermore, this allowed the identification of CPS serotypes which share significant structural features (epitopes) with these strains, as this can potentially indicate cross-reactivity.

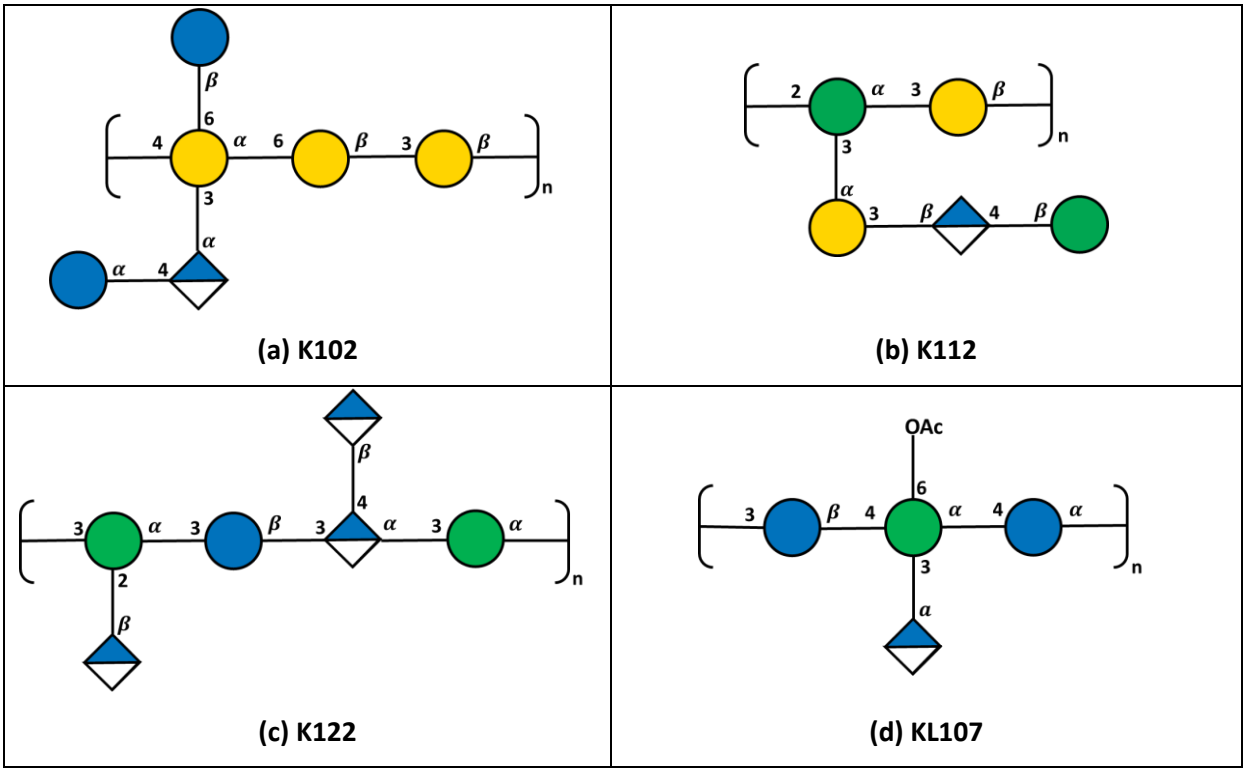


Figure 6.1: Proposed capsular polysaccharide repeating unit structures of **(a)** K102; **(b)** K112; **(c)** K122; and **(d)** KL107 represented in SNFG format.

The CPS RU of K102 (**Figure 6.1a**) was found to be novel among the K-antigens of other serotypes. Furthermore, it has a unique 3 + 2 + 1 repeating unit type compared to other *K. pneumoniae* K-antigens with a hexasaccharide repeating unit. However, it also shares significant epitopes with K15 and K18, which can potentially provide cross-reactivity (**Section 3.5**). As noted, the NMR data was able to confirm the D-absolute configurations of all sugar residues of K102 CPS, except for residue A (α-D-Glc*p*), based on the obtained glycosylation shifts and Shashkov's configurational factors.[119] However, this was not possible for the remaining capsular polysaccharides. The pentasaccharide CPS of K112 (**Figure 6.1b**) was found to be novel among other K-serotypes and has an unusual 2+3 RU type. Although this was a novel CPS RU, it was interestingly found that it has an almost identical structure

to K20, which only differs with the presence of the terminal β-Man residue on its side chain. This suggested that these serotypes are likely to be serologically cross-reactive. Additionally, it also shares some significant epitopes with K21; K31; and K43, which may potentially provide cross-reactivity (as described in Section 4.4). K122 CPS (**Figure 6.1**) was found to have a novel hexasaccharide repeating unit which has a 4 + 1 + 1 RU type as the repeating units of K15 and K27. A tetrasaccharide fragment within this hexasaccharide RU of K122 CPS is also found in the CPS structure of K35, which suggests that they may be potentially cross-reactive. KL107 CPS was the only one which turned out to be not a new K-antigen. This was found to be the same as the K-antigen of serotype K2 and suggested that it originated from a strain of K2, which is associated with hypervirulent *K. pneumoniae* isolates.[49] The repeating unit of K2 (KL107 CPS RU) is structurally similar to K13; K30 (K33); K35; and K69, with the major difference being an additional sugar residue, a pyruvate substituted galactose. Their similar structures explain the cross-reactivities which have been reported between K2; K13 and K30.[55]

It was not possible to confirm the D-configurations of the sugar residues of K112, K122, and KL107 capsular polysaccharides as the disaccharide units in their structures did not conform to the criteria of configurational factors in the literature. Nonetheless, as it was established from the constructed database of the capsular polysaccharides of *K. pneumoniae* serotypes (**Table 1A** in the **Appendix**), common aldohexose and uronic acid residues produced by *K. pneumoniae* serotypes generally have D-configurations apart from the 6-deoxy sugars (rhamnose and fucose). For KL107 CPS, which is identical to K2, the experimental NMR chemical shifts data agreed with the K2 chemical shifts in the literature, for which chemical analysis for the determination of absolute configurations has been conducted.[48] Absolute configuration chemical studies can be performed in the future to provide further confirmation of the expected D-configurations of the sugar residues of the capsular polysaccharides of K102; K112; and K122. Furthermore, the cross-reactivities postulated based on structural similarities will be investigated in-depth by our collaborators, the GVGH group.

Overall, the NMR characterisation and structural elucidation of the *K. pneumoniae* capsular polysaccharides of K102; K112; and K122 were successfully accomplished, and these were found to be novel strains, in agreement with the genotyping approach that was used to identify them. However, the NMR characterisation revealed that KL107 strain was not novel, as its CPS was found to be identical to the CPS of K2. This suggested that KL107 is the same as serotype K2 and shows that, although genotyping is a more practical and effective approach for identification of isolates, it may not accurately distinguish all K-antigens. This work also demonstrated the efficiency and power of NMR spectroscopy for the chemotyping of polysaccharides. This is crucial for clear understanding of the immunological behaviour of K-antigens and is an important initial step in the development of

conjugate vaccines, as NMR characterisation provides data for confirmation of serotype identity, purity, and integrity testing of known K-antigens to be included in the vaccine formulation.

References

[1] C.K. Kumar, K. Sands, T.R. Walsh, S. O'Brien, M. Sharland, J.A. Lewnard, H. Hu, P. Srikantiah, R. Laxminarayan, Global, regional, and national estimates of the impact of a maternal Klebsiella pneumoniae vaccine: A Bayesian modelling analysis, PLoS Med 20 (2023). https://doi.org/10.1371/journal.pmed.1004239.

[2] C.J. Murray, K.S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, et. al., Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis, The Lancet 399 (2022) 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.

[3] M. Choi, N. Hegerle, J. Nkeze, S. Sen, S. Jamindar, S. Nasrin, S. Sen, J. Permala-Booth, J. Sinclair, M.D. Tapia, J.K. Johnson, S. Mamadou, J.T. Thaden, V.G. Fowler, A. Aguilar, E. Terán, D. Decre, F. Morel, K.A. Krogfelt, A. Brauner, E. Protonotariou, E. Christaki, Y. Shindo, Y.-T. Lin, A.L. Kwa, S. Shakoor, A. Singh-Moodley, O. Perovic, J. Jacobs, O. Lunguya, R. Simon, A.S. Cross, S.M. Tennant, The Diversity of Lipopolysaccharide (O) and Capsular Polysaccharide (K) Antigens of Invasive Klebsiella pneumoniae in a Multi-Country Collection, Front Microbiol 11 (2020). https://doi.org/10.3389/fmicb.2020.01249.

[4] C. Opoku-Temeng, S.D. Kobayashi, F.R. DeLeo, Klebsiella pneumoniae capsule polysaccharide as a target for therapeutics and vaccines, Comput Struct Biotechnol J 17 (2019) 1360–1366. https://doi.org/10.1016/j.csbj.2019.09.011.

[5] R. Edelman, D.N. Talrlor, S.S. Wasserman, J.B. Mcclain, A.S. Cross, J.C. Sadoff, J.U. Que, S.J. Cryz, Phase 1 trial of a 24-valent Klebsiella capsular polysaccharide vaccine and an eight-valent Pseudomonas O-polysaccharide conjugate vaccine administered simultaneously, n.d.

[6] R.R. Wick, E. Heinz, K.E. Holt, K.L. Wyres, Kaptive Web: User-Friendly Capsule and Lipopolysaccharide Serotype Prediction for Klebsiella Genomes, J Clin Microbiol 56 (2018). https://doi.org/10.1128/JCM.00197-18.

[7] M.M.C. Lam, R.R. Wick, L.M. Judd, K.E. Holt, K.L. Wyres, Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the Klebsiella pneumoniae species complex, Microb Genom 8 (2022). https://doi.org/10.1099/mgen.0.000800.

[8] C. Friedländer, About Schizomycetes in acute fibrous pneumonia, Archives for Pathological Anatomy and Physiology and for Clinical Medicine 87 (1882) 319–324.

[9] M.K. Paczosa, J. Mecsas, Klebsiella pneumoniae: Going on the Offense with a Strong Defense, Microbiology and Molecular Biology Reviews 80 (2016) 629–661. https://doi.org/10.1128/mmbr.00078-15.

[10] D. Chang, L. Sharma, C.S. Dela Cruz, D. Zhang, Clinical Epidemiology, Risk Factors, and Control Strategies of Klebsiella pneumoniae Infection, Front Microbiol 12 (2021). https://doi.org/10.3389/fmicb.2021.750662.

[11] O.D. Chapman, The genus Klebsiella, J Bacteriol 51 (1946) 637. https://doi.org/10.1007/0-387-30746-x_8.

[12] R. Podschun, U. Ullmann, Klebsiella spp. as Nosocomial Pathogens: Epidemiology, Taxonomy, Typing Methods, and Pathogenicity Factors, 1998.

[13] R.M. Martin, M.A. Bachman, Colonization, infection, and the accessory genome of Klebsiella pneumoniae, Front Cell Infect Microbiol 8 (2018). https://doi.org/10.3389/fcimb.2018.00004.

[14] J. Ashurst, A. Dawson, Klebsiella Pneumonia, In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing (2023). https://www.ncbi.nlm.nih.gov/books/NBK519004/ (accessed November 19, 2023).

[15] M. Pollack, RogerE. Nieman, JosephA. Reinhardt, P. Charache, MasonP. Jett, PaulH. Hardy, Factors influencing colonisation and antibiotic-resistance patterns of Gram-negative bacteria in hospital patients, The Lancet 300 (1972) 668–671. https://doi.org/10.1016/S0140-6736(72)92084-3.

[16] T.A. Russo, C.M. Marr, Hypervirulent Klebsiella pneumoniae, 2019. https://journals.asm.org/journal/cmr.

[17] E. Tacconelli, E. Carrara, A. Savoldi, D. Kattula, F. Burkert, Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics, n.d. http://www.cdc.gov/drugresistance/threat-report-2013/.

[18] F. Micoli, P. Costantino, R. Adamo, Potential targets for next generation antimicrobial glycoconjugate vaccines, FEMS Microbiol Rev 42 (2018) 388–423. https://doi.org/10.1093/femsre/fuy011.

[19] D. Gu, N. Dong, Z. Zheng, D. Lin, M. Huang, L. Wang, E.W.-C. Chan, L. Shu, J. Yu, R. Zhang, S. Chen, A fatal outbreak of ST11 carbapenem-resistant hypervirulent Klebsiella pneumoniae in a Chinese hospital: a molecular epidemiological study, Lancet Infect Dis 18 (2018) 37–46. https://doi.org/10.1016/S1473-3099(17)30489-9.

[20] Y. Zhang, J. Zeng, W. Liu, F. Zhao, Z. Hu, C. Zhao, Q. Wang, X. Wang, H. Chen, H. Li, F. Zhang, S. Li, B. Cao, H. Wang, Emergence of a hypervirulent carbapenem-resistant Klebsiella pneumoniae isolate from clinical infections in China, Journal of Infection 71 (2015) 553–560. https://doi.org/10.1016/j.jinf.2015.07.010.

[21] V. Sundaram, S. Agrawal, S. Chacham, K. Mukhopadhyay, S. Dutta, P. Kumar, Klebsiella pneumoniae Brain Abscess in Neonates: A Report of 2 Cases, J Child Neurol 25 (2010) 379–382. https://doi.org/10.1177/0883073809338326.

[22] U.A. Qureshi, N.A. Wani, B.A. Charoo, T. Kosar, M.A. Qurieshi, U. Altaf, Klebsiella brain abscess in a neonate, Arch Dis Child Fetal Neonatal Ed 96 (2011) F19–F19. https://doi.org/10.1136/adc.2010.194993.

[23] A.L. Shane, P.J. Sánchez, B.J. Stoll, Neonatal sepsis, The Lancet 390 (2017) 1770–1780. https://doi.org/10.1016/S0140-6736(17)31002-4.

[24] N.A. Mohd Asri, S. Ahmad, R. Mohamud, N. Mohd Hanafi, N.F. Mohd Zaidi, A.A. Irekeola, R.H. Shueb, L.C. Yee, N. Mohd Noor, F.H. Mustafa, C.Y. Yean, N.Y. Yusof, Global Prevalence of Nosocomial Multidrug-Resistant Klebsiella pneumoniae: A Systematic Review and Meta-Analysis, Antibiotics 10 (2021) 1508. https://doi.org/10.3390/antibiotics10121508.

[25] K. Luo, J. Tang, Y. Qu, X. Yang, L. Zhang, Z. Chen, L. Kuang, M. Su, D. Mu, Nosocomial infection by Klebsiella pneumoniae among neonates: a molecular epidemiological study, Journal of Hospital Infection 108 (2021) 174–180. https://doi.org/10.1016/j.jhin.2020.11.028.

[26] R. Aggarwal, N. Sarkar, A.K. Deorari, V.K. Paul, Sepsis in the newborn, The Indian Journal of Pediatrics 68 (2001) 1143–1147. https://doi.org/10.1007/BF02722932.

[27] T. You, H. Zhang, L. Guo, K.-R. Ling, X.-Y. Hu, L.-Q. Li, Differences in clinical characteristics of early- and late-onset neonatal sepsis caused by *Klebsiella* pneumoniae, Int J Immunopathol Pharmacol 34 (2020) 205873842095058. https://doi.org/10.1177/2058738420950586.

[28] B.L. Ligon, Penicillin: Its Discovery and Early Development, Semin Pediatr Infect Dis 15 (2004) 52–57.

[29] K. Bush, Bench-to-bedside review: The role of beta-lactamases in antibiotic-resistant Gram-negative infections, Critic Care 14 (2010). https://doi.org/10.1186/cc8892.

[30] WHO Library Cataloguing-in-Publication Data Global Action Plan on Antimicrobial Resistance, 2015. www.paprika-annecy.com.

[31] K. Bush, P.A. Bradford, β-lactams and β-lactamase inhibitors: An overview, Cold Spring Harb Perspect Med 6 (2016). https://doi.org/10.1101/cshperspect.a025247.

[32] M. Shyaula, C. Khadka, P. Dawadi, M.R. Banjara, Systematic Review and Meta-analysis on Extended-Spectrum β-lactamases Producing *Klebsiella pneumoniae* in Nepal, Microbiol Insights 16 (2023) 117863612211451. https://doi.org/10.1177/11786361221145179.

[33] R.F. Potter, A.W. D'Souza, G. Dantas, The rapid spread of carbapenem-resistant Enterobacteriaceae, Drug Resistance Updates 29 (2016) 30–46. https://doi.org/10.1016/j.drup.2016.09.002.

[34] S.T. Hlophe, N.H. McKerrow, Hospital-acquired Klebsiella pneumoniae Infections in a Paediatric Intensive Care Unit, South African Journal of Child Health 8 (2014).

[35] N.Z.Z. Malinga, C.O. Shobo, C. Molechan, D.G. Amoako, O.T. Zishiri, L.A. Bester, Molecular Surveillance and Dissemination of Klebsiella pneumoniae on Frequently Encountered Surfaces in South African Public Hospitals, Microbial Drug Resistance 28 (2022) 306–316. https://doi.org/10.1089/mdr.2020.0546.

[36] H. Buys, R. Muloiwa, C. Bamford, B. Eley, Klebsiella pneumoniae bloodstream infections at a South African children's hospital 2006-2011, a cross-sectional study, BMC Infect Dis 16 (2016). https://doi.org/10.1186/s12879-016-1919-y.

[37] H. Buys, R. Muloiwa, C. Bamford, B. Eley, Klebsiella pneumoniae bloodstream infections at a South African children's hospital 2006-2011, a cross-sectional study, BMC Infect Dis 16 (2016). https://doi.org/10.1186/s12879-016-1919-y.

[38] H. Ismail, O. Perovic, Overview of antimicrobial susceptibility patterns of ESKAPE organisms isolated from patients with bacteremia in South Africa, 2016-2018, n.d. http://www.nicd.ac.za.

[39] M. Mendelson, M. van Vuuren, M. Molefe, O. Perovic, C. Govind, K. Faure, A. Brink, B. Kgope, B. Makhafola, B. Durham, C. Lyle, G. Reubenson, G. Richards, H. Finlayson, M. Molefe, M. Govind, P. Nkumbule, R. Gordon, S. Mehtar, A. Whitelaw, H. Ismail, J. Black, M. Mocke-Richter, M. Makiwane, M. Gizjelaar, N. Schellack, N. Qekwana, Surveillance for antimicrobial resistance and consumption of antimicrobials in South Africa, 2021, n.d.

[40] F. Micoli, L. Del Bino, R. Alfini, F. Carboni, M.R. Romano, R. Adamo, Glycoconjugate vaccines: current approaches towards faster vaccine design, Expert Rev Vaccines 18 (2019) 881–895. https://doi.org/10.1080/14760584.2019.1657012.

[41] B. Lepenies, Carbohydrate-Based Vaccines Methods and Protocols Methods in Molecular Biology 1331, n.d. http://www.springer.com/series/7651.

[42] S. Herget, P. V. Toukach, R. Ranzinger, W.E. Hull, Y.A. Knirel, C.W. Von Der Lieth, Statistical analysis of the bacterial carbohydrate structure data base (BCSDB): Characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans, BMC Struct Biol 8 (2008). https://doi.org/10.1186/1472-6807-8-35.

[43] R. Follador, E. Heinz, K.L. Wyres, M.J. Ellington, M. Kowarik, K.E. Holt, N.R. Thomson, The diversity of Klebsiella pneumoniae surface polysaccharides, Microb Genom 2 (2016) e000073. https://doi.org/10.1099/mgen.0.000073.

[44] J.P. Kamerling, G.J. Gerwig, Strategies for the structural analysis of carbohydrates, Comprehensive Glycoscience (2007) 1–68. http://dx.doi.org/10.1016/b978-044451967-2/00032-5.

[45] L.P.P. Patro, T. Rathinavelan, Targeting the Sugary Armor of Klebsiella Species, Front Cell Infect Microbiol 9 (2019). https://doi.org/10.3389/fcimb.2019.00367.

[46] C. Whitfield, Bacterial extracellular polysaccharides, n.d. www.nrcresearchpress.com.

[47] I. Orskov, M.A. Fife-Asbury, New Klebsiella Capsular Antigen, K82, and the Deletion of Five of Those Previously Assigned, Int J Syst Bacteriol 27 (1977) 386–387. https://doi.org/10.1099/00207713-27-4-386.

[48] O.G. Ovchinnikova, L.P. Treat, T. Teelucksingh, B.R. Clarke, T.A. Miner, C. Whitfield, K.A. Walker, V.L. Miller, Hypermucoviscosity Regulator RmpD Interacts with Wzc and Controls Capsular Polysaccharide Chain Length, MBio 14 (2023). https://doi.org/10.1128/mbio.00800-23.

[49] M.F. Feldman, A.E. Mayer Bridwell, N.E. Scott, E. Vinogradov, S.R. McKee, S.M. Chavez, J. Twentyman, C.L. Stallings, D.A. Rosen, C.M. Harding, A promising bioconjugate vaccine against hypervirulent *Klebsiella pneumoniae*, Proceedings of the National Academy of Sciences 116 (2019) 18655–18663. https://doi.org/10.1073/pnas.1907833116.

[50] W.-L. Yu, W.-C. Ko, K.-C. Cheng, C.-C. Lee, C.-C. Lai, Y.-C. Chuang, Comparison of prevalence of virulence factors for Klebsiella pneumoniae liver abscesses between isolates with capsular K1/K2 and non-K1/K2 serotypes, Diagn Microbiol Infect Dis 62 (2008) 1–6. https://doi.org/10.1016/j.diagmicrobio.2008.04.007.

[51] V. Arato, M.M. Raso, G. Gasperini, F. Berlanda Scorza, F. Micoli, Prophylaxis and Treatment against Klebsiella pneumoniae: Current Insights on This Emerging Anti-Microbial Resistant Global Threat, Int J Mol Sci 22 (2021) 4042. https://doi.org/10.3390/ijms22084042.

[52] D. Artyszuk, W. Jachymek, R. Izdebski, M. Gniadkowski, J. Lukasiewicz, The OL101 O antigen locus specifies a novel Klebsiella pneumoniae serotype O13 structure, Carbohydr Polym 326 (2024) 121581. https://doi.org/10.1016/j.carbpol.2023.121581.

[53] C.-T. Fang, Y.-J. Shih, C.-M. Cheong, W.-C. Yi, Rapid and Accurate Determination of Lipopolysaccharide O-Antigen Types in Klebsiella pneumoniae with a Novel PCR-Based O-

Genotyping Method, J Clin Microbiol 54 (2016) 666–675. https://doi.org/10.1128/JCM.02494-15.

[54] L. Mazgaeen, P. Gurung, Recent Advances in Lipopolysaccharide Recognition Systems, Int J Mol Sci 21 (2020) 379. https://doi.org/10.3390/ijms21020379.

[55] P. Pieroni, R.P. Renniet, H.G. Deneer, The use of bacteriophages to differentiate serologically cross-reactive isolates of Klebsiella pneumoniae, 1994.

[56] B. Ayling-Smith, T.L. Pitt, State of the art in typing: Klebsiella spp., Journal of Hospital Infection 16 (1990) 287–295. https://doi.org/10.1016/0195-6701(90)90001-5.

[57] I. Orskov, F. Orskov, Serotyping of Klebsiella, Methods in microbiology 14 (1984) 143–164.

[58] Y.-J. Pan, H.-C. Fang, H.-C. Yang, T.-L. Lin, P.-F. Hsieh, F.-C. Tsai, Y. Keynan, J.-T. Wang, Capsular Polysaccharide Synthesis Regions in Klebsiella pneumoniae Serotype K57 and a New Capsular Serotype, J Clin Microbiol 46 (2008) 2231–2240. https://doi.org/10.1128/JCM.01716-07.

[59] Y.J. Pan, T.L. Lin, Y.H. Chen, C.R. Hsu, P.F. Hsieh, M.C. Wu, J.T. Wang, Capsular types of Klebsiella pneumoniae revisited by wzc sequencing, PLoS One 8 (2013). https://doi.org/10.1371/journal.pone.0080670.

[60] R.-W. Tsay, L.K. Siu, C.-P. Fung, F.-Y. Chang, Characteristics of Bacteremia Between Community-Acquired and Nosocomial Klebsiella pneumoniae Infection, Arch Intern Med 162 (2002) 1021. https://doi.org/10.1001/archinte.162.9.1021.

[61] M.M. Ochoa-Díaz, S. Daza-Giovannetty, D. Gómez-Camargo, Bacterial genotyping methods: From the basics to modern, in: Methods in Molecular Biology, Humana Press Inc., 2018: pp. 13–20. https://doi.org/10.1007/978-1-4939-7604-1_2.

[62] S. Brisse, V. Passet, A.B. Haugaard, A. Babosan, N. Kassis-Chikhani, C. Struve, D. Decré, wzi Gene Sequencing, a Rapid Method for Determination of Capsular Type for Klebsiella Strains, J Clin Microbiol 51 (2013) 4073–4078. https://doi.org/10.1128/JCM.01924-13.

[63] L.P.P. Patro, K.U. Sudhakar, T. Rathinavelan, K-PAM: a unified platform to distinguish Klebsiella species K- and O-antigen types, model antigen structures and identify hypervirulent strains, Sci Rep 10 (2020). https://doi.org/10.1038/s41598-020-73360-1.

[64] Z. Chen, M. Liu, Y. Cui, L. Wang, Y. Zhang, J. Qiu, R. Yang, C. Liu, D. Zhou, A novel PCR-based genotyping scheme for clinical *Klebsiella pneumoniae*, Future Microbiol 9 (2014) 21–32. https://doi.org/10.2217/fmb.13.137.

[65] S. Brisse, S. Issenhuth-Jeanjean, P.A.D. Grimont, Molecular Serotyping of *Klebsiella* Species Isolates by Restriction of the Amplified Capsular Antigen Gene Cluster, J Clin Microbiol 42 (2004) 3388–3398. https://doi.org/10.1128/JCM.42.8.3388-3398.2004.

[66] K.L. Wyres, R.R. Wick, C. Gorrie, A. Jenney, R. Follador, N.R. Thomson, K.E. Holt, Identification of Klebsiella capsule synthesis loci from whole genome data, Microb Genom 2 (2016). https://doi.org/10.1099/mgen.0.000102.

[67] D. Lin, J. Chen, Y. Yang, J. Cheng, C. Sun, Epidemiological study of carbapenem-resistant Klebsiella pneumoniae, Open Medicine (Poland) 13 (2018) 460–466. https://doi.org/10.1515/med-2018-0070.

[68] S. Dhar, E.T. Martin, P.R. Lephart, J.P. McRoberts, T. Chopra, T.T. Burger, R. Tal-Jasper, K. Hayakawa, H. Ofer-Friedman, T. Lazarovitch, R. Zaidenstein, F. Perez, R.A. Bonomo, K.S. Kaye, D. Marchaim, Risk Factors and Outcomes for Carbapenem-Resistant Klebsiella pneumoniae Isolation, Stratified by Its Multilocus Sequence Typing: ST258 Versus Non-ST258, Open Forum Infect Dis 3 (2016). https://doi.org/10.1093/ofid/ofv213.

[69] J. Wang, D. Xu, B. Qu, C. Geng, Adult intracranial infection caused by an extended-spectrum-beta-lactamase-producing strain of hypervirulent Klebsiella pneumoniae: a case report, Ann Transl Med 10 (2022) 941–941. https://doi.org/10.21037/atm-22-3805.

[70] Y.-J. Pan, T.-L. Lin, C.-T. Chen, Y.-Y. Chen, P.-F. Hsieh, C.-R. Hsu, M.-C. Wu, J.-T. Wang, Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of Klebsiella spp, Sci Rep 5 (2015) 15573. https://doi.org/10.1038/srep15573.

[71] I. Speciale, A. Notaro, P. Garcia-Vello, F. Di Lorenzo, S. Armiento, A. Molinaro, R. Marchetti, A. Silipo, C. De Castro, Liquid-state NMR spectroscopy for complex carbohydrate structural analysis: A hitchhiker's guide, Carbohydr Polym 277 (2022). https://doi.org/10.1016/j.carbpol.2021.118885.

[72] C. Fontana, G. Widmalm, Primary Structure of Glycans by NMR Spectroscopy, Chem Rev 123 (2023) 1040–1102. https://doi.org/10.1021/acs.chemrev.2c00580.

[73] B. Lindberg, F. Lindh, J. Lönngren, W. Nimmich, Structural studies of the capsular polysaccharide of Klebsiella type 33, Carbohydr Res 70 (1979) 135–144. https://doi.org/10.1016/S0008-6215(00)83277-2.

[74] P. Cescutti, N. Ravenscroft, S. Ng, Z. Lam, G.G.S. Dutton, Structural investigation of the capsular polysaccharide produced by a novel Klebsiella serotype (SK1). Location of O-acetyl substituents using NMR and MS techniques, Carbohydr Res 244 (1993) 325–340. https://doi.org/10.1016/0008-6215(83)85011-3.

[75] N. Ravenscroft, L.A.S. Parolis, H. Parolis, Bacteriophage degradation of Klebsiella K30 capsular polysaccharide. An NMR investigation of the 3,4-pyruvated galactose-containing repeating oligosaccharide, Carbohydr Res 254 (1994) 333–340. https://doi.org/10.1016/0008-6215(94)84268-X.

[76] S.W. Cui, Food carbohydrates: chemistry, physical properties, and applications, Taylor & Francis, 2005.

[77] D. Marion, An introduction to biological NMR spectroscopy, Molecular and Cellular Proteomics 12 (2013) 3006–3025. https://doi.org/10.1074/mcp.O113.030239.

[78] J. Duus, C.H. Gotfredsen, K. Bock, Carbohydrate structural determination by NMR spectroscopy: Modern methods and limitations, Chem Rev 100 (2000) 4589–4614. https://doi.org/10.1021/cr990302n.

[79] G. Widmalm, General NMR Spectroscopy of Carbohydrates and Conformational Analysis in Solution, in: Comprehensive Glycoscience (2007) 101–132. http://dx.doi.org/10.1016/B978-044451967-2/00025-8

[80] C. Dybowski, Nuclear Magnetic Resonance and Electron Spin Resonance Spectroscopy: Introduction, in: Encyclopedia of Analytical Chemistry, Wiley, (2016) 1–7. https://doi.org/10.1002/9780470027318.a6101.pub3.

[81] G. Proctor, F.C. Yu, The Dependence of a Nuclear Magnetic Resonance Frequency Upon Chemical Compound, Physical Review 77.5 (1950) 717.

[82] T.M. Eads, S.D. Kennedy, R.G. Bryant, Solvent Suppression in High-Resolution Proton Nuclear Magnetic Resonance Based on Control of Transverse Relaxation Rate, UTC, 1986. https://pubs.acs.org/sharingguidelines.

[83] C.S. Johnson Jr, Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications, 1999.

[84] N. Ravenscroft, P. Costantino, P. Talaga, R. Rodriguez, W. Egan, Glycoconjugate Vaccines, in: Vaccine Analysis: Strategies, Principles, and Control, Springer Berlin Heidelberg, Berlin, Heidelberg, (2015). 301–381. https://doi.org/10.1007/978-3-662-45024-6_8.

[85] J. Hütter, B. Lepenies, Carbohydrate-Based Vaccines: An Overview, in: Lepenies, B. (eds) Carbohydrate-Based Vaccines. Methods in Molecular Biology, 2015: pp. 1–10. https://doi.org/10.1007/978-1-4939-2874-3_1.

[86] A. Weintraub, Immunology of bacterial polysaccharide antigens, Carbohydr Res 338 (2003) 2539–2547. https://doi.org/10.1016/j.carres.2003.07.008.

[87] D. Goldblatt, Conjugate vaccines, Clinical & Experimental Immunology 119.1 (2000) 1-3.

[88] F. Micoli, M.R. Romano, F. Carboni, R. Adamo, F. Berti, Strengths and weaknesses of pneumococcal conjugate vaccines, Glycoconj J 40 (2023) 135–148. https://doi.org/10.1007/s10719-023-10100-3.

[89] R. Steffen, E. Caumes, Three novel pentavalent meningococcal vaccines, J Travel Med (2023). https://doi.org/10.1093/jtm/taad152.

[90] P.C. McCarthy, A. Sharyan, L. Sheikhi Moghaddam, Meningococcal Vaccines: Current Status and Emerging Strategies, Vaccines (Basel) 6 (2018) 12. https://doi.org/10.3390/vaccines6010012.

[91] I. Frost, H. Sati, P. Garcia-Vello, M. Hasso-Agopsowicz, C. Lienhardt, V. Gigante, P. Beyer, The role of bacterial vaccines in the fight against antimicrobial resistance: an analysis of the preclinical and clinical development pipeline, Lancet Microbe 4 (2023) e113–e125. https://doi.org/10.1016/S2666-5247(22)00303-2.

[92] P. Ranjbarian, Z. Sobhi Amjad, R. Chegene Lorestani, A. Shojaeian, M. Rostamian, Klebsiella pneumoniae vaccine studies in animal models, Biologicals 82 (2023) 101678. https://doi.org/10.1016/j.biologicals.2023.101678.

[93] M. Choi, S.M. Tennant, R. Simon, A.S. Cross, Progress towards the development of Klebsiella vaccines, Expert Rev Vaccines 18 (2019) 681–691. https://doi.org/10.1080/14760584.2019.1635460.

[94] T.A. Ahmad, L.H. El-Sayed, M. Haroun, A.A. Hussein, E.S.H. El Ashry, Development of immunization trials against Klebsiella pneumoniae, Vaccine 30 (2012) 2411–2420. https://doi.org/10.1016/j.vaccine.2011.11.027.

[95] S.J. Cryz, P.M. Mortimer, V. Mansfield, R. Germanier, Seroepidemiology of Klebsiella bacteremic isolates and implications for vaccine development, J Clin Microbiol 23 (1986) 687–690. https://doi.org/10.1128/jcm.23.4.687-690.1986.

[96] M. Granström, B. Wretlind, B. Markman, S. Cryz, Enzyme-linked immunosorbent assay to evaluate the immunogenicity of a polyvalent Klebsiella capsular polysaccharide vaccine in humans, J Clin Microbiol 26 (1988) 2257–2261. https://doi.org/10.1128/jcm.26.11.2257-2261.1988.

[97] S.J. Cryz, A.S. Cross, G.C. Sadof, J.U. Que, Human IgG and IgA subclass response following immunization with a polyvalent Klebsiella capsular polysaccharide vaccine, Eur J Immunol 18 (1988) 2073–2075. https://doi.org/10.1002/eji.1830181230.

[98] R.J. Jones, E.A. Roe, Vaccination Against 77 Capsular Types of Klebsiella Aerogenes with Polyvalent Klebsiella Vaccines, J Med Microbiol 18 (1984) 413–421. https://doi.org/10.1099/00222615-18-3-413.

[99] P.L. Wantuch, C.J. Knoot, L.S. Robinson, E. Vinogradov, N.E. Scott, C.M. Harding, D.A. Rosen, Capsular polysaccharide inhibits vaccine-induced O-antigen antibody binding and function across both classical and hypervirulent K2:O1 strains of Klebsiella pneumoniae, PLoS Pathog 19 (2023). https://doi.org/10.1371/journal.ppat.1011367.

[100] A. Cross, A. Artenstein, J. Que, T. Fredeking, E. Furer, J.C. Sadoff, S. J. Cryz Jr., Safety and Immunogenicity of a Polyvalent Escherichia Coli Vaccine In Human Volunteers, Journal of Infectious Diseases 170 (1994) 834–840. https://doi.org/10.1093/infdis/170.4.834.

[101] L.-A. Barel, L.A. Mulard, Classical and novel strategies to develop a *Shigella* glycoconjugate vaccine: from concept to efficacy in human, Hum Vaccin Immunother 15 (2019) 1338–1356. https://doi.org/10.1080/21645515.2019.1606972.

[102] A. Naini, M.P. Bartetzko, S.R. Sanapala, F. Broecker, V. Wirtz, M.P. Lisboa, S.G. Parameswarappa, D. Knopp, J. Przygodda, M. Hakelberg, R. Pan, A. Patel, L. Chorro, A. Illenberger, C. Ponce, S. Kodali, J. Lypowy, A.S. Anderson, R.G.K. Donald, A. von Bonin, C.L. Pereira, Semisynthetic Glycoconjugate Vaccine Candidates against *Escherichia coli* O25B Induce Functional IgG Antibodies in Mice, JACS Au 2 (2022) 2135–2151. https://doi.org/10.1021/jacsau.2c00401.

[103] M. Trautmann, T.K. Held, A.S. Cross, O antigen seroepidemiology of Klebsiella clinical isolates and implications for immunoprophylaxis of Klebsiella infections, Vaccine 22 (2004) 818–821. https://doi.org/10.1016/j.vaccine.2003.11.026.

[104] P.L. Wantuch, C.J. Knoot, L.S. Robinson, E. Vinogradov, N.E. Scott, C.M. Harding, D.A. Rosen, Capsular polysaccharide inhibits vaccine-induced O-antigen antibody binding and function across both classical and hypervirulent K2:O1 strains of Klebsiella pneumoniae, PLoS Pathog 19 (2023) e1011367. https://doi.org/10.1371/journal.ppat.1011367.

[105] P.L. Wantuch, C.J. Knoot, L.S. Robinson, E. Vinogradov, N.E. Scott, C.M. Harding, D.A. Rosen, A heptavalent O-antigen bioconjugate vaccine exhibits differential functional antibody responses against diverse Klebsiella pneumoniae isolates, BioRxiv (2023) 2023.12.12.571344. https://doi.org/10.1101/2023.12.12.571344.

[106] T.K. Held, N.R.M. Jendrike, T. Rukavina, R. Podschun, M. Trautmann, Binding to and Opsonophagocytic Activity of O-Antigen-Specific Monoclonal Antibodies against Encapsulated and Nonencapsulated *Klebsiella pneumoniae* Serotype O1 Strains, Infect Immun 68 (2000) 2402–2409. https://doi.org/10.1128/IAI.68.5.2402-2409.2000.

[107] S. Merino, S. Camprubí, S. Albertí, V.J. Benedí, J.M. Tomás, Mechanisms of Klebsiella pneumoniae resistance to complement-mediated killing, Infect Immun 60 (1992) 2529–2535. https://doi.org/10.1128/iai.60.6.2529-2535.1992.

[108] N. Ravenscroft, F. Berti, NMR characterization of bacterial glycans and glycoconjugate vaccines, in: Recent Trends in Carbohydrate Chemistry: Synthesis and Biomedical Applications of Glycans and Glycoconjugates, Elsevier, 2020: pp. 239–281. https://doi.org/10.1016/B978-0-12-820954-7.00007-4.

[109] C. Abeygunawardana, T.C. Williams, J.S. Sumner, J.P. Hennessey, Development and Validation of an NMR-Based Identity Assay for Bacterial Polysaccharides, Anal Biochem 279 (2000) 226–240. https://doi.org/10.1006/abio.1999.4470.

[110] G. Widmalm, General NMR Spectroscopy of Carbohydrates and Conformational Analysis in Solution, in: Comprehensive Glycoscience, Elsevier, 2007: pp. 101–132. https://doi.org/10.1016/B978-044451967-2/00025-8.

[111] W.A. Bubb, NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity, Concepts Magn Reson Part A Bridg Educ Res 19 (2003) 1–19. https://doi.org/10.1002/cmr.a.10080.

[112] H.E. Gottlieb, V. Kotlyar, A. Nudelman, NMR chemical shifts of common laboratory solvents as trace impurities, Journal of Organic Chemistry 62 (1997) 7512–7515. https://doi.org/10.1021/jo971176v.

[113] G. Zheng, W.S. Price, Solvent signal suppression in NMR, Prog Nucl Magn Reson Spectrosc 56 (2010) 267–288. https://doi.org/10.1016/j.pnmrs.2010.01.001.

[114] H. Mo, D. Raftery, Pre-SAT180, a simple and effective method for residual water suppression, Journal of Magnetic Resonance 190 (2008) 1–6. https://doi.org/10.1016/j.jmr.2007.09.016.

[115] E.M. Vilén, M. Klinger, C. Sandström, Application of diffusion-edited NMR spectroscopy for selective suppression of water signal in the determination of monomer composition in alginates, Magnetic Resonance in Chemistry 49 (2011) 584–591. https://doi.org/10.1002/mrc.2789.

[116] K. Bock, C. Pedersen, Carbon-13 Nuclear Magnetic Resonance Spectroscopy of Monosaccharides, in: Advances in Carbohydrate Chemistry and Biochemistry, 1983: pp. 27–66. https://doi.org/10.1016/S0065-2318(08)60055-4.

[117] C. Jones, B. Mulloy, The Application of Nuclear Magnetic Resonance to Structural Studies of Polysaccharides, in: Spectroscopic Methods and Analyses, Humana Press, New Jersey, n.d.: pp. 149–168. https://doi.org/10.1385/0-89603-215-9:149.

[118] M. Lundborg, G. Widmalm, Structural analysis of glycans by NMR chemical shift prediction, Anal Chem 83 (2011) 1514–1517. https://doi.org/10.1021/ac1032534.

[119] A.S. Shashkov, G.M. Lipkind, Y.A. Knirel, N.K. Kochetkov, N.D. Zelinsky, Stereochemical Factors Determining the Effects of Glycosylation on the ^{13}C Chemical Shifts in Carbohydrates, 1988.

[120] G.M. Lipkind, A.S. Shashkov, Y.A. Knirel, E. V Vinogudov, N.K. Kochetkov, N.D. Zeiinsky, A computer-assisted structural analysis of regular poly-saccharides on the basis of r3C-N.M.R. data, Elsevier Science Publishers B.V, 1988.

[121] D.M. Doddrell, D.T. Pegg, M.R. Bendall, Distortionless Enhancement of NMR Signals by Polarization Transfer, Journal of Magnetic Resonance (1969) 48.2 (1982) 323-327.

[122] K. Gheysen, C. Mihai, K. Conrath, J.C. Martins, Rapid Identification of Common Hexapyranose Monosaccharide Units by a Simple TOCSY Matching Approach, Chemistry – A European Journal 14 (2008) 8869–8878. https://doi.org/10.1002/chem.200801081.

[123] C. Roumestand, C. Delay, J.A. Gavin, D. Canet, A practical approach to the implementation of selectivity in homonuclear multidimensional NMR with frequency selective-filtering techniques. Application to the chemical structure elucidation of complex oligosaccharides, Magnetic Resonance in Chemistry 37 (1999) 451–478. https://doi.org/10.1002/(SICI)1097-458X(199907)37:7<451::AID-MRC487>3.0.CO;2-U.

[124] C. Dalvit, G. Bovermann, Pulsed field gradient one-dimensional NMR selective ROE and TOCSY experiments, Magnetic Resonance in Chemistry 33 (1995) 156–159. https://doi.org/10.1002/mrc.1260330214.

[125] J.H. Prestegard, C.M. Bougault, A.I. Kishore, Residual Dipolar Couplings in Structure Determination of Biomolecules, Chem Rev 104 (2004) 3519–3540. https://doi.org/10.1021/cr030419i.

[126] A. Kumar, R.R. Ernst, K. Wüthrich, A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules, Biochem Biophys Res Commun 95 (1980) 1–6. https://doi.org/10.1016/0006-291X(80)90695-6.

[127] Atta-ur-Rahman, M.I. Choudhary, Recent Developments in NMR Spectroscopy, in: Solving Problems with NMR Spectroscopy, Elsevier, 1996: pp. 365–390. https://doi.org/10.1016/B978-012066320-0/50009-4.

[128] G. Bodenhausen, D.J. Ruben, Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, Chem Phys Lett 69 (1980) 185–189. https://doi.org/10.1016/0009-2614(80)80041-8.

[129] C. Fontana, G. Widmalm, Primary Structure of Glycans by NMR Spectroscopy, Chem Rev 123 (2023) 1040–1102. https://doi.org/10.1021/acs.chemrev.2c00580.

[130] L. Lerner, A. Bax, Sensitivity-enhanced two-dimensional heteronuclear relayed coherence transfer NMR spectroscopy, Journal of Magnetic Resonance (1969) 69 (1986) 375–380. https://doi.org/10.1016/0022-2364(86)90091-0.

[131] R. Wagner, S. Berger, Heteronuclear Edited Gradient Selected 1D and 2D NOE Spectra: Determination of the NOE Effect between Chemically Equivalent Protons, Magnetic Resonance in Chemistry 35 (1997) 199–202. https://doi.org/10.1002/(SICI)1097-458X(199703)35:3<199::AID-OMR55>3.0.CO;2-1.

[132] Ad. Bax, M.F. Summers, Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR, J Am Chem Soc 108 (1986) 2093–2094. https://doi.org/10.1021/ja00268a061.

[133] M.M. Corsaro, C. De Castro, T. Naldi, M. Parrilli, J.M. Tomás, M. Regué, 1H and 13C NMR characterization and secondary structure of the K2 polysaccharide of Klebsiella pneumoniae

strain 52145, Carbohydr Res 340 (2005) 2212–2217. https://doi.org/10.1016/j.carres.2005.07.006.

Appendix

Table 1A: Database of known *Klebsiella pneumoniae* capsular polysaccharide (K-antigen) repeating unit structures.

K-antigen Serotype	Repeating unit structure	Repeating unit size	Repeating unit type (branching structure)	References
K1	→4)-[2,3-Pyr]-β-D-Glc*p*A-(1→4)-α-L-Fuc*p*-(1→3)-β-D-Glc*p*-(1→	3	3	Erbing, C.; Kenne, L.; Lindberg, B.; Lönngren, J.; W. Sutherland, I. *Carbohydrate Research*, **1976**, *50(1)*, 115–20. http://dx.doi.org/10.1016/s0008-6215(00)84088-4.
K2	→3)-β-D-Glc*p*-(1→4)-[α-D-Glc*p*A-(1→3)]β-D-Man*p*-(1→4)-α-D-Glc*p*-(1→	4	3+1	Ovchinnikova, O.G.; Treat, L.P.; Teelucksingh, T.; Clarke, B.R.; Miner, T.A.; Whitfield, C.; Walker, K.A.; Miller, V.L. *mBio*, **2023**, *14(3)*. http://dx.doi.org/10.1128/mbio.00800-23.
K3	→2)-[[4,6-Pyr]α-D-Man*p*-(1→4)]α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→	5	4+1	Dutton, G.G.S.; Parolis, H.; Joseleau, J.-P.; Marais, M.-F. *Carbohydrate Research*, **1986**, *149(2)*, 411–23. http://dx.doi.org/10.1016/s0008-6215(00)90061-2.
K4	→3)-α-D-Glc*p*-(1→2)-α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	4	4	Merrifield, E.H.; Stephen, A.M. *Carbohydrate Research*, **1981**, *96(1)*, 113–20. http://dx.doi.org/10.1016/s0008-6215(00)84701-1.

K5	→4)-β-D-Glc*p*A-(1→4)-β-D-Glc*p*2OAc-(1→3)-[4,6-Pyr]β-D-Man*p*-(1→	3	3	Dutton, G.G.; Yang, M.-T. *Canadian Journal of Chemistry*, **1973**, *51(11)*, 1826–32. http://dx.doi.org/10.1139/v73-272.
K6	→3)-α-L-Fuc*p*-(1→3)-β-D-Glc*p*-(1→3)-β-D-Man*p*-(1→4)-α-D-Glc*p*A-(1→	4	4	Elsässer-Beile, U.; Friebolin, H.; Stirm, S. *Carbohydrate Research*, **1978**, *65(2)*, 245–9. http://dx.doi.org/10.1016/s0008-6215(00)84316-5.
K7	→3)-β-D-Glc*p*A-(1→2)-[α-D-Gal*p*-(1→3)]α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-[4,6-(S)-Pyr]β-D-Glc*p*-(1→3)-β-D-Glc*p*-(1→	6	5+1	Kjellberg, A.; Widmalm, G.; Jansson, P.-E.; Nimmich, W. *Carbohydrate Research*, **1995**, *273(1)*, 53–62. http://dx.doi.org/10.1016/0008-6215(95)00060-7.
K8	→3)-[β-D-Glc*p*A-(1→4)]β-D-Gal*p*-(1→3)-α-D-Gal*p*-(1→3)-[4,6-(R)-Pyr]β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→	5	4+1	Erbing, B.; Jansson, P.-E.; Widmalm, G.; Nimmich, W. *Carbohydrate Research*, **1995**, *273(2)*, 197–205. http://dx.doi.org/10.1016/0008-6215(95)00099-f.
K9	→3)-α-L-Rha*p*-(1→3)[β-D-Glc*p*A-(1→4)]-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-α-D-Gal*p*-(1→	5	4+1	Isaac, D.H.; Atkins, E.D.T.; Stirm, S. *International Journal of Biological Macromolecules*, **1981**, *3(3)*, 165–70. http://dx.doi.org/10.1016/0141-8130(81)90058-1.
K10	→4)-β-D-Glc*p*A-(1→2)-[α-D-Gal*p*-(1→3)]α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-α-D-Gal*p*-(1→	6	5+1	Dutton, G.G.S.; Ng, S.K.; Parolis, L.A.S.; Parolis, H.; Chakraborty, A.K. *Carbohydrate Research*, **1989**, *193*, 147–55. http://dx.doi.org/10.1016/0008-6215(89)85114-6.
K11	→3)-β-D-Glc*p*-(1→3)-[[4,6-Pyr]α-D-Gal*p*-(1→4)]β-D-Glc*p*A-(1→3)-α-D-Gal*p*-(1→	4	3+1	Thurow, H.; Choy, Y.-M.; Frank, N.; Niemann, H.; Stirm, S. *Carbohydrate Research*, **1975**, *41(1)*, 241–55. http://dx.doi.org/10.1016/s0008-6215(00)87023-8.

K12	→3)-α-D-Gal*p*-(1→2)-[[4,6-Pyr]β-D-Gal*p*-(1→4)-β-D-Glc*p*A-(1→3)]β-D-Gal*f*-(1→6)-α-D-Glc*p*-(1→3)-α-L-Rha*p*-(1→	6	4+2	Dutton, G.G.S.; Savage, A.V. *Carbohydrate Research*, **1980**, *83(2)*, 351–62. http://dx.doi.org/10.1016/s0008-6215(00)84547-4.
K13	→4)-[[3,4-Pyr]β-D-Gal*p*-(1→4)-α-D-Glc*p*A-(1→3)]β-D-Man*p*-(1→4)-α-D-Glc*p*-(1→3)-β-D-Glc*p*-(1→	5	3+2	Niemann, H.; Frank, N.; Stirm, S. *Carbohydrate Research*, **1977**, *59(1)*, 165–77. http://dx.doi.org/10.1016/s0008-6215(00)83303-0.
K14	→4)-β-D-Glc*p*A-(1→3)-β-D-Gal*f*-(1→3)-β-D-Glc*p*-(1→4)-[[4,6-Pyr]β-D-Glc*p*-(1→2)][α-L-Rha*p*-(1→3)]β-D-Man*p*-(1→	6	4+2	Dutton, G.G.S.; Parolis, H.; Parolis, L.A.S. *Carbohydrate Research*, **1985**, *140(2)*, 263–75. http://dx.doi.org/10.1016/0008-6215(85)85127-2.
K15	→4)-[β-D-Glc*p*A-(1→3)][β-D-Glc*p*-(1→6)]β-D-Gal*p*-(1→3)-α-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→	6	4+1+1	Parolis, H.; Parolis, L.A.S.; Whittaker, D.V. *Carbohydrate Research*, **1992**, *231*, 93–103. http://dx.doi.org/10.1016/0008-6215(92)84011-g.
K16	→3)-[β-D-Gal*p*-(1→4)]-α-D-Glc*p*-(1→4)-β-D-Glc*p*A-(1→4)-α-L-Fuc*p*-(1→	4	3+1	Chakraborty, A.K.; Friebolin, H.; Niemann, H.; Stirm, S. *Carbohydrate Research*, **1977**, *59(2)*, 525–9. http://dx.doi.org/10.1016/s0008-6215(00)83189-4.
K17	→4)-β-D-Glc*p*-(1→2)-[α-L-Rha*p*-(1→3)]α-L-Rha*p*-(1→4)-α-D-Glc*p*A-(1→3)-β-L-Rha*p*-(1→	5	4+1	Dutton, G.G.S.; Folkman, T.E. *Carbohydrate Research*, **1980**, *80(1)*, 147–61. http://dx.doi.org/10.1016/s0008-6215(00)85322-7.
K18	→3)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→4)-[α-D-Glc*p*-(1→4)-β-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→3)]α-D-Glc*p*-(1→	6	3+3	Dutton, G.G.; Savage, A.V.; Vignon, M. *Canadian Journal of Chemistry*, **1980**, *58(23)*, 2588–91. http://dx.doi.org/10.1139/v80-413.

K19	→2)-α-L-Rha*p*-(1→2)-α-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→3)-[α-L-Rha*p*-(1→4)]α-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→	6	5+1	Buerret, M.; Vignon, M.; Joseleau, J.-P. *Carbohydrate Research*, **1986**, *157*, 13–25. http://dx.doi.org/10.1016/0008-6215(86)85057-1.
K20	→2)-[β-D-Glc*p*A-(1→3)-α-D-Gal*p*-(1→3)]α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→	4	2+2	Choy, Y.M.; Dutton, G.G. *Canadian Journal of Chemistry*, **1973**, *51(2)*, 198–207. http://dx.doi.org/10.1139/v73-031.
K21	→3)-[[4,6-Pyr]α-D-Gal*p*-(1→4)]α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→	5	4+1	Choy, Y.M.; Dutton, G.G. *Canadian Journal of Chemistry*, **1973**, *51(2)*, 198–207. http://dx.doi.org/10.1139/v73-031.
K21b	→3)-[[4,6-(R)-Pyr]α-D-Gal*p*-(1→4)]β-D-Glc*p*A2OAc-(1→3)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→	5	4+1	Chowdhury, T.A.; Jansson, P.-E.; Lindberg, B.; Lindquist, U. *Carbohydrate Research*, **1989**, *190(1)*, 145–51. http://dx.doi.org/10.1016/0008-6215(89)84154-0.
K22	→3)-[[4-O-(S)-1-carboxyethyl]β-D-Glc*p*A-(1→6)-α-D-Glc*p*-(1→4)]β-D-Gal*p*6OAc-(1→4)-β-D-Glc*p*-(1→	4	2+2	Parolis, L.A.S.; Parolis, H.; Niemann, H.; Stirm, S. *Carbohydrate Research*, **1988**, *179*, 301–14. http://dx.doi.org/10.1016/0008-6215(88)84126-0.
K23	→3)-[β-D-Glc*p*A-(1→6)-α-D-Glc*p*-(1→2)]α-L-Rha*p*-(1→3)-β-D-Glc*p*-(1→	4	2+2	Dutton, G.G.S.; Mackie, K.L.; Savage, A.V.; Stephenson, M.D. *Carbohydrate Research*, **1978**, *66(1)*, 125–31. http://dx.doi.org/10.1016/s0008-6215(00)83245-0.
K24	→2)-[β-D-Man*p*-(1→4)]α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	5	4+1	Choy, Y.-M.; Dutton, G.G.; Zanlungo, A.M. *Canadian Journal of Chemistry*, **1973**, *51(11)*, 1819–25. http://dx.doi.org/10.1139/v73-271.

K25	→3)-[β-D-Glc*p*-(1→2)-β-D-Glc*p*A-(1→4)]β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→	4	2+2	Niemann, H.; Kwiatkowski, B.; Westphal, U.; Stirm, S. *Journal of Bacteriology*, **1977**, *130(1)*, 366–74. http://dx.doi.org/10.1128/jb.130.1.366-374.1977.
K26	→3)-β-D-Gal*p*-(1→2)-[[4,6-Pyr]β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)]α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→	7	4+3	Di Fabio, J.; Dutton, G.G.S. *Carbohydrate Research*, **1981**, *92(2)*, 287–98. http://dx.doi.org/10.1016/s0008-6215(00)80399-7.
K27	→3)-[4,6-Pyr]β-D-Glc*p*-(1→3)-D-Gal*p*-(1→3)-[D-Glc*p*-(1→4)][β-D-Glc*p*A-(1→6)]D-Gal*p*-(1→6)-β-D-Glc*p*-(1→	6	4+1+1	Churms, S.C.; Merriefield, E.H.; Stephen, A.M. *Carbohydrate Research*, **1980**, *81(1)*, 49–58. http://dx.doi.org/10.1016/s0008-6215(00)85676-1.
K28	→2)-α-D-Gal*p*-(1→3)-[β-D-Glc*p*-(1→3)-β-D-Glc*p*A-(1→2)]α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	6	4+2	Curvall, M.; Lindberg, B.; Lönngren, J.; Nimmich, W. *Carbohydrate Research*, **1975**, *42(1)*, 95–105. http://dx.doi.org/10.1016/s0008-6215(00)84103-8.
K30	→4)-β-D-Glc*p*-(1→4)-[[3,4-Pyr]β-D-Gal*p*-(1→6)][α-D-Glc*p*A-(1→3)]β-D-Man*p*-(1→4)-β-D-Man*p*6OAc-(1→	5	3+1+1	Ravenscroft, N.; Parolis, L.A.S.; Parolis, H. *Carbohydrate Research*, **1994**, *254*, 333–40. http://dx.doi.org/10.1016/0008-6215(94)84268-x.
K31	→3)-β-D-Glc*p*-(1→3)-[[4,6-Pyr]β-D-Glc*p*-(1→2)-α-D-Man*p*-(1→4)]α-D-GlcpA-(1→3)-β-D-Gal*p*-(1→	5	3+2	Cho-Chak, C.; Sui-Lam, W.; Yuen-Min, C. *Carbohydrate Research*, **1979**, *73(1)*, 169–74. http://dx.doi.org/10.1016/s0008-6215(00)85486-5.

K32	→3)-α-D-Gal*p*-(1→2)-[3,4-Pyr]α-L-Rha*p*-(1→3)-β-L-Rha*p*-(1→4)-α-L-Rha*p*-(1→	4	4	Dutton, G.G.S.; Mackie, K.L.; Savage, A.V.; Rierger-Hug, D.; Stirm, S. *Carbohydrate Research*, **1980**, *84(1)*, 161–70. http://dx.doi.org/10.1016/s0008-6215(00)85439-7.
K33	→4)-β-D-Glc*p*-(1→4)-[α-D-Glc*p*A-(1→3)][3,4-Pyr][β-D-Gal*p*-(1→6)]β-D-Man*p*-(1→4)-β-D-Man*p*6OAc-(1→	5	3+1+1	Lindberg, B.; Lindh, F.; Lönngren, J.; Nimmich, W. *Carbohydrate Research*, **1979**, *70(1)*, 135–44. http://dx.doi.org/10.1016/s0008-6215(00)83277-2.
K34	→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Glc*p*-(1→3)-[α-L-Rha*p*-(1→4)]α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→	6	5+1	Joseleau, J.-P.; Michon, F.; Vignon, M. *Carbohydrate Research*, **1982**, *101(2)*, 175–85. http://dx.doi.org/10.1016/s0008-6215(00)80998-2.
K35	→3)-[4,6-Pyr]α-D-Gal*p*-(1→3)-α-D-Man*p*-(1→3)-[β-D-Glc*p*A-(1→2)]α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	5	4+1	Dutton, G.G.S.; Lim, A.V.S. *Carbohydrate Research*, **1985**, *145(1)*, 67–80. http://dx.doi.org/10.1016/s0008-6215(00)90413-0.
K36	→3)-β-D-Gal*p*-(1→3)-[[4,6-Pyr]β-D-Glc*p*-(1→4)-β-D-Glc*p*A-(1→2)]α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→	6	4+2	Dutton, G.G.S.; Mackie, K.L. *Carbohydrate Research*, **1977**, *55(1)*, 49–63. http://dx.doi.org/10.1016/s0008-6215(00)84442-0.
K37	→3)-[4-O-Lac-β-D-Glc*p*A-(1→6)-α-D-Glc*p*-(1→4)]β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→	4	2+2	Lindberg, B.; Lindqvist, B.; Lönngren, J.; Nimmich, W. *Carbohydrate Research*, **1977**, *58(2)*, 443–51. http://dx.doi.org/10.1016/s0008-6215(00)84371-2.

K38	→6)-β-D-Glc*p*-(1→3)-[3-deoxy-L-glycero-pentulosonic acid-(2→3)][β-D-Glc*p*-(1→2)]β-D-Gal*p*-(1→4)-α-D-Gal*p*-(1→	5	3+1+1	Lindberg, B.; Samuelsson, K.; Nimmich, W. *Carbohydrate Research*, **1973**, *30(1)*, 63–70. http://dx.doi.org/10.1016/s0008-6215(00)82173-4.
K39	→3)-β-D-Glc*p*-(1→4)-β-D-Glc*p*A-(1→2)-α-D-Man*p*-(1→4)-β-D-Glc*p*A-(1→2)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	6	6	Nixon Anderson, A.; Parolis, H.; Dutton, G.G.S.; Leek, D.M. *Carbohydrate Research*, **1987**, *167*, 279–90. http://dx.doi.org/10.1016/0008-6215(87)80285-9.
K40	→4)-α-D-Glc*p*A-(1→4)-[β-D-Gal*p*-(1→2)-β-L-Rha*p*-(1→3)]β-D-Gal*p*-(1→2)-α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→	7	5+2	Maddali, U.B.; Ray, A.K.; Roy, N. *Carbohydrate Research*, **1990**, *208*, 59–66. http://dx.doi.org/10.1016/0008-6215(90)80085-h.
K43	→3)-[β-D-Man*p*-(1→4)-β-D-Glc*p*A-(1→2)]α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-α-D-Gal*p*-(1→	5	3+2	Aereboe, M.; Parolis, H.; Parolis, L.A.S. *Carbohydrate Research*, **1993**, *248*, 213–23. http://dx.doi.org/10.1016/0008-6215(93)84128-s.
K44	→3)-β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-β-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→	5	5	Dutton, G.G.S.; Folkman, T.E. *Carbohydrate Research*, **1980**, *78(2)*, 305–15. http://dx.doi.org/10.1016/0008-6215(80)90011-7.
K45	→3)-β-D-Glc*p*-(1→2)-[β-D-Glc*p*A-(1→3)]α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→	5	4+1	Dutton, G.G.S.; Di Fabio, J.L.; Zanlungo, A.B. *Carbohydrate Research*, **1982**, *106(1)*, 93–100. http://dx.doi.org/10.1016/s0008-6215(00)80735-1.

K46	→3)-[β-D-Glc*p*-(1→3)-[4,6-Pyr]β-D-Man*p*-(1→4)]α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→3)-α-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→	6	4+2	Okutani, K.; Dutton, G.G.S. *Carbohydrate Research*, **1980**, *86(2)*, 259–71. http://dx.doi.org/10.1016/s0008-6215(00)85903-0.
K47	→3)-β-D-Gal*p*-(1→4)-[α-L-Rha*p*-(1→4)-β-D-Gc*p*A-(1→3)]α-L-Rha*p*-(1→	4	2+2	Björndal, H.; Linberg, B.; Lönngren, J.; Rossel, K.-G.; Nimmich, W. *Carbohydrate Research*, **1973**, *27(2)*, 373–8. http://dx.doi.org/10.1016/s0008-6215(00)81319-1.
K48	→3)-β-D-Gc*p*-(1→3)-[α-D-Gal*p*A-(1→2)]α-L-Rha*p*-(1→4)-α-D-Glc*p*-(1→2)-α-L-Rha*p*-(1→	5	4+1	Joseleau, J.-P.; Marais, M.-F. *Carbohydrate Research*, **1988**, *179*, 321–6. http://dx.doi.org/10.1016/0008-6215(88)84128-4.
K49	→3)-α-D-Gal*p*-(1→2)-[α-D-Gal*p*OAc-(1→3)]α-D-Man*p*-(1→3)-α-D-Gal*p*-(1→	4	3+1	Joseleau, J.-P. *Carbohydrate Research*, **1985**, *142(1)*, 85–92. http://dx.doi.org/10.1016/s0008-6215(00)90735-3.
K50	→3)-β-D-Gal*p*-(1→3)-[β-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→6)]α-D-Glc*p*-(1→4)-α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→	7	5+2	Altman, E.; Dutton, G.G.S. *Carbohydrate Research*, **1983**, *118*, 183–94. http://dx.doi.org/10.1016/0008-6215(83)88046-x.
K51	→3)-α-D-Gal*p*-(1→3)-[α-D-Glc*p*A-(1→6)-α-D-Glc*p*-(1→4)]α-D-Gal*p*-(1→	4	2+2	Chakraborty, A.K.; Dąbrowski, U.; Geyer, H.; Geyer, R.; Stirm, S. *Carbohydrate Research*, **1982**, *103(1)*, 101–5. http://dx.doi.org/10.1016/s0008-6215(82)80010-4.

K52	→3)-[α-D-Gal*p*-(1→2)]α-D-Gal*p*-(1→4)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→2)-α-L-Rha*p*-(1→4)-β-D-Gal*p*-(1→	6	5+1	Stenutz, R.; Erbing, B.; Widmalm, G.; Jansson, P.-E.; Nimmich, W. *Carbohydrate Research*, **1997**, *302(1–2)*, 79–84. http://dx.doi.org/10.1016/s0008-6215(97)00106-7.
K53	→4)-β-D-Glc*p*A-(1→2)-[α-L-Rha*p*-(1→3)]α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→2)-α-L-Rha*p*-(1→	6	5+1	Dutton, G.G.S.; Paulin, M. *Carbohydrate Research*, **1980**, *87(1)*, 107–17. http://dx.doi.org/10.1016/s0008-6215(00)85195-2.
K54	→2)-α-D-GlcpA-(1→3)-α-L-Fucp2OAc-(1→3)-[β-D-Glcp-(1→4)]β-D-Glcp-(1→	8	3+1	Dutton, G.G.S.; Merrifield, E.H. *Carbohydrate Research*, **1982**, *105(2)*, 189–203. http://dx.doi.org/10.1016/s0008-6215(00)84967-8.
K55	→4)-β-D-Glc*p*-(1→4)-[α-D-Glc*p*A-(1→3)-α-D-Gal*p*-(1→3)]α-L-Rha*p*2OAc-(1→	4	2+2	Bebault, G.M.; Dutton, G.G.S. *Carbohydrate Research*, **1978**, *64*, 199–213. http://dx.doi.org/10.1016/s0008-6215(00)83701-5.
K56	→3)-[4,6-Pyr]β-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→3)-[α-L-Rha*p*-(1→2)]α-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→	5	4+1	Choy, Y.-M.; Dutton, G.G. *Canadian Journal of Chemistry*, **1973**, *51(18)*, 3021–6. http://dx.doi.org/10.1139/v73-450.
K57	→3)-β-D-Gal*p*-(1→3)-[α-D-Man*p*-(1→4)]α-D-Gal*p*A-(1→2)-α-D-Man*p*-(1→	4	3+1	Kamerling, J.P.; Lindberg, B.; Lönngren, J.; Nimmich, W.; Servin, R.; Sternerup, H. *Acta Chemica Scandinavica*, **1975**, *29b*, 593–8. http://dx.doi.org/10.3891/acta.chem.scand.29b-0593.
K58	→3)-α-D-Glc*p*-(1→4)-[2,3-Pyr]β-D-Glc*p*A-(1→4)-[α-D-Gal*p*-(1→3)]α-L-Fuc*p*-(1→	4	3+1	Dutton, G.G.S.; Savage, A.V. *Carbohydrate Research*, **1980**, *84(2)*, 297–305. http://dx.doi.org/10.1016/s0008-6215(00)85559-7.

K59	→3)-β-D-Glc*p*-(1→3)-[β-D-Glc*p*A-(1→4)]β-D-Gal*p*-(1→2)-α-D-Man*p*6OAc-(1→3)-α-D-Man*p*6OAc-(1→	5	4+1	Lindberg, B.; Lönngren, J.; Rudén, U.; Nimmich, W. *Carbohydrate Research*, **1975**, *42(1)*, 83–93. http://dx.doi.org/10.1016/s0008-6215(00)84102-6.
K60	→3)-β-D-Glc*p*-(1→3)-[α-D-Glc*p*-(1→4)]β-D-Glc*p*A-(1→3)-[β-D-Glc*p*-(1→2)]α-D-Gal*p*-(1→3)-[β-D-Glc*p*-(1→2)]α-D-Man*p*-(1→	7	4+1+1+1	Dutton, G.G.S.; Di Fabio, J. *Carbohydrate Research*, **1980**, *87(1)*, 129–39. http://dx.doi.org/10.1016/s0008-6215(00)85197-6.
K61	→4)-β-D-Glc*p*A-(1→2)-[α-D-Gal*p*-(1→3)]α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→	5	4+1	Rao, A.S.; Roy, N. *Carbohydrate Research*, **1979**, *76(1)*, 215–24. http://dx.doi.org/10.1016/0008-6215(79)80020-8.
K62	→4)-α-D-Glc*p*-(1→2)-β-D-Glc*p*A-(1→2)-[α-D-Man*p*-(1→3)]α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→	5	4+1	Dutton, G.G.S.; Mo-Tai, Y. *Carbohydrate Research*, **1977**, *59(1)*, 179–92. http://dx.doi.org/10.1016/s0008-6215(00)83304-2.
K63	→3)-α-D-Gal*p*-(1→3)-α-D-Gal*p*A-(1→3)-α-L-Fuc*p*-(1→	3	3	Joseleau, J.-P.; Marais, M.-F. *Carbohydrate Research*, **1979**, *77(1)*, 183–90. http://dx.doi.org/10.1016/s0008-6215(00)83804-5.
K64	→4)-α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→4)-[[4,6-Pyr]β-D-Glc*p*-(1→2)][α-L-Rha*p*(1→3)]α-D-Man*p*-(1→	6	4+2	Merrifield, E.H.; Stephen, A.M. *Carbohydrate Research*, **1979**, *74(1)*, 241–57. http://dx.doi.org/10.1016/s0008-6215(00)84780-1.
K66	→3)-α-D-Man*p*-(1→3)-α-D-Gal*p*-(1→2)-[4-O-Lac-β-D-Glc*p*-(1→3)]α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→	5	4+1	Jansson, P.-E.; Lindberg, B.; Lönngren, J.; Ortega, C.; Nimmich, W. *Carbohydrate Research*, **1984**, *132(2)*, 297–305.

				http://dx.doi.org/10.1016/0008-6215(84)85226-x.
K67	→3)-α-L-Rha*p*-(1→3)-[α-L-Rha*p*-(1→4)-[β-D-Gal*p*-(1→3)]β-D-Glc*p*A-(1→2)]α-D-Man*p*-(1→3)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	7	4+3	Dutton, G.G.S.; Karunaratne, D.N. *Carbohydrate Research*, **1983**, *119*, 157–69. http://dx.doi.org/10.1016/0008-6215(83)84054-3.
K68	→2)-[[4,6-Pyr]-α-D-Man*p*-(1→4)]α-D-Gal*p*A-(1→2)-α-D-Man*p*-(1→3)-β-D-Gal*p*-(1→	4	3+1	Dutton, G.G.S.; Parolis, H.; Parolis, L.S. *Carbohydrate Research*, **1986**, *152*, 249–59. http://dx.doi.org/10.1016/s0008-6215(00)90305-7.
K69	→4)-β-D-Glc*p*-(1→4)-[α-D-Glc*p*A-(1→3)][[4,6-(R)-Pyr]β-D-Gal*p*-(1→6)]β-D-Man*p*-(1→4)-β-D-Man*p*6OAc-(1→	5	3+1+1	Hackland, P.L.; Parolis, H.; Parolis, L.A.S. *Carbohydrate Research*, **1988**, *172(2)*, 209–16. http://dx.doi.org/10.1016/s0008-6215(00)90855-3.
K70	→4)-β-D-Glc*p*-(1→4)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→2)-α-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→2)-[4,6-Pyr]α-L-Rha*p*-(1→	6	6	Dutton, G.G.S.; Mackie, K.L. *Carbohydrate Research*, **1978**, *62(2)*, 321–35. http://dx.doi.org/10.1016/s0008-6215(00)80879-4.
K71	→3)-α-L-Rha*p*-(1→3)-β-D-Glc*p*-(1→2)-[β-D-Glc*p*-(1→2)-[β-D-Glc*p*A-(1→3)]α-L-Rha*p*-(1→3)]α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→	7	4+3	Jackson, G.E.; Ravenscroft, N.; Stephen, A.M. *Carbohydrate Research*, **1990**, *200*, 409–28. http://dx.doi.org/10.1016/0008-6215(90)84207-b.

K72	→3)-β-D-Glc*p*-(1→3)-α-L-Rha*p*-(1→2)-[3,4-Pyr]-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→	4	4	Choy, Y.-M.; Dutton, G.G. *Canadian Journal of Chemistry*, **1974**, *52(4)*, 684–7. http://dx.doi.org/10.1139/v74-106.
K74	→3)-β-D-Gal*p*-(1→2)-[[4,6-Pyr]-β-D-Gal*p*-(1→4)-α-D-Glc*p*A-(1→3)]-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→	5	3+2	Dutton, G.G.S.; Paulin, M. *Carbohydrate Research*, **1980**, *87(1)*, 119–27. http://dx.doi.org/10.1016/s0008-6215(00)85196-4.
K79	→3)-β-D-Gal*p*-(1→3)-[α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)]β-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→	7	5+2	Dutton, G.G.S.; Lim, A.V.S. *Carbohydrate Research*, **1985**, *144(2)*, 263–76. http://dx.doi.org/10.1016/s0008-6215(00)90674-8.
K80	→3)-β-D-Gal*p*-(1→2)-[[3,4-Pyr]-β-L-Rha*p*-(1→4)-α-D-Glc*p*A-(1→3)]α-D-Man*p*-(1→2)-α-D-Man*p*-(1→	5	3+2	Dutton, G.G.S.; Karunaratne, D.N. *Carbohydrate Research*, **1984**, *134(1)*, 103–14. http://dx.doi.org/10.1016/0008-6215(84)85026-0.
K81	→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→4)-β-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→	6	6	Curvall, M.; Lindberg, B.; Lönngren, J.; Nimmich, W. *Carbohydrate Research*, **1975**, *42(1)*, 73–82. http://dx.doi.org/10.1016/s0008-6215(00)84101-4.
K82	→3)-β-D-Glc*p*-(1→3)-[β-D-Glc*p*A-(1→4)]α-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→	4	3+1	Dutton, G.G.S.; Lim, A.V.S. *Carbohydrate Research*, **1983**, *123(2)*, 247–57. http://dx.doi.org/10.1016/0008-6215(83)88481-x.