Structural and Conformational Studies of Oligo- and Polysaccharides

Mona Zaccheus
“What day is it?” asked Pooh.
"It's today," squeaked Piglet.
"My favourite day," said Pooh.”
— A.A. Milne
Abstract

The focus of this thesis is to examine the structural properties of polysaccharides produced by bacteria, as well as the dynamic and conformational behavior of a synthetically derived oligosaccharide.

The primary structures of the O-polysaccharide repeating units of four different *Escherichia coli* (*E. coli*) strains, namely O175, O177, O103 and TD2158, as well as the first report of a capsular polysaccharide produced by lactic acid bacteria *Leuconostoc mesenteroides* ssp. *cremoris* PIA2 are reported in paper I–V. Structural analyses have been performed using a combination of nuclear magnetic resonance spectroscopy and chemical component analysis.

The elucidated structures in paper I–III, as well as paper V, are composed of linear repeating units of varying composition and length. In paper IV, the structure of the O-polysaccharide repeating unit of *E. coli* TD2158 is determined to be a branched hexasaccharide structure with a heterogeneous substitution pattern, with either a β-D-GlcNAc or β-D-Glc residue branching to the backbone chain. Incubation with bacteriophage HK620 tailspike protein shows that the polysaccharide is selectively cleaved at the α-D-GlcNAc-(1→2)-α-L-Rhap-linkage of the backbone chain, yielding a 9:1 ratio of β-D-GlcNAc/β-D-Glcp containing hexasaccharides after digestion.

In paper VI the conformational properties of a trisaccharide, which constitutes an internal epitope of the Le^a^Le^b^ hexasaccharide over-expressed on the surface of squamous lung cancer cells, have been analyzed using NMR spectroscopy and molecular dynamics simulations. The β-(1→3)-linkage of the trisaccharide was shown to be highly flexible.
List of publications

This thesis is based on the following publications, referred to in the text by their Roman numerals I-VI. Reprints were made with the kind permission of the publisher.

I. Structural elucidation of the O-antigenic polysaccharide from \textit{Escherichia coli} O175
Mona V. Svensson, Andrej Weintraub and Göran Widmalm

II. Structural studies of the O-antigenic polysaccharide from \textit{Escherichia coli} O177
Mona V. Svensson, Andrej Weintraub and Göran Widmalm

III. Genetic and structural relationships of \textit{Salmonella} O55 and \textit{Escherichia coli} O103 O-antigens and identification of a 3-hydroxybutanoyltransferase gene involved in the synthesis of a Fuc3N derivative
Bin Liu §, Andrei V. Perepelov §, Mona V. Svensson §, Sergei D. Shevelev, Dan Guo, Sof'ya N. Senchenkova, Alexander S. Shashkov, Andrej Weintraub, Lu Feng, Göran Widmalm, Yuriy A. Knirel and Lei Wang
\textit{Glycobiology}, \textbf{2010}, 20, 679 – 688

IV. Structural studies of the O-antigen polysaccharide from \textit{Escherichia coli} TD2158 having O18 serogroup specificity and aspects of its interaction with the tailspike endoglycosidase of the infecting bacteriophage HK620
Mona V. Zaccheus†, Nina Lorenzen, Magnus Lundborg, Charlotte Utrecht, Stefanie Barbirz and Göran Widmalm
\textit{Submitted for publication}

§ Authors contributed equally to this work
† née Svensson
V. Structural Studies of the Capsular Polysaccharide Produced by *Leuconostoc mesenteroides* ssp. *cremoris* PIA2
Mona V. Svensson, Xue Zhang, Eine Huttunen, and Göran Widmalm
*Biomacromolecules*, **2011**, 12, 2496 – 2501

VI. Conformational dynamics of a central trisaccharide fragment of the Le"Le" tumor associated antigen studied by NMR spectroscopy and molecular dynamics simulations
Mona V. Zaccheus†, Robert Pendrill, Trudy A. Jackson, An Wang, France-Isabelle Auzanneau and Göran Widmalm
*Submitted for publication*

Related papers by the author, not included in the thesis:

Complete $^1$H and $^{13}$C NMR chemical shift assignments of mono-, di-, and trisaccharides as basis for NMR chemical shift predictions of polysaccharides using the computer program CASPER
Mattias U. Roslund, Elin Säwén, Jens Landström, Jerk Rönnols, K. Hanna M. Jonsson, Magnus Lundborg, Mona V. Svensson, and Göran Widmalm

*Brucella melitensis* 16M produces a mannan and other extracellular matrix components typical of a biofilm
*FEMS Immunology and Medical Microbiology*, **2010**, 59, 364 – 377
Table of contents

Abstract .................................................................................................................................. v

List of publications .............................................................................................................. vii

1 Introduction .......................................................................................................................... 1
  1.1 Introduction to carbohydrates ....................................................................................... 1
  1.2 Bacterial polysaccharides ............................................................................................ 2
  1.3 Biosynthesis of O-polysaccharides .............................................................................. 3
    1.3.1 Glycosyltranferases ............................................................................................ 3
    1.3.2 Transportation across the plasma membrane ...................................................... 4
    1.3.3 The Wzx/Wzy-dependent pathway .................................................................... 4
    1.3.4 The ABC-transporter dependent pathway ......................................................... 5
  1.4 Gram-negative Escherichia coli ................................................................................... 5

2 Experiments for structural determination ........................................................................ 7
  2.1 Introduction to NMR spectroscopy .............................................................................. 7
    2.1.1 Scalar couplings in carbohydrates ................................................................... 8
    2.1.2 NMR experiments ............................................................................................ 9
  2.2 Chemical component analysis ...................................................................................... 10

3 Carbohydrate conformational analysis .......................................................................... 12
  3.1 Carbohydrate conformation ....................................................................................... 12
  3.2 Molecular dynamics simulations in conformational analysis ................................... 13
  3.3 NMR measurements .................................................................................................. 14
    3.3.1 Nuclear Overhauser Effect .............................................................................. 14
    3.3.2 Trans-glycosidic coupling constants ................................................................. 16

4 Structural studies of the O-polysaccharides from diarrheal Escherichia coli O175 and O177 (Papers I and II) ......................................................................................... 17
  4.1 E. coli O175 (Paper I) ................................................................................................. 17
    4.1.1 Result and discussion ....................................................................................... 17
  4.2 E. coli O177 (Paper II) ............................................................................................... 21
    4.2.1 Results ............................................................................................................ 21
  4.3 Conclusions ................................................................................................................ 23

5 Structural and genetic relationships of Escherichia coli O103 and Salmonella enterica O55 (Paper III) ........................................................................................................... 24
Abbreviations

BASHD  Band-selective, homonuclear-decoupled
BS-CT  Band-selective constant-time
CPS    Capsular polysaccharide
CSSF   Chemical shift selective filter
E. coli Escherichia coli
EAEC   Enteroaggregative E. coli
EHEC   Enterohemorrhagic E. coli
EPS    Exopolysaccharide
ER     Endoplasmic reticulum
ETEC   Enterotoxigenic E. coli
FID    Free induction decay
GC     Gas-liquid chromatography
H2BC   Heteronuclear two-bond correlation
HETCOR Heteronuclear correlation
HMBC   Heteronuclear multiple-bond correlation
HSQC   Heteronuclear single-quantum correlation
ISPA   Isolated spin-pair approximation
J-HMBC J-resolved heteronuclear multiple-bond correlation
LAB    Lactic acid bacteria
LPS    Lipopolysaccharide
mAb    Monoclonal antibodies
MD     Molecular dynamics
Me     Methyl
NAc    N-acetyl
NMR    Nuclear magnetic resonance
NOE    Nuclear Overhauser effect
STEP   Selective TOCSY edited preparation
TACA   Tumor-associated carbohydrate antigen
TOCSY  Total correlation spectroscopy
TSP    Sodium 3-trimethylsilyl-(2,2,3,3-2H4)-propanoate
Und-P  Undecaprenyl phosphate
1 Introduction

1.1 Introduction to carbohydrates

The many and diverse roles of carbohydrates in nature, from structural components to cell-signaling and energy storage, not only make them the most abundant biological molecules, but also give a hint to the great variety of their combinations. With the possibility to form linear or branched oligo- or polysaccharides and further decorating them with non-carbohydrate substituents, like acetates or phosphates, a multitude of combinations can be created from only a handful of monosaccharides.

Carbohydrates are polyhydroxylated aldehydes and ketones termed aldoses and ketoses, respectively. The aldoses are the most common type of structures and the focus of this thesis, see Figure 1.1. The carbohydrates in open-chain form are composed of a carbonyl group, a number of secondary alcohols and, in many cases, a primary alcohol. Monosaccharides that are not hydroxylated at all carbons in the chain are denoted with a deoxy function. All of the carbons in the secondary alcohols are chiral centers and the number of carbons and their configuration determines the carbohydrate identity.

*Figure 1.1* Carbohydrate monosaccharides discussed in this thesis, all in the \(^4\text{C}_1\) conformation: a) β-D-Fucp3NHb, b) α-D-GalpNAc, c) α-D-GlcpA

Carbohydrates containing five or six carbons are denoted pentoses and hexoses, respectively, and form furanose or pyranose rings, depending on if whether the cyclic hemiacetal is created with O4 or O5. In the cyclic form the carbonyl carbon becomes sterogenic and two configurations are possible, α and β. The most stable conformations for many carbohydrate residues in solution is the chair conformation \(^4\text{C}_1\) or \(^1\text{C}_4\), where superscript and subscript numbering denotes above and below the plane of the chair, respectively. In
these conformations the hydroxyl group attached to C1 is positioned axial in the α- and equatorial in the β-configuration. Monosaccharides can adopt two absolute configurations, D or L, where the descriptor is related to the optical chirality of glyceraldehyde and determined by the configuration of the highest numbered chiral carbon in the chain. The L-hexoses have the opposite configurations to the D-hexoses at all chiral centers. The most abundant type of carbohydrates in nature is with the D-configuration; however some L-configured carbohydrates are more common than others, like L-rhamnose and L-fucose.

1.2 Bacterial polysaccharides

The bacterial surface is to a large extent covered by polysaccharides conjugated to proteins (e.g. glycoproteins) and lipids (e.g. glycolipids). Surface glycoconjugates attached to the bacterial cell wall are essential both for the cell wall structure as well as the signaling of the bacterium. As the first means of contact between the bacterium and the environment, the glycoconjugates also work as protection against host immune defense.

The definition of Gram-positive or Gram-negative bacteria refers to a staining method, named after Hans Christian Gram, where the bacterial species are colored differently. Gram-positive bacteria have a thick peptidoglycan cell wall while Gram-negative bacteria have a double membrane, made up of an inner peptidoglycan cell wall and an outer membrane containing mainly lipopolysaccharides (LPS). Some microorganisms produce extracellular polysaccharides, also referred to as exopolysaccharides (EPS), either in the form of a capsule that envelopes the cell, or as excreting material.

LPS, also known as endotoxins, are attached to the outer layer of Gram-negative bacteria and protrude from the cell surface. The LPS covers approximately 75% of the surface of the bacterium and can be divided into three parts; the lipid A, the core and the O-polysaccharide (also referred to as the O-antigen), see Figure 1.2.

![Figure 1.2 Schematic structure of the lipopolysaccharide.](Image)

In small amounts, the LPS can stimulate the immune system and protect the host from other infections. In larger amounts, however, the LPS of a pathogenic strain can cause high fever, increased heart rate and lead to septic
The lipid A is the hydrophobic component of the LPS and consists of glucosamine phosphate and fatty acids. The lipid A anchors the LPS to the outer membrane of the bacterial cell wall, and is structurally conserved in different Gram-negative bacteria. The endotoxic response against the LPS is caused by the lipid A and the toxicity is related to the response from the immune system and amount and length of the fatty acids.

The core consists of non-repetitive oligosaccharides and is divided into the inner and the outer core. The inner core of the LPS is structurally alike for Escherichia coli, Salmonella enterica and Shigellae strains, and connects to the lipid A by 3-deoxy-α-D-manno-oct-2-ulopyranosonic acid (Kdo). The Kdo is present in all known core regions, and is in most cases coupled to three L-glycero-D-manno-heptospyranoses (L,D-Hep). The structure of the outer core is not as structurally conserved as the inner core, and is branched with Glcp, Galp and their derivatives. The outer core connects to the O-antigen, and both the core and the O-antigen repeating units are synthesized on the cytoplasmic side of the inner membrane.

The O-polysaccharide is specific for each serogroup and consists of repeating oligosaccharides, up to 60 units of 2-8 carbohydrate residues. It is recognized by the host organism immune system, but also helps to shield the bacterial cell and protects the bacteria from the effect of antibiotics. The diversity of O-antigen repeating units is large; the units can be branched or linear, the carbohydrate residues can differ in type, connectivity and stereochemistry. In addition, substituents like O- and N-acyl groups and phosphate groups are common. During biosynthesis of the O-polysaccharide, subunits of varying length are synthesized and added to the growing polysaccharide chain.

1.3 Biosynthesis of O-polysaccharides

The polysaccharides are assembled in the cytosol and exported to the periplasmic side of the inner membrane. Once full-length polysaccharides are polymerized, they are transported through the periplasm and across the outer membrane to the cell surface. The length of the polysaccharides as well as the glycosyl residues is determined by how the saccharide is assembled and exported from the cytosol to the periplasm. The polysaccharides are transferred from the lipid carrier to the lipid A core.

1.3.1 Glycosyltransferases

Glycosyltransferases are a large family of enzymes that assemble carbohydrate monosaccharide units into poly- or oligosaccharide chains. A limited number of different biosynthesis and assembly pathways are responsible for the large structural diversity in bacterial glycoconjugates, indicating the im-

1.3 Biosynthesis of O-polysaccharides

The polysaccharides are assembled in the cytosol and exported to the periplasmic side of the inner membrane. Once full-length polysaccharides are polymerized, they are transported through the periplasm and across the outer membrane to the cell surface. The length of the polysaccharides as well as the glycosyl residues is determined by how the saccharide is assembled and exported from the cytosol to the periplasm. The polysaccharides are transferred from the lipid carrier to the lipid A core.

1.3.1 Glycosyltransferases

Glycosyltransferases are a large family of enzymes that assemble carbohydrate monosaccharide units into poly- or oligosaccharide chains. A limited number of different biosynthesis and assembly pathways are responsible for the large structural diversity in bacterial glycoconjugates, indicating the im-

1.3 Biosynthesis of O-polysaccharides

The polysaccharides are assembled in the cytosol and exported to the periplasmic side of the inner membrane. Once full-length polysaccharides are polymerized, they are transported through the periplasm and across the outer membrane to the cell surface. The length of the polysaccharides as well as the glycosyl residues is determined by how the saccharide is assembled and exported from the cytosol to the periplasm. The polysaccharides are transferred from the lipid carrier to the lipid A core.

1.3.1 Glycosyltransferases

Glycosyltransferases are a large family of enzymes that assemble carbohydrate monosaccharide units into poly- or oligosaccharide chains. A limited number of different biosynthesis and assembly pathways are responsible for the large structural diversity in bacterial glycoconjugates, indicating the im-
portance of specific and diverse enzymes to catalyze the polymerization of the polysaccharide chains.\textsuperscript{11}

The glycosyltransferases are located in the translocon, which is a cluster of proteins, in the endoplasmic reticulum (ER) membrane and the Golgi apparatus. They transfer nucleotide sugar substrates (e.g. UDP-Glc, UDP-Gal, GDP-Fuc) to the acceptors, and generally work in sequence so that the product of the work of one glycosyltransferase is the acceptor substrate for another.

1.3.2 Transportation across the plasma membrane

Three pathways for transportation of the saccharide chains to the periplasmic side of the membrane are known to date\textsuperscript{10}; the Wzx/Wzy-dependent pathway, the ABC transporter-dependent pathway and the synthase dependent pathway. The synthase-dependent pathway is not discussed further in this thesis.

For the Wzx/Wzy-dependent and the ABC-transporter pathway, the saccharide chains are synthesized from sugar nucleotides and assembled on the membrane-bound carrier lipid, undecaprenyl phosphate (Und-P), on the cytosolic side of the plasma membrane.\textsuperscript{12} Both pathways lead to lipid-bound polysaccharide chains located on the periplasmic side of the inner membrane, however the steps in which the chains are elongated and transported differ.

1.3.3 The Wzx/Wzy-dependent pathway

In the Wzx/Wzy-dependent pathway the enzyme WecA homolog catalyzes the production of Und-PP-GlcNAc, and in some cases Und-PP-GalNAc, which act as a primer for chain elongation.\textsuperscript{10} Individual polysaccharide repeating units linked to Und-PP-GlcNAc are assembled and exported across the membrane by Wzx flippase. Polymerization by one repeating unit at a time to the reducing end is performed by Wzy polymerase, on the periplasmic side of the plasma membrane. The outcome of the blockwise elongation of the polysaccharide chains is a GlcNAc, or GalNAc, residue in each repeating unit.
Figure 1.3 Schematic picture the synthesis and export of bacterial polysaccharides using the Wzx/Wzy-dependent and the ABC transporter-dependent pathway. Figure adapted from Cuthbertson et al. \(^{12}\)

1.3.4 The ABC-transporter dependent pathway

In the ABC-transporter dependent pathway the complete polysaccharide chain linked to Und-PP-GlcNAc is polymerized on the cytosolic side of the membrane by addition of glycosyl residues to the nonreducing end of the growing chain.\(^{13}\) After polymerization the lipid-bound polysaccharide is transported across the membrane through the ABC transporter. In bacteria the ABC transporters are involved in both uptake of nutrients as well as export of diverse molecules (e.g. lipids, proteins, oligo- and polysaccharides).\(^{12}\) Polysaccharides polymerized with the ABC-transporter dependent pathway have only one GlcNAc residue per chain.

1.4 Gram-negative *Escherichia coli*

*Escherichia coli* (*E. coli*) is a Gram-negative bacterium commonly found in the lower intestines of humans and animals. Most strains of *E. coli* are avirulent and exist in symbiotic relation with their host, but some strains have acquired different sets of virulence genes.\(^{14}\) Infection with pathogenic *E. coli* strains can cause three general syndromes: sepsis/meningitis, urinary tract infection and enteric/diarrheal disease.

Characterization of *E. coli* is made by virotyping and serotyping. Virotyping describes how the bacterium attaches to, and how this affects, host cells. It also describes if toxins are produced and how invasive the bacterium is. There are six diarrheagenic virotypes, all affecting the host cells differently, with the primary difference in how these strains adhere to the host cells and by their shared virulence. The different strains are: i) enterotoxigenic *E. coli* (ETEC), ii) enteroaggregative *E. coli* (EAEC), iii) diffusely adherent *E. coli* (DAEC), iv) enteropathogenic *E. coli* (EPEC), v) enterohemorrhagic *E. coli*
(EHEC) and vi) enteroinvasive (EIEC) *E. coli*. Two virotypes of *E. coli* are discussed in this thesis; EAEC (paper I) and EHEC (papers II and III) strains.

The serotype classification used for bacteria is based on the cell surface antigens. There are three fundamental serotyping antigens; the LPS (O-antigen), the polysaccharide capsule (K-antigen) and the protein flagella (H-antigen). The identification of an *E. coli* strain is normally made by both the O- and the H-antigens, and the K-antigen if present. To this date, 174 *E. coli* serogroups have been identified and more than 100 of the structures determined. Further information regarding the determined structures, as well as the serogroups where the glycosyltransferases coding for the O-antigen have been determined but the structure has not, can be found in ECODAB. With the results from this thesis, three additional structures are contributed to the structure determined *E. coli* serogroups. The fourth *E. coli* structure elucidated, that of *E. coli* TD2158 (paper IV), has not yet been serologically determined.
2 Experiments for structural determination

2.1 Introduction to NMR spectroscopy

Defining the structure of compounds, synthesized as well as of unknown biological material, is one of the challenges in chemistry. A valuable, non-destructive and reasonably sensitive technique for structural and conformational determination is nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy is built upon the principle that the magnetic spin of each nucleus will align and precess at a frequency specific for the nucleus, the Larmor frequency, in an applied external magnetic field. When applying an electromagnetic pulse at the Larmor frequency of the desired nucleus, the spins will absorb the energy and be perturbed.

All nuclei have a nuclear spin quantum number, \( I \). If the spin quantum number is not zero, the nuclei can be studied by NMR spectroscopy. For spin-\( \frac{1}{2} \) nuclei, like \(^1\)H and \(^{13}\)C, placed in a strong magnetic field there are two possible spin states, \( +\frac{1}{2} \) and \( -\frac{1}{2} \), often referred to as the \( \alpha \)- and \( \beta \)-spin states. The two energetic states are separated by an energy \( \Delta E \), which is the energy required to be absorbed to change the spin state from \( \alpha \) to \( \beta \). The energy difference is dependent on the magnetogyric constant, \( \gamma \), of the nucleus and the applied magnetic field. Larger \( \gamma \)-nuclei have a larger energy difference and is more sensitive in NMR spectroscopy. \(^1\)H has one of the largest magnetogyric constants of nuclei frequently studied by NMR spectroscopy.

NMR is measured by the perturbation of the equilibrium bulk magnetization created in a strong magnetic field by applying electromagnetic radiation at the frequency of the nucleus. The energy that is absorbed by the nuclei is radiated back at the resonance frequency of each nucleus. The result of the experiment is a spectrum, with lines appearing at the chemical shift of the nucleus. The chemical shift scale is given in parts per million (ppm) instead of in frequency as to make it possible to compare spectra produced at different magnetic fields. The power, duration and amount of the electromagnetic pulses, combined with the delays in between them, are manipulated to satisfy the conditions of the experiment.
2.1.1 Scalar couplings in carbohydrates

Scalar couplings, also called $J$-couplings or spin-spin-couplings, is caused by the magnetic interactions of one nucleus with its surrounding nuclei, and the result is a splitting of the line based on the spin quantum number of the coupled nucleus. When referring to a specific coupling it is customary to specify over how many bonds the coupling $J$ is, in superscript, and also which atoms are involved, in subscript. Scalar couplings are measured in Hz, the size of the coupling related to the electron distribution around the nuclei. The Karplus-type equations\(^{20}\), which are discussed further in section 3.3.2, relates the angle between the scalar coupled nuclei to the size of the coupling.

The anomic configuration of hexopyranosyl residues can often be determined from the scalar coupling; the $^3J_{\text{H1,H2}}$ for the glucol/galacto-configuration is $\sim$4 Hz for $\alpha$ and $\sim$8 Hz for $\beta$. In the manno-configuration, however, both $\alpha$- and $\beta$-configuration give rise to small couplings (1.6 and 0.8 Hz, respectively).\(^{21}\) Heteronuclear coupling constants also give an indication of the anomic configuration; 170–177 Hz is the approximate coupling for $^1J_{\text{C1,H1}}$ in the $\alpha$-configuration in hexopyranosyl residues, while the $\beta$-configuration gives rise to a coupling around 159–168 Hz. The difference of about 10 Hz between $\alpha$- and $\beta$-configurations makes it possible to distinguish them from each other. Both $\alpha$- and $\beta$-furanosyl residues give rise to couplings in around 173–175 Hz, and cannot be determined using the $^1J_{\text{C1,H1}}$ coupling.

When the coupling constant between two nuclei is in the same order of magnitude as the chemical shift difference between them, they are said to share a strong coupling, causing the line splitting to be smaller than the actual coupling constant. In paper VI, strong coupling between $^3J_{\text{H3,H4}}$ hinders the measurement of a trans-glycosidic heteronuclear coupling involving H4. The strong coupling phenomena further complicate the carbohydrate spectra, making the determination of the exact coupling constants and chemical shifts difficult. Computational NMR prediction and spectral simulation programs, where a defined spin-system, the experimental NMR spectrum and the approximate chemical shifts as well as coupling constants of the molecule are introduced, can help determine the correct couplings and chemical shifts of strongly coupled systems. A simulated spectrum is created and fitted to the experimental spectrum and when the spectra are identical the exact coupling constants and chemical shifts can be extracted. An example of a simulated and experimental spectrum using the spin simulation program PERCH\(^{22}\) is shown in Figure 8.3.
2.1.2 NMR experiments

Carbohydrate monosaccharide structures give rise to rather complicated NMR spectra. Polysaccharides are much larger molecules than the monosaccharides and tumble slower in solution, causing the excited spins to relax more rapidly through spin-spin relaxation ($T_2$), which in turn results in broad resonances. Due to the line broadening, spectra of polysaccharides are even more crowded than the monosaccharides. In addition, remains of the core and RNA from incomplete purification often further complicate the spectra. The carbohydrate ring protons resonate at 3.2–4.2 ppm, while the anomic signals appear at 4.4–5.6 ppm.\textsuperscript{23} N-acetyl and O-acetyl groups are found at ~2.0 and ~2.1 ppm, respectively, while the 6-deoxysugar methyl groups resonate at ~1.1–1.3 ppm. Information about the ratio N-acetyl, O-acetyl or 6-deoxysugar methyl groups can be gathered by integration relative to the anomic peaks. It is also possible to gain information about the carbohydrate ring size by the $^{13}$C chemical shift, which in glycosylated pyranoses resonates at 95–105 ppm and in furanoses 105–110 ppm, approximately.\textsuperscript{24} In addition to 1D $^1$H and $^{13}$C experiments, there are several NMR experiments that are of great value when performing a structural investigation of synthetic as well as biological material. The experiments listed here, save from the BS-CT-HMBC, have been used in the assignments of all structures in paper I–VI, as well as for the structures given in Table A1 and A2 in the Appendix.

A decoupled multiplicity-edited $^1$H,$^{13}$C-HSQC, which reveals the carbon/proton pairs and at the same time quickly identifies the CH$_2$-groups, is a good starting point.\textsuperscript{25} To extract the $^1J_{C,H}$, which can give further insight to the anomic configuration, it is convenient to run a coupled HSQC as well. If the carbons have similar chemical shifts, a $^{13}$C,$^1$H-HETCOR gives better resolution in the carbon dimension, however, it is a more time-consuming experiment.\textsuperscript{26}

Each carbohydrate residue is a unique spin system, and the 2D $^1$H,$^1$H-TOCSY is therefore a suitable experiment to deduce the chemical shifts of the ring protons.\textsuperscript{27,28} Increasing the mixing time allows the magnetization to be transferred through-out the ring, thus making it possible to assign the proton/carbon pairs found in HSQC. Commonly, five mixing times are used in structural determination of polysaccharides: 10, 30, 60, 90 and 120 ms. The shortest mixing time usually shows the correlation to the neighboring proton, while the longest mixing time allows for the magnetization to reach the end of the chain.

In some carbohydrate residues, the $^3J_{H,H}$ is too small for the magnetization to be transferred by the TOCSY sequence. This is the case with the $^3J_{H_4,H_5}$ of the GalpNAc residue in paper I, for example. In these cases, the $^1$H,$^1$H-NOESY is of help, as it will give crosspeaks for protons close in space.\textsuperscript{29} This way, the sub-spin systems found in TOCSY can be linked to together.
Another way to assign the connected ring atoms is by using the H2BC experiment, which gives heteronuclear correlations separated by two bonds.\textsuperscript{30} It is a valuable experiment in the instances that the chemical shifts of protons are close and the assignments of them cannot be carried out with the TOCSY experiment.

The chemical shift for the carbons and protons in the residues are compared to those of the corresponding monosaccharides.\textsuperscript{31} A large chemical shift difference indicates a glycosylation position, and is referred to as a glycosylation shift.

When the carbohydrate residues have been identified, the linkage positions can be confirmed by the $^1$H,$^{13}$C-HMBC experiment.\textsuperscript{32} This experiment gives heteronuclear correlations over two to three bonds, which makes it a suitable experiment to reveal the glycosidic linkage positions.

Furthermore, in the cases where more than one N-acyl group is identified or two carbonyl signals are very close in $^{13}$C chemical shift, a band-selective constant-time HMBC (BS-CT-HMBC) can be of great help as it makes it easier to identify which methyl protons correlate to which carbonyl carbon.\textsuperscript{33} A regular HMBC would produce a spectrum with large peak overlap, but using a carbon pulse, which excites only the carbonyl region, increases the resolution. However, this reveals the homonuclear $J$-couplings, which are not a problem in low-resolved HMBC, creating a spectrum with severely tilted cross-peaks in the F\textsubscript{1}-dimension. When allowing the long-range couplings to evolve for a constant time period a fully resolved spectrum with nice peak shapes is produced.

\section*{2.2 Chemical component analysis}

To gain more information about the polysaccharide there are several sugar component analysis experiments that are of value. In this thesis sugar analysis and absolute configuration determination is used.

Sugar analysis is performed to identify the carbohydrate residues in the polysaccharide repeating unit, and the relative ratio of the different residues. Sugar analysis using alditol acetates is the common approach. The polysaccharide is first hydrolyzed to monosaccharides using trifluoroacetic acid (TFA), followed by reduction of the carbonyl groups with sodium borohydride ($\text{NaBH}_4$) and acetylation of all hydroxyl groups using acetic anhydride in pyridine. The compounds are then separated by gas-liquid chromatography (GLC) and their retention times are compared to those of standards. The ratio of the carbohydrate residues may be misrepresented, since the ratio of one residue is dependent on the hydrolysis of the residue next to it. For example, uronic acids hydrolyze slowly so the residue coupled to it will have a lower ratio. Partial N-deacetylation of 2-acetamido-2-deoxy-hexoses occurs in the hydrolysis reaction, and the resulting 2-aminosugars do not hydrolyze
at all. The result is an under-representation of both the 2-acetamido-2-deoxy-
hexose and the residue they are linked to.

If the repeating unit contains a uronic acid, as in paper I, sugar analysis
using alditol acetates is not convenient since acidic carbohydrates cannot be
observed. Consequently, sugar analysis with residues containing acid moi-
eties is performed by converting the uronic acids to methyl glycosides methyl
esters. This method also starts with hydrolysis using TFA to give the mono-
saccharide residues, followed by acid solvolysis using methanol and small
amounts of acetyl chloride, which yields methyl glycosides with an ester
function instead of the acid. Subsequent acetylation with acetic anhydride
and pyridine gives fully acetylated methyl glycosides methyl esters which
are then separated by GC.

To determine if the carbohydrate residues are of D- or L- configuration
they can be derivatized with an optically active alcohol, thus creating a new
chiral center. The procedure starts with TFA-hydrolysis of the glycosidic
bonds in the repeating unit. If N-acetyl carbohydrate residues are present,
they are carefully re-acetylated with acetic anhydride followed by reaction of
the monosaccharides with acid and an optically active alcohol, in this case
(S)-(+)-2-BuOH. The residues are then acetylated and separated by GLC and
compared to standards. The diastereomers that are created, one of two possi-
ble for each residue, have different physical properties and therefore will
separate by GC, thus making it possible to determine the absolute configura-
tion of the carbohydrate residue.

If other components are present, as for example the 3-hydroxybutanoic
acid in paper II, appropriate methods like gas chromatography-mass spec-
trometry (GC-MS) or liquid chromatography-mass spectrometry (LC-MS)
are commonly used to identify them.
3 Carbohydrate conformation analysis

3.1 Carbohydrate conformation

Carbohydrate recognition is essential in many biological functions. To understand the mechanism for their interaction, knowledge of the carbohydrate three-dimensional structure is of vast importance. The 3D structure of carbohydrates is described by its configuration as well as conformation, the latter which is defined mainly by the spatial arrangement of the atoms involved in the glycosidic linkages. Three torsion angles are of importance for the conformation of oligosaccharides, namely the $\phi$ (phi), $\psi$ (psi) and $\omega$ (omega) torsion angles, pictured in Figure 3.1.

The torsion angles in this thesis are defined by the non-IUPAC nomenclature where the $\phi$ torsion angle is described by H1–C1–O–C_{aglycon} and the $\psi$ torsion angle is defined as C1–O–C_{aglycon}–H_{aglycon}. The $\phi$ torsion angle is influenced by steric and the $\epsilon$-exo-anomeric effect, which is an electronic effect that describes the ability of the lone pairs of the exocyclic oxygen to interact with the antibonding $\sigma^*$ endocyclic C–O orbital. The $\psi$ torsion angle is mainly influenced by steric effects.

![Figure 3.1](image-url) The torsion angles $\omega$, $\phi$ and $\psi$ for the disaccharide $\beta$-D-Glc$p$-(1→3)-$\beta$-D-Gal$p$.

The $\omega$ torsion angle is the most flexible of the three torsion angles, defined by O5–C5–C6–O6 and influenced by steric and electronic effects. The main electronic effect is the gauche effect, which describes the preference of vicinal electronegative groups to adopt a gauche relationship, rather than the usual antiperiplanar relation. The effect is caused by the ability for hyperconjugation from the C-H $\sigma$-bonding orbital into the antibonding $\sigma^*$ C-O orbital, thus stabilizing the conformation. The main steric factor is the repulsion between the hydroxyl groups in C4 and C6.
The three most common conformations of the C5-C6 fragment are staggered and termed gg (gauche-gauche), gt (gauche-trans) and tg (trans-gauche), referring to the spatial arrangement of the electronegative groups to each other and the relative arrangement of O6 to C4 of the ring. The CH₂ group is prochiral, and the two hydrogens are denoted pro-R and pro-S. The most preferred conformations are gt and gg.

![Figure 3.2 The three staggered rotamers describing the ω-torsion angle.](image)

In gt and gg the torsion angle for ω is +60° and −60°, respectively. In the tg rotamer, the torsion angle is 180°, leading to the unfavorable trans-relation between O5 and O6. The H₆pro-R and H₆pro-S are assigned in paper I, based on these properties.

To gain an understanding of the structural properties of a carbohydrate system in a biologically relevant environment, the carbohydrates are preferably studied in water. However, complete three-dimensional representations of carbohydrate moieties are not accessible using NMR spectroscopy alone. Conformations of glycosidic linkages are difficult to determine as the properties extracted from a NMR experiment is a time average for the different conformations the linkages populate. For complete insight into the three-dimensional structure of oligosaccharides the dynamic properties around the glycosidic linkages need to be considered.

3.2 Molecular dynamics simulations in conformational analysis

Molecular dynamics (MD) simulations is a computational technique used to predict and describe the motions of a molecule as a function of time. Comparisons of the results to data extracted from experimental techniques tests the validity of the simulation. In combination with experimental data, MD simulation is a powerful tool that can give insights into a number of properties of a model system.
In a MD simulation the molecule is first solvated, the potential energy of the system minimized and the temperature raised by adding kinetic energy to set the molecule in motion. A force field calculates the forces arising from interactions of bonded and non-bonded atoms to simulate molecular motions. The atoms are thus assigned velocities and allowed to adjust their locations in small steps under the influence of the force field, populating low energy conformations. The procedure is repeated for a sufficient amount of time, usually in the range of ns to μs. The result of a simulation is often represented as a trajectory of for example a distance plotted against time or a scatter plot of glycosidic torsion angles, see Figure 8.6.

A number of different force fields parameterized to describe carbohydrate interactions are currently in use. In this thesis three different force fields were employed; GLYCAM06, PARM22/SU01 and CHARMM2011.

3.3 NMR measurements

3.3.1 Nuclear Overhauser Effect

The nuclear Overhauser effect (NOE) is the change in spin population and consequently signal intensity through cross-relaxation when the spin population of a nuclei close in space is perturbed. Perturbation can be brought about either through saturation or inversion, using a steady-state or transient experiment, respectively.

The size and sign of the NOE is dependent of the nuclei observed, the rotational correlation time of the molecule, τc, and the spectrometer frequency used. The correlation time in turn is dependent of the size of the molecule and the temperature and viscosity of the solution. Figure 3.3 shows the dependency of the magnitude and sign of the transient NOE (ηtransient) on ωτc for a nucleus, where ω is the Larmor frequency for the nucleus and τc the time it takes for the molecule to rotate one radian. At $\omega \tau_c = \sqrt{5}/4$ the ηtransient cross a zero-point and goes from a positive to negative sign, see Figure 3.3. By changing the temperature and spectrometer frequency the ηtransient move from the intermediate region, either into extreme narrowing region with faster tumbling and smaller $\omega \tau_c$, or into the region of slow tumbling and larger magnitude NOE.
NOE is a sensitive probe for short distances and the effects are inversely proportional to the sixth power of the distance between the nuclei, a property that makes it useful in conformational analysis.

In paper VI effective proton-proton distances has been determined using the isolated spin-pair approximation (ISPA)\(^42\), which states that the distance between two nuclei, \(r_{ij}\), can be calculated if a reference distance, \(r_{ref}\), and the cross-relaxation rate for the reference distance is known, as well as the cross-relaxation rate from the interaction between \(i\) and \(j\).

\[
 r_{ij} = r_{ref} \left( \sigma_{ref} / \sigma_{ij} \right)^{1/6}
\]  

(3.1)

To retrieve the cross-relaxation rates, selective proton excitation using an array of NOE experiments with different mixing times can be performed. Proton-proton cross-relaxation rates, as defined by Dixon et al., are calculated by comparing the intensity of the perturbed peak, \(I_j\), with that of the detected peak, \(I_i\), multiplied with the mixing time, \(\tau_{mix}\), during which the relaxation occurs.\(^43\)

\[
 \sigma_{ij}^{\tau_{mix}} = -I_j^{\tau_{mix}} / \left( \tau_{mix} I_i^{\tau_{mix}} \right)
\]  

(3.2)

By plotting the cross-relaxation rate of the different protons affected versus the mixing time, slopes where the intercept of the curves corresponds to the cross-relaxation rate for the transition are created, see Figure 8.4.
3.3.2 Trans-glycosidic coupling constants

The trans-glycosidic three-bond scalar coupling constants, $^{3}J_{\text{COCH}}$ and $^{3}J_{\text{COCC}}$, can be used to help determine the torsion angles $\phi$ and $\psi$. The coupling constants can be related to their conformational relationship as shown by Karplus$^{20}$ with the general formula:

$$^{3}J_{\phi} = A \cos^2 \theta + B \cos \theta + C \quad (3.3)$$

Where A, B and C are constants and $\theta$ the torsion angle. The equation holds for both hetero- and homonuclear coupling constants and has been re-parameterized several times. In paper VI the most accurate Karplus equations to date have been used for calculation, and a plot of the function can be seen in Figure 3.4.$^{44}$

![Figure 3.4 Karplus-type relationship for trans-glycosidic $^{1}H,^{13}C$ heteronuclear couplings for a $\beta$-D linkage as a function of the $\phi$ torsion angle, as parameterized by Säwén et al.$^{44}$](image)

The $^{3}J_{\text{COCH}}$ can be experimentally determined using, among others, the 2D $^{1}H,^{13}C$-J-HMBC NMR experiment which utilizes a low-pass $J$-filter to suppress $^{1}J_{\text{C,H}}$ couplings.$^{45}$ The coupling constants are extracted from the F$_1$ dimension anti-phase peaks, separated by $\kappa J$, where $\kappa$ is a scaling factor employed due to the low resolution of the indirect dimension.
4 Structural studies of the O-polysaccharides from diarrheal *Escherichia coli* O175 and O177 (Papers I and II)

4.1 *E. coli* O175 (Paper I)

The strain O175:H28 was first isolated from a case of human diarrhea and belongs to the enteroaggregative (EAEC) group of *E. coli*. EAEC strains are often the cause of outbreaks in persistent pediatric diarrhea in developing countries. The effect of EAEC infection is watery, occasionally bloody, diarrhea with low-grade fever. \(^{46}\) EAEC colonize the intestinal mucosa and secrete enterotoxins and cytotoxins, and studies have shown that EAEC induce significant mucosal damage, although mild. \(^{47}\)

_E. coli_ O175:H28 was reported to cross-react with antisera against *E. coli* O22\(^{48}\) and O83\(^{49}\), indicating structural similarities in their respective O-antigen polysaccharides. \(^{50}\) The repeating unit for both O-antigens is a linear pentasaccharide of pyranoses, containing a \((1\rightarrow6)-\alpha-D-Glc\betap-(1\rightarrow4)-\beta-D-Glc\alpha p\) structural element. Due to the cross-reactivity, it is probable that this carboxylic functionality is also present in the O-antigen of *E. coli* O175.

4.1.1 Result and discussion

Sugar analysis of the polysaccharide of O175:H28 using alditol acetates showed that the sample contained glucose, mannose and 2-amino-2-deoxy-galactose. Absolute configurations of the residues were determined by butanolysis showing that all residues in the sample had the D-configuration.

A \(^1\text{H}\) spectrum revealed five anomic signals, see Figure 4.1, which was confirmed by the multiplicity-edited \(^1\text{H},\ ^{13}\text{C}-\text{HSQC}\) spectrum, see Figure 4.2.
The five anomeric residues had $^{13}$C chemical shifts indicating hexopyranosyl rings, but only four sets of CH$_2$-groups, revealing that not all residues contained a primary hydroxyl group. In combination with a resonance in the carbonyl region of the $^{13}$C spectrum this indicated that one residue contained a carboxylate functionality, as expected from the cross-reactivity with the E. coli O22 and O83 antigens. A cross-peak at 2.06/23.3 ppm also showed the presence of an N-acetyl group. The carbohydrate residues were denoted A-E in order of decreasing $^1$H chemical shifts. The anomeric configuration of residue E was recognized as β from the large $^3$$J_{HH,2}$ of 8.3 Hz, while large $^1$$J_{C1,H1}$ of 172-176 Hz showed that the remaining residues A-D were α-linked.

**Figure 4.1** Selected region from the presaturated $^1$H spectrum of the O-antigen of E. coli O175.

**Figure 4.2** Part of the multiplicity-edited $^1$H, $^{13}$C-HSQC NMR spectrum of the O-antigen polysaccharide of E. coli O175 with anomeric and substitution positions annotated. CH$_2$-signals are shown in grey.
Using the $^{1}H,^{1}H$-TOCSY experiment with different mixing times, in combination with the $^{1}H,^{13}C$-HSQC, the chemical shifts for the five different residues were revealed. By analyzing the TOCSY spectra it was also possible to get clues as to which carbohydrate moieties the sample contained. Residue A was determined to be glucose, and a large glycosylation shift in $^{13}C$ showed that it was substituted at C2. Residue B was assigned to be a 4-substituted glucuronic acid, due to the high $^{1}H$ chemical shift of H5. The spin systems of residues C and D were both assigned as mannose based on their high H2 chemical shift (4.00 and 4.23 ppm) and substituted at C2 and C3, respectively. Finally, residue E was determined to be a galactose moiety, and a low chemical shift of C2 (51.9 ppm) indicated that the residue was N-acetylated, thus residue E was assigned as 2-acetamido-2-deoxy-galactose substituted at C3.

The $^{1}H,^{13}C$-HMBC experiment was used in combination with the $^{1}H,^{1}H$-NOESY experiment to deduce the linkage pattern of the residues, showing that the repeating unit produced by *E. coli* O175:H28 is a linear pentasaccharide. The BS-CT-HMBC was used to give correlations between the N-acetyl carbonyl carbon and H2 of residue E as well as the methyl protons, see Figure 3 in paper I. The experiment also correlated the uronic carboxylic carbon to H5 of residue B. For all chemical shifts of O175, see paper I, Table 1. The repeating unit of O175:H28 is:

$$\rightarrow 2)\alpha-D-Glc(p-(1 \rightarrow 4)\alpha-D-Glc(pA-(1 \rightarrow 3)\alpha-D-Manp-(1 \rightarrow 2)\alpha-D-Manp-(1 \rightarrow 3)\beta-D-GalpNAc-(1 \rightarrow)$$

**A** | **B** | **D** | **C** | **E**

Comparing the O-antigen repeating unit from *E. coli* O175 to the repeating units of O22 and O83, all three polysaccharides are linear pentasaccharides containing hexopyranosyl residues, with an $\alpha-D-Glc(p-(1 \rightarrow 4)-D-Glc(pA$ linkage, see Figure 4.3.

![Figure 4.3](image-url) Schematic of the repeating units of *E. coli* O175 (top), O22 (middle) and O83 (bottom) in CFG notation.

A spin system of low intensity was identified, with chemical shifts similar to an unsubstituted $\alpha-D-Glc(p$ and thus denoted $A'$. We propose that the residue is located at the terminal end of the polysaccharide, which means that the 3-substituted 2-acetamido-2-deoxy-galactose is located at the reducing end of
the repeating unit and that the biological repeating unit is defined. Integration of the anomeric resonances of residue A and A’ showed that the repeating unit consists of about 20 residues on average.

H1 in residue C shows NOESY correlations to H1 and H5 in residue D, as well as correlations to H3 and H4 in residue E, see Figure 4.4. The close proximity of H4 in residue E to H1 of residue C cause steric interactions, which give rise to lower carbon chemical shifts of the carbons attached than would otherwise have been expected. The effect becomes evident when comparing the glycosylation shift of C1 of residue C (ΔδC 0.7 ppm) to that of any one of the other residues (ranging from ΔδC 5.9 ppm to 8.1 ppm). This is a known phenomenon referred to as the gamma-(γ) gauche effect. When carbons, separated by three bonds, are in gauche relation to each other the steric crowding from the protons attached causes an upfield shift of the chemical shifts of said carbons, compared to their trans arrangement.

\[ \text{Figure 4.4 Part of the } ^1\text{H},^1\text{H-NOESY NMR spectrum of the O-antigen polysaccharide from E. coli O175 of (a) the anomic region with trans-glycosidic correlations annotated and (b) correlations from the methyl group of the } N\text{-acetyl group of the } \beta-\text{D-GalpNAc residue.} \]

The NOESY spectrum showed cross-peaks from the N-acetyl methyl group of residue E to H5 as well as to H6\textsubscript{pro-S} and H6\textsubscript{pro-R} of residue C, see Figure 4.4b. When inspecting a molecular model of the disaccharide α-D-Manp-(1→3)-β-D-GalpNAc-(1→ (residue C–E) it is evident that H6\textsubscript{pro-S} in residue C is closer than H6\textsubscript{pro-R} to the N-acetyl methyl group on residue E on average. As envisioned in Figure 3.2, the gt and gg conformations are the most preferred. In the gt conformations both H6-protons are positioned with approximately the same distance from the N-acetyl methyl group, while in the gg conformation H6\textsubscript{pro-S} in residue C is closer to the methyl group, and thus gives a stronger NOESY correlation. See Figure 5b in paper I for a schematic picture of the D–C–E part of the pentasaccharide repeating unit, illustrated with the gg conformation for the ω torsion angle in residue C.
4.2 *E. coli* O177 (Paper II)

*E. coli* strain O177:H25 has been isolated in cattle, goats and humans and belong to the group of verocytotoxin-producing strains, which belongs to the EHEC virotype. The strain cross-reacts weakly with *E. coli* O18ab and O15.\textsuperscript{50} A previous study determined the functions of the glycosyl transferases used in the biosynthesis of *E. coli* O177, based on the action of similar *E. coli* glycosyl transferases.\textsuperscript{18}

4.2.1 Results

Chemical component analysis showed that the polysaccharide contained D-GlcN, L-FucN and L-Rha. Full elucidation of the polysaccharide structure was performed using 1D and 2D NMR experiments.

![Figure 4.5 The ¹H NMR spectrum of the O-antigen polysaccharide from *E. coli* O177.](image)

Four anomic signals with chemical shifts between 5.08 and 4.66 ppm could be identified in the ¹H NMR spectrum and the residues were denoted from A to D in order of falling ¹H chemical shifts, see Figure 4.5. Three signals of ¹H chemical shifts of 1.30–1.20 indicated 6-deoxy functionalities, and three signals of chemical shifts of 2.06–2.00 indicated N-acetylated amino sugars. The ¹³C chemical shifts of the anomic signals were indicative for hexopyranosyl residues.

¹H,¹H-TOCSY experiments with varying mixing times were used in combination with a ¹H,¹³C-HSQC experiment to assign the residues, and analysis showed that residue A is L-Rhap, B and C are L-FucpNAc and D is D-GlcpNAc. The anomic ¹J_{CH} for residue A–C are over 173 Hz while residue D has a coupling of 164 Hz, showing that the residues are α- and β-linked, respectively.
$^{13}$C glycosylation shifts showed that residue A is 2-substituted and residues B–D are 3-substituted. A $^1$H,$^{13}$C-HMBC experiment, see Figure 4.6, was used in combination with a $^1$H,$^1$H-NOESY experiment to deduce the substitution pattern.

**Figure 4.6** Selected region of the $^1$H,$^{13}$C-HMBC spectrum of the O-polysaccharide from *E. coli* 0177 with pertinent annotations.

The O-antigen repeating unit of O177:H25 is a linear tetrasaccharide with the following structure:

$$\rightarrow 2)-\alpha-L-Rhap(1\rightarrow 3)-\alpha-L-FucpNAc-(1\rightarrow 3)-\alpha-L-FucpNAc-(1\rightarrow 3)-\beta-D-GlpNAc-(1\rightarrow$$

The repeating unit for *E. coli* O177 shares the $\rightarrow 3)-\alpha-L-FucpNAc-(1\rightarrow 3)-\beta-D-GlpNAc$ structural motif with *E. coli* O15, which may explain the cross-reactivity.$^{52}$ As can be seen in the $^1$H spectrum, a signal of low intensity is present in the anomic region. Analysis of the spin system originating from this resonance, using $^1$H,$^{13}$C-HSQC, $^1$H,$^1$H-TOCSY as well as $^1$H,$^{13}$C-H2BC, showed chemical shifts similar to that of $\alpha-L$-Rhap-OME$^{53}$, and the spin system was determined as A'. The location of this residue at the terminal end of the repeating unit was further corroborated by the anomic $^1$H and $^{13}$C resonances width at half maximum, which were smaller than those for resonance A. The narrower line width of A' indicates that $T_2$ is longer than in A, as a result of the larger flexibility of the terminal residue compared to the residues in the chain. This phenomenon has previously been observed for O-polysaccharides, and was studied with $^{13}$C NMR relaxation experiments for the O-antigen of *E. coli* O91.$^{54}$
4.3 Conclusions

The O-antigen repeating units for *E. coli* strains O175:H28 and O177:H25 have been determined by chemical analysis and NMR spectroscopy.

*E. coli* O175:H28 is composed of a linear pentasaccharide with approximately 20 repeating units on average in the sample investigated. Structural similarities in the primary structure of the repeating units of the O-antigen from *E. coli* O175 and that of O22 and O83 explain the previously reported cross-reactivity. Several trans-glycosidic NOEs were observed, which could be used for conformational studies in combination with MD simulations and NMR measurements.55

*E. coli* O177:H25 was shown to be a linear tetrasaccharide of about 20 repeating units on average with a structural element similar to that of *E. coli* O15, which could explain the cross-reactivity previously shown.
5 Structural and genetic relationships of *Escherichia coli* O103 and *Salmonella enterica* O55 (Paper III)

5.1 Introduction

Diarrheal disease is a large and often fatal problem in the developing world.\(^{56}\) It is also a cause for nuisance in the industrialized world, increasingly so as traveling to areas where pathogenic strains are common increase.

The O-antigen of O103:K+:H8 belongs to the serogroup enterohemorrhagic *E. coli* (EHEC).\(^{57}\) EHEC was first recognized as a serogroup involved in human disease in 1982, and found to cause bloody diarrhea as well as hemolytic uremic syndrome (HUS).\(^{47}\) EHEC secrete Shiga toxin, which damage the renal (kidney) endothelial cells and in turn can lead to HUS. Some of the consequences of HUS are kidney failure and low red erythrocyte and trombocyte count. An extremely low dose of EHEC bacteria is sufficient for infection, and consequently is passed on easily by person-to-person transmission. The bovine intestinal tract is a common breeding ground for EHEC, thus initially the EHEC strains were linked to outbreaks associated with uncooked meat. However, other food items have also been associated to EHEC strains.

*E. coli* and *Salmonella enterica* (*S. enterica*) are closely related, but only a few cases have shown identical or close to identical O-antigen structures.\(^{58}\) The gene cluster of *E. coli* O103 was determined by Fratamico\(^ {57}\) et al. in 2005 and is in the present study showed to be almost identical to the gene cluster of *S. enterica* O55.

5.2 Structural assignment of the *E. coli* O103 O-antigen

The LPS of O103:K+:H8 was treated with base to yield the *O*-deacylated LPS, and subsequently purified by gel-permeation chromatography. Sugar analysis of the *E. coli* O103 O-antigen showed glucose, glucosamine, galactosamine as well as 3-amino-3-deoxy-fucose. The sugar analysis of *S. ente-
rical O55 gave the same result, but *E. coli* O103 contained more glucosamine and less glucose. The carbohydrate residues for *S. enterica* O55 were shown to have the D absolute configuration, and it was assumed to be the same for *E. coli* O103.

Figure 5.1 The $^1$H NMR spectrum of the anomeric region of the O-deacylated LPS from *E. coli* O103, annotated in falling chemical shift order.

Chemical shifts of the LPS-OH from *E. coli* O103 were taken from $^1$H and $^{13}$C NMR spectra and a $^{13}$C,$^1$H-HETCOR was employed to give the carbon/proton pairs. See Figure 5.1 for the anomeric region of the $^1$H NMR spectrum, and Figure 5.2 for selected regions from the HETCOR NMR spectrum. The anomeric $^{13}$C chemical shifts indicated five hexopyranosyl residues, and four cross-peaks at approximately 50 ppm were indicative of N-acylated carbons.

Figure 5.2 Selected regions from the $^{13}$C,$^1$H-HETCOR NMR spectrum showing a) the substituent Hb2 in E2, b) the nitrogen-bearing carbons, c) the anomeric residues of the O-deacylated LPS from *E. coli* O103.
Three methyl groups with chemical shifts typical for N-acetyl groups were present, as well as a cross-peak at $\delta_{H}/\delta_{C}$ 1.25/16.2 ppm indicating the presence of a 6-deoxy carbohydrate residue. The $^3J_{H_1,H_2}$ showed that residue D and E were $\beta$-linked, with couplings of 6.9 and 7.6 Hz, respectively. Analyzing the $^1J_{C_1,H_1}$ it was apparent that residue C was also $\beta$-linked, as the coupling was 164 Hz, while residue A and B had couplings which were larger than 171 Hz and thus $\alpha$-linked.

Using $^1$H,$^1$H-TOCSY with different mixing times along with the 2D $^1$H,$^13$C-H2BC, the spin system of the five residues were assigned. The fact that the magnetization did not travel beyond H4 in residue A, in combination with the rather high proton chemical shift of H4, 4.06 ppm, indicated that it was a galactose moiety. The low carbon chemical shift of C2 showed that it was N-acylated, and the high glycosylation shift showed that the residue was linked in C4. Residues B, C and D were all glucose moieties, glycosylated at C6, C3 and C2, respectively. Residues B and C were also N-acylated at C2. For residue E, the magnetization did not travel further than to H4 in the TOCSY experiments and in combination with the 6-deoxy functionality it was determined to be a fucosyl moiety, N-acylated in C3 and glycosidically linked at position C2. In addition to these findings a chain of three protonated carbons was identified in the TOCSY experiments, an apparent substituent of one of the N-acylated residues.

The $^1$H,$^13$C-HMBC was used, in combination with the $^1$H,$^1$H-NOESY, to confirm the glycosylation positions and how the residues were linked to each other. The unknown substituent was identified as a hydroxybutanoic acid (Hb), see Figure 1.1a, and due to the high chemical shift of the $\beta$-carbon it was determined to be $\beta$-hydroxylated. Thus, the substituent is a 3-hydroxybutanoic acid, linked to the carbohydrate residue through an amide linkage. A BS-CT-HMBC correlated the carboxyl carbon and C3 of residue E. The experiment also showed the correlations between the methyl groups and the remaining three carbonyl carbons, as well as the linkage position to the corresponding residue.

The repeating unit for the O-antigen of E. coli O103 is a non-branched pentasaccharide with the following structure:

\[
\rightarrow 2\beta-D\text{Glc}(1\rightarrow 2)\beta-D\text{Fucp}(1\rightarrow 6)\alpha-D\text{GalpNAc}(1\rightarrow 4)\alpha-D\text{GalpNAc}(1\rightarrow 3)\beta-D\text{Glc}p\text{NAc}(1\rightarrow 4)\text{E}\ D\ E\ B\ A\ C
\]

Full assignment of the repeating unit can be seen in Table 6.1
Table 6.1. $^1$H and $^{13}$C NMR chemical shifts (ppm) of the resonances from the O-antigen polysaccharide of E. coli O103 and inter-residue correlations from $^1$H,$^1$H-NOESY and $^1$H,$^{13}$C-HMBC spectra. NMR spectra were recorded at 340 K.

<table>
<thead>
<tr>
<th>Sugar residue</th>
<th>$^1$H/$^{13}$C</th>
<th>Correlation to atom (from anomic atom)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-α-D-GalpNAc(1→</td>
<td>A</td>
<td>5.49 [3.7]</td>
</tr>
<tr>
<td>(0.21)</td>
<td>(0.08)</td>
<td>(-0.04)</td>
</tr>
<tr>
<td>97.99 (176)</td>
<td>50.68</td>
<td>67.96</td>
</tr>
<tr>
<td>(6.04)</td>
<td>(-0.48)</td>
<td>(-0.44)</td>
</tr>
<tr>
<td>6-α-D-GlcNAc(1→</td>
<td>B</td>
<td>4.95</td>
</tr>
<tr>
<td>(0.26)</td>
<td>(0.06)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>99.19 (171)</td>
<td>54.93</td>
<td>71.26</td>
</tr>
<tr>
<td>(7.42)</td>
<td>(-0.07)</td>
<td>(-0.48)</td>
</tr>
<tr>
<td>3-β-D-GlcNAc(1→</td>
<td>C</td>
<td>4.82</td>
</tr>
<tr>
<td>(0.10)</td>
<td>(0.11)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>102.22 (164)</td>
<td>55.47</td>
<td>79.35</td>
</tr>
<tr>
<td>(6.37)</td>
<td>(-2.39)</td>
<td>(4.54)</td>
</tr>
<tr>
<td>2-β-D-Glc(1→</td>
<td>D</td>
<td>4.71 [6.9]</td>
</tr>
<tr>
<td>(0.07)</td>
<td>(0.21)</td>
<td>(-0.01)</td>
</tr>
<tr>
<td>101.06 (163)</td>
<td>82.32</td>
<td>77.18</td>
</tr>
<tr>
<td>(4.22)</td>
<td>(7.12)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>2-β-D-Fucp3N-(R)-3Hb6-(1→</td>
<td>E</td>
<td>4.55 [7.6]</td>
</tr>
<tr>
<td>(0.09)</td>
<td>(0.43)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>103.47 (162)</td>
<td>74.18</td>
<td>55.62</td>
</tr>
<tr>
<td>(5.82)</td>
<td>(3.72)</td>
<td>(-0.12)</td>
</tr>
</tbody>
</table>

$J_{H1,H2}$ values are given in hertz in square brackets and $J_{H1,C1}$ values in braces. Chemical shift differences as compared to corresponding monosaccharides are given in parentheses. * Chemical shift differences as compared to β-D-FucpNAc.
5.3 *S. enterica* O55 O-antigen and gene manipulation

The assignment of the repeating unit for the O-antigen of *S. enterica* O55 was made in a similar way as for *E. coli* O103, and had the following structure:

\[\rightarrow 2)\beta-D-Glc\beta-(1\rightarrow)\beta-D-Fucp3NAc-(1\rightarrow)\alpha-D-Glc\beta-(1\rightarrow)\alpha-D-GalpNAc-(1\rightarrow)\beta-D-Glc\beta-(1\rightarrow)\]

The similarities between the two structures are remarkable. The only differences are in the N-acyl group on the fucose moiety, and the residue the fucose links to, which is N-acetylated in *E. coli* O103 and not in *S. enterica* O55. The biosynthesis of the O-antigen is made by a cluster of genes, and for *E. coli*, *S. enterica* and also *Citrobacter freundii*, these genes are located in the same region on the chromosome. As the repeating unit of both O-antigens are so much alike, the gene cluster of *S. enterica* O55 was sequenced and compared to that of *E. coli* O103 sequenced by Fratamico et al. in 2005. It was discovered that the two gene sequences were very similar but for one gene believed to be an acyl transferase gene. A mutant was then created of *S. enterica* O55, where the gene for the acyl transferase was deleted and an insertion of the gene believed to be responsible for the acyl attachment in *E. coli* O103 was made. The strain expressed by the mutant *S. enterica* O55 bacterium, denoted H1988, was grown and analyzed, and the structure was:

\[\rightarrow 2)\beta-D-Glc\beta-(1\rightarrow)\beta-D-Fucp3NHb-(1\rightarrow)\alpha-D-Glc\beta-(1\rightarrow)\alpha-D-GalpNAc-(1\rightarrow)\beta-D-Glc\beta-(1\rightarrow)\]

The mutant H1988 O-antigen was substituted with 3-hydroxybutanoic acid in the Fuc3N-residue instead of the acetyl group present in the wild-type parental strain.

5.4 Conclusions

Structure determination of the *E. coli* O103 O-antigen has been performed using NMR spectroscopy and chemical analysis, and comparison to that of *S. enterica* O55 showed high similarity of the O-polysaccharide repeats.

Gene manipulation of the gene cluster for *S. enterica* O55 resulted in a O-antigen repeating unit close to identical to that of *E.coli* O103. This result supports the hypothesis that the gene, which differs in the two bacteria, is in fact the acyl transferase. It was suggested by Ochman and Wilson that *Salmonella* and *E. coli* diverged between 120 and 160 million years ago. This dates back to approximately the time when mammals first appeared, which is reasonable since *E. coli* ferments milk sugar lactose in the colon of mammals and *Salmonellae* neither ferment lactose nor live in the colon. The
results from this study show high gene similarity of the O-antigen biosynthesis of *E. coli* O103 and *S. enterica* O55. O-antigen modification may play a role in enhancing the survival and pathogenicity of a strain, and we suggest that these gene clusters share a common ancestor. After divergence of the bacteria the gene for acyltransferase in *E. coli* O103 was developed separately.
6 Structure determination of the O-antigen polysaccharide from *Escherichia coli* TD2158 digested by tailed bacteriophage HK620 (Paper IV)

6.1 Introduction

Tailed bacteriophages are viruses that infect bacteria, recognizing and cleaving the carbohydrate repeating units attached to the cell membrane at specific linkages. The tailed bacteriophages have an icosahedral head containing viral doubled stranded DNA and a protein tail that binds to and penetrates the LPS layer of the host bacterium. The phage inserts the DNA through the double cell membrane, then reproduces and ultimately kills the host cell. Sf6, P22 and HK620, which infect *Shigella flexneri*, *Salmonella enterica* and *Escherichia coli*, respectively, are all examples of tailed bacteriophages shown to be closely related.

The receptor-binding moiety of the tailed bacteriophage HK620 has been crystallized together with a hexasaccharide, produced by digestion of *E. coli* TD2158 in the receptor pocket showing where on the O-polysaccharide the bacteriophage cleaves. *E. coli* TD2158 was determined to be of the serological subgroup O18A1, however, the chemical shifts of the repeating unit of O18A1 have never been reported. To further investigate the linkage cleaved by the bacteriophage a full structural assignment of both the O-antigen polysaccharide as well as a thorough investigation of the oligosaccharide produced by the bacteriophage was performed by NMR spectroscopy and chemical analysis.

6.2 Structure elucidation of the O-polysaccharide

The LPS of *E. coli* TD2158 was first isolated from a sewage sample from Hong Kong and delipidated to produce the O-antigen polysaccharide. Sugar analysis and absolute configuration determination confirmed that the O-
polysaccharide contained the same sugar residues as the serogroups of \textit{E. coli} O18; L-rhamnose, D-glucose, D-galactose as well as D-glucosamine.

The $^1$H NMR spectrum of the O-polysaccharide of \textit{E. coli} TD2158 showed a complex anomic region, indicating heterogeneity in the sample, see Figure 6.1. With the help of a $^1$H,$^1$C-HSQC in combination with a $^1$C,$^1$H-HETCOR NMR spectrum, it was possible to make out twelve anomic signals of the same intensity from pyranose residues.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6_1.png}
\caption{Selected region of the multiplicity-edited $^1$H,$^1$C-HSQC NMR spectrum of the O-antigen polysaccharide from \textit{E. coli} TD2158. The sugar residues are denoted A–L and cross-peaks in the anomic region are annotated.}
\end{figure}

The anomic signals could be divided into five pairs with very similar $^1$H and $^1$C chemical shifts and two additional crosspeaks, indicating that the polysaccharide could have an alternating substitution pattern thus creating two different repeating units in a 1:1 ratio. The signals were denoted in falling $^1$H chemical shift order, from A to L. Residues A–H were recognized as $\alpha$-coupled due to the large $^1J_{C1,H1}$ of 171–173 Hz. Residues K and L both had a small $^1J_{C1,H1}$ of 161 Hz, and were thus denoted as $\beta$-coupled residues.

Five 2D $^1$H,$^1$H-TOCSY NMR spectra with mixing times ranging from 10 to 120 ms were used together with a $^1$H,$^1$C-H2BC and the $^1$H,$^1$C-HSQC NMR spectrum to assign the glycosyl moieties, while $^1$C glycosylation shifts revealed which positions were glycosylated. $^1$H,$^1$H-BASHD-TOCSY experiments, where a narrow spectral region can be selected in the $F_1$-dimension accompanied by homonuclear decoupling along the evolution dimension during $t_1$, were used to overcome problems with resonance overlap in the anomic region for residues C, D, E, F and G.\textsuperscript{65} See Figure 6.2 for the selective excitations of E, F and G. Note that residue E shows a TOCSY signal to ~4.3 ppm, which is H4 of residue A, while E and G clearly do not.
Figure 6.2 Selected regions of the BASHD-TOCSY NMR spectrum with selective excitations of E (red), F (green) and G (blue), as well as the $^1$H NMR spectrum (bottom). The chemical shift scale has been slashed and the intensities of the ring TOCSY signals are reduced relative to the anomeric excitations.

Residues A and B were identified as galactose residues linked in C3 and C4 while C and D were N-acetylated glucosamine residues linked in C3 and C6. Residues E–H were all glucose moieties, with residue E and H linked in C6 and F and G unsubstituted. Residues I and J were both rhamnoses glycosylated in position C2, while L was glucose and K an N-acetylated glucose, both residues only substituted at the anomeric position.

With the help of a $^1$H,$^1^3$C-HMBC and the $^1$H,$^1$H-NOESY NMR spectrum the glycosylation pattern, K–B, L–A, A–C, C–I, I–H, H–B, B–D, D–J, J–E and E–A, could be deduced for the repeating unit. See Table 1 in paper IV for chemical shifts and glycosylation shifts. The repeating unit for E. coli TD2158 is a hexasaccharide with a backbone of four carbohydrate residues, substituted at two positions. One of the branching residues is either β-D-GlcNAc or β-D-Glcp, substituted in approximately the same ratio. The repeating unit for the serogroup O18A1 is described by the hexasaccharide substituted by the β-D-GlcNAc.

\[
\begin{align*}
\text{F,G} & \quad \alpha\text{-d-Glc}p-(1\rightarrow 6) \\
\text{L,J} \quad \text{E,H} \quad \text{A,B} \quad \alpha\text{-d-Glc}p-(1\rightarrow 6) \quad \alpha\text{-d-Glcp}-(1\rightarrow 4) \quad \alpha\text{-d-Galp}-(1\rightarrow 3) \quad \alpha\text{-d-Glc}p\text{NAc}-(1\rightarrow) \\
& \beta\text{-d-Glcp/β-d-GlcpNAc}-(1\rightarrow 3) \\
& \text{L K}
\end{align*}
\]
6.3 Oligosaccharide produced by tailed bacteriophage

Tailspike endoglycosidase interaction with the O-polysaccharide from *E. coli* TD2158 produced an α/β-mixture of two similar hexasaccharides as a major and minor product. The major hexasaccharide was produced to approximately 90 percent and contained two D-GlcNAc residues; the branching β-D-GlcNAc and the backbone D-GlcNAc as an α/β-mixture, see Figure 6.3.

![Figure 6.3](image_url)

*Figure 6.3* Selected regions of the $^1$H,$^1^3$C-HSQC NMR spectrum of a hexasaccharide produced from the *E. coli* TD2158 O-antigen polysaccharide by the cleavage action of the tailspike endoglycosidase of *E. coli* phage HK620. a) The anomeric region, cross-peaks from the reducing end residue annotated to anomeric configuration; b) the region for D-GlcNAc C2 resonances with corresponding annotations.

The minor hexasaccharide contained a branching β-D-Glc and the backbone D-GlcNAc.

Oligosaccharides of 1–4 repeating units were analyzed by MALDI-TOF mass spectrometry, showing that homogenous β-D-GlcNAc-substituted oligosaccharides decreased with increasing oligosaccharide length, and no β-D-GlcNAc-substituted tetrasaccharides were observed.

The large amount of β-D-GlcNAc-substituted hexasaccharide can be explained by the binding and cleavage site on the tailspike protein HK620, where crystal structure analysis showed that the β-D-GlcNAc-(1→3)-linked oligosaccharides establishes a large amount of hydrogen bonding that are not possible to the same extent if the β-D-Glc substitutes the backbone. The catalytic amino acids Asp339 and Glu372 are located at the C-terminal end of the binding pocket, cleaving the polysaccharide between α-D-GlcNAc-(1→2)-α-L-Rhap-linkage. The result of the poor affinity of β-D-Glc-
substituted oligosaccharides is a reduced cleavage rate at the reducing end of these repeating units and a subsequent accumulation of β-D-GlcNAc-substituted oligosaccharides.

Figure 6.4 E. coli TD2158 hexasaccharide binding site on the HK620 tailspike protein. View of the full binding site with active site residues marked as pink sticks and the β-D-GlcNAc-(1→3)-linked residue in black.

6.4 Conclusions

The repeating unit for the E. coli TD2158 O-polysaccharide was determined by NMR spectroscopy and chemical analysis. The carbohydrate composition was elucidated by 1H and 13C chemical shifts, in combination with glycosylation shifts and coupling patterns. Several 2D NMR experiments were used to assign the twelve carbohydrate residues, including a series of 1H,1H-BASHD-TOCSY experiments to overcome problems with overlap in the anomic region.

The E. coli TD2158 O-polysaccharide is composed of irregularly hexasaccharide repeating units with a β-D-GlcNAc or β-D-GlcNAc branching residue, alternatively, where the former of these two hexasaccharides is the same as for the O18A1 serogroup.

The study showed that the tailed bacteriophage HK620 digests the E. coli TD2158 polysaccharide, preferably so if the branching residue is β-D-GlcNAc, creating the above-mentioned hexasaccharide by cleaving the α-D-GlcNAc-(1→2)-α-L-Rhap-linkage.
7 Structural elucidation of the capsular polysaccharide repeating unit from *Leuconostoc mesenteroides* ssp. *cremoris* PIA2 (Paper V)

7.1 Introduction

Probiotics are living microorganisms associated with positive effects on human and animal health. They function by increasing the concentration of beneficial microorganisms in the intestines and minimizing the amounts of pathogenic bacteria, thus preventing bacterial overgrowth. Probiotic microorganisms are usually consumed in the form of yogurts and fermented milks and health effects include faster recovery from acute diarrhea in children, relief of irritable bowel syndrome and modulation of the immune response.

The *Leuconostoc* genera are gram-positive bacteria often used as natural starter cultures in, for example, milk fermentation and belong to the lactic acid bacteria (LAB). The production of LAB exopolysaccharides (EPS) affects the consistency of fermented dairy products, and their physical properties are related to the structure of the EPS. The traditional Finnish fermented milk viili has a thick consistency and is inoculated with a culture of, among others, *Leuconostoc* strains. The term exopolysaccharides refers to polysaccharides secreted outside the bacterial cell wall, and the polymers can be tightly attached to the cell wall in the form of capsules or secreted in the form of unattached material. Some strains of the *Leuconostoc* genera are known to produce both unattached as well as capsule EPS, while other produce only the unattached material. *Leuconostoc* strains producing only the capsule EPS has not been confirmed, and CPS-producing LAB has not been studied as extensively as those producing only unattached material.

The strain *Leuconostoc mesenteroides* ssp. *cremoris* PIA2 has been recognized as a potential probiotic strain with immunomodulatory effects. Transmission electron microscopy has showed that the strain PIA2 produces only capsular exopolysaccharides (CPS).
7.2 Structural assignment of the CPS from *Leuconostoc mesenteroides* subsp. *cremoris* PIA2

The CPS was isolated from *Leuconostoc mesenteroides* subsp. *cremoris* PIA2 grown in skim milk. Sugar analysis and absolute configuration determination showed that the CPS contained D-galactose residues. A multiplicity-edited $^1$H, $^{13}$C-HSQC NMR spectrum of the CPS repeating unit revealed five anomic residues, with four residues with $^{13}$C chemical shifts of 110–107 ppm and one residue with a shift of 103 ppm, indicating four hexofuranosyl and one hexopyranosyl ring, see Figure 7.1. The carbohydrate residues were denoted A–E in order of decreasing $^1$H chemical shifts.

![Figure 7.1](image.png)

*Figure 7.1* Multiplicity-edited $^1$H, $^{13}$C-HSQC NMR spectrum of the CPS from *Leuconostoc mesenteroides* ssp. *cremoris* PIA2 showing (a) the anomic region and (b) the region for ring and hydroxymethyl atoms. CH$_2$-signals are shown in grey.

The chemical shifts of the sugar residues in the *Leuconostoc mesenteroides* subsp. *cremoris* PIA2 CPS were assigned using a combination of 1D and 2D NMR techniques. The $^1$H and $^{13}$C NMR chemical shifts of the five sugars in the repeating unit of the CPS are given in Table 1 in paper V.

Using the 2D $^1$H,$^1$H-TOCSY experiment with different mixing times, in combination with the $^1$H,$^{13}$C-HSQC, the chemical shifts for the five different residues were revealed. The high anomic $^{13}$C chemical shift of the four furanosyl moieties (A–D) indicated β-substitution, further supported by the rather small half-height linewidth of the $^1$H anomic resonances. The large
$^3J_{\text{H1,H2}}$ and small $^1J_{\text{C1,H1}}$ of residue E showed that it is also β-linked. From the analysis of $^1$H,$^1$H-TOCSY NMR spectra and $^1$H and $^{13}$C chemical shifts, it was deduced that all residues have the galacto-configuration.

The substitution positions for the sugar residues were identified from their $^{13}$C NMR glycosylation shifts. The $^1$H,$^1$C-HMBC experiment was used in combination with a 2D $^1$H,$^1$H-NOESY experiment to deduce the linkage-pattern of the residues, see Figure 7.2.

Figure 7.2 Part of a $^1$H,$^1$H-NOESY NMR spectrum of the CPS from Leuconostoc mesenteroides ssp. cremoris PIA2. Trans-glycosidic correlations from anomeric protons are annotated.

The spectra showed that the Leuconostoc mesenteroides subsp. cremoris PIA2 CPS is a linear pentasaccharide, and the repeating unit is:

$$\rightarrow3)\beta-D\text{-Gal} \rightarrow 1\rightarrow6)\beta-D\text{-Gal} \rightarrow 1\rightarrow2)\beta-D\text{-Gal} \rightarrow 1\rightarrow6)\beta-D\text{-Gal} \rightarrow 1\rightarrow3)\beta-D\text{-Galp} \rightarrow 1\rightarrow$$

D B C A E

The six-substituted furanosidic residues have larger $^3J_{\text{H1,H2}}$ than do the two- and three-substituted furanoses, with couplings of 1.7 Hz for the former and approximately 1 Hz for the later, indicating differences in their conformational preferences.

7.3 Conclusions

The CPS of Leuconostoc mesenteroides subsp. cremoris PIA2 was determined by chemical analysis and NMR spectroscopy to be composed of an unbranched galactan polysaccharide. Only a few homopolysaccharides have been identified as produced by LAB. There are studies of LAB producing both EPS and CPS, but LAB producing solely CPS has not been shown previously to this study to the best of our knowledge. Exactly how the CPS and unattached EPS of LAB affect the immune system of the host is unknown and further studies are needed to unveil the characteristics of their interaction.
8 Conformation and dynamics of a trisaccharide studied by NMR spectroscopy and MD simulation (paper VI)

8.1 Introduction

The Le\(^a\)Le\(^x\) hexasaccharide is a tumor-associated carbohydrate antigen (TACA) over-expressed at the surface of human squamous lung carcinoma cells. Cell surface oligosaccharides in normal cell tissue can differ from that of tumor tissue, making TACAs interesting as targets for cancer vaccines.\(^{77}\) Using the Le\(^a\)Le\(^x\) hexasaccharide as an antigen would likely induce an autoimmune reaction since the terminal non-reducing Le\(^a\) trisaccharide is commonly expressed by healthy tissue. Monoclonal antibodies (mAb) raised against the hexasaccharide have shown to selectively recognize Le\(^a\)Le\(^x\) while only binding weakly to Le\(^a\).\(^{78}\) The indication that the mAb recognizes internal fragments of the Le\(^a\)Le\(^x\) hexasaccharide, see Figure 8.1, makes the different epitopes interesting probes for study. With the end goal to develop necessary targeted vaccines against squamous lung carcinoma, it is imperative to fully characterize the internal epitopes of the Le\(^a\)Le\(^x\) hexasaccharide.

\[ \beta-D-GlcNAc-(1\rightarrow3)-\beta-D-Galp-(1\rightarrow4)-\beta-D-GlcNAc-OMe, \text{marked in black.} \]

The conformational properties of one such epitope, the trisaccharide \(\beta-D-GlcNAc-(1\rightarrow3)-\beta-D-Galp-(1\rightarrow4)-\beta-D-GlcNAc-OMe\), (1), was in the present study investigated using an extensive array of NMR experiments as
well as MD simulation by three different force fields. See Figure 8.2 for a schematic picture of trisaccharide 1.

![Figure 8.2 Trisaccharide β-D-GlcNAc-(1→3)-β-D-Galp-(1→4)-β-D-GlcNAc-OMe (1) with relevant torsion angles and anomeric protons labeled as described.](image)

### 8.2 Conformational studies

$^1$H and $^13$C chemical shifts of the trisaccharide were assigned using a combination of 1D and 2D NMR experiments. $^1$H chemical shifts and coupling constants were refined by spin simulation of a 1D $^1$H NMR spectrum using the PERCH software, see Figure 8.3 below and Table 1 in paper VI.

![Figure 8.3 Selected region of $^1$H NMR spectra of (a) the simulated spectrum and (b) the experimental spectrum of trisaccharide 1 in D$_2$O at 280 K and 700 MHz with pertinent atoms labeled.](image)

Several spectrometer frequencies and experiment temperatures were considered for studying the interresidue NOE interactions, before deciding for 700 MHz and 280 K. Diffusion experiment showed that the translational diffusion, $D_t$, is $1.71\times10^{-10}$ m$^2$/s and that the correlation time, $\tau_c$, of the molecule is 540 ps at 280 K. For a spectrometer frequency of 700 MHz this corresponds to a $\omega\tau_c$ of 2.4, see Figure 3.3, resulting in a good signal-to-noise ratio for the subsequent NOE experiments.
8.2.1 NMR measurements

Proton-proton cross-relaxation rates, $\sigma_{\text{NOE}}$, for the $\beta$-GlcNAc-(1→3)-Gal linkage were obtained by selectively exciting $H1''$ and $H4'$, respectively, using a series of 1D $^1H,^1H$-NOESY experiments with mixing times ranging from 60 to 300 ms. The experiment was repeated five times with varying mixing times, see Figure 8.4 for a plot from one of the experiments.

![Figure 8.4](image1)

*Figure 8.4* Plots of the cross-relaxation rate versus mixing time for $H3'$ (red square), $H3''$ (green triangle) and $H4'$ (blue diamond). Selective excitation of $H1''$ was performed with the 1D $^1H,^1H$ NOESY experiment and six mixing times between 130 and 300 ms.

Effective proton-proton distances were calculated using the ISPA analysis\textsuperscript{42}; the data is collected together with cross-relaxation rates in Table 8.1.

The $H1'$ and $H1$ protons were completely overlapped, see Figure 8.3, at all temperatures and spectral frequencies. Cross-correlation rates for the $\beta$-(1→4)-linkage could therefore not be studied by the 1D $^1H,^1H$-NOESY experiment, and was instead examined using the 1D $^1H,^1H$-STEP-NOESY experiment.\textsuperscript{80} The experiment involves two selective excitations and two mixing periods, starting with a selective TOCSY edited preparation period (STEP), followed by another selective excitation and a NOESY mixing period.

![Figure 8.5](image2)

*Figure 8.5* 1D $^1H,^1H$-STEP-NOESY experiment with selective excitation of $H4'$ followed by 100 ms TOCSY isotropic mixing with subsequent selective excitation of $H1'$ and a NOESY mixing period of 600 ms performed at 280 K and 700 MHz.
Selective excitation of H4' and an isotropic mixing period of 100 ms allowed the transfer of magnetization to H1'. A subsequent selective excitation of H1' followed by a NOESY mixing period gave NOE correlations arising from H1' alone, see Figure 8.5. Four different delays, ranging from 300 to 600 ms, were used in the NOESY mixing. H1' showed two inter-glycosidic NOE correlations to H3' and H5', as well as two trans-glycosidic NOE correlations to H4 and H3. All four signals were overlapped, so to extract the cross-relaxation rate from H1' to H4 the three remaining cross-relaxation rates were calculated using the ISPA analysis and subtracted from the total cross-relaxation rate, as showed valid by Jonsson et al. The reference distance \( r_{H1,H5} \) was taken from the PARM22/SU01 MD simulation and the \( \sigma_{H1,H5} \) was calculated from a 1D \( ^1\text{H}, ^1\text{H} \)-NOESY experiment with excitation of H1 and H1'.

To further investigate the conformational properties of the trisaccharide, the NOEs resulting from the N-acetyl methyl group of the terminal residue (Me") was examined. As Me" is separated from the methyl group of the N-acetyl in the reducing end (Me) by only 5 Hz at 700 MHz, a \( ^1\text{H}, ^1\text{H} \)-CSSF-NOESY was employed. \cite{82,83} A chemical shift selective filter (CSSF) is utilized by setting the peak of interest exactly on-resonance and letting all other chemical shifts evolve for a maximum delay, \( t_{\text{max}} \). When the FIDs are co-added, the magnetization of the on-resonance peak is added together while the magnetization of all other chemical shifts cancel each other out. The application of a NOE mixing period gives NOEs arising from the on-resonance peak only. The experiment showed NOEs between Me" and H1", as well as H6\text{pro-S}, see Table 8.1.

2D J-HMBC experiments were performed to obtain the trans-glycosidic \( J_{C,H} \) coupling constants, see Table 8.2. Strong coupling between H3 and H4 hindered the extraction of \( J_{C1',H4} \).

### 8.2.2 MD simulation

A total of three different MD simulations were carried out for trisaccharide 1. Two simulations were performed at 280 K, for the duration of 400 ns using the PARM22/SU01\cite{40} and for 200 ns using the CHARMM2011\cite{41} force field. A third simulation was performed for 20 ns at 300 K using the GLY-CAM06\cite{39} force field.

The translational diffusion constant, \( D_t \), was calculated\cite{84} to 1.77×10\(-10\) m\(^2\)/s and the correlation time, \( \tau_c \), to approximately 470 ps using the average from the two MD simulations performed at 280 K.

The simulations show some differences in how flexible the molecule is at the two linkages and how much the anti-\( \psi \) states are populated. All three force fields, however, show similar conformational preferences for trisaccharide 1, with the major state where \( \phi \approx 40^\circ \), which corresponds to the exo-anomeric conformation, and where the \( \psi \) torsion angles are in syn-
conformations, which is when H1 and the $H_{\text{glyco}}$ are on the same side of the trisaccharide.

In the PARM22/SU01 force field the $\psi_3$ torsion angle interconvert between approximately $-50^\circ$ and $+50^\circ$, see Figure 8.6 below and Figure 5a in paper VI, and the timescale of this conversion was calculated to 625 ps using a number correlation function with correction for the low viscosity of the computed water.

The PARM22/SU01 also showed populations of the non-exo-$\phi_4$ conformation, which was not seen in the other two simulations. Further analysis of the conformations populated for $\omega$ vs. $\phi_4$ shows several different hydrogen bonds during the simulation; see Figure 6 in paper VI. One of the hydrogen bonds is that between HO2' and O6, see Figure 6 and 8c in paper VI, which occurs when $\phi_4 \approx 0^\circ$ and could be related to the transitions between the exo- and the non-exo-anomeric conformations. Trans-glycosidic hydrogen bonding between $\beta(1\rightarrow3)$ and $\beta(1\rightarrow4)$ linked D-GlcNAc and $\beta$-D-GlcpA residues was recently showed by Nestor et al using NMR experiments, further supported by MD simulations.\(^8\)

The trans-glycosidic $^3J_{\text{CH}}$ were calculated by taking the average from the trajectories (Table 8.1). The Karplus-type equations used for calculation are given in Paper VI, equations 1 and 2.

### 8.2.3 Comparing results from NMR experiments and MD simulations

The translational diffusion at 280 K in D$_2$O was determined to be $1.77 \times 10^{-10}$ m$^2$/s from MD simulations, in very good agreement with the experimentally determined value, calculated to $1.71 \times 10^{-10}$ m$^2$/s. The correlation time, $\tau_c$, was
calculated to 540 ps from the diffusion experiment, and to 470 ps in MD simulations. The overall reorientation of trisaccharide 1 was shown to be slightly faster than the $\psi_3$ interconversion process, which was calculated from MD simulations to have a correlation time of 625 ps.

The trans-glycosidic proton-proton distances simulated at 280 K are in good agreement, calculated within 0.3 Å of the experimentally determined effective distances, as can be seen in Table 8.1. For distances calculated from the GLYCAM06 simulation at 300 K, see Table 2 in paper VI.

Table 8.1 Cross-relaxation rates and effective $^1$H,$^1$H distances of 1 derived from selective excitations of the first resonance in the atom pair using 1D NOESY experiments and MD simulations at 280 K.

<table>
<thead>
<tr>
<th>Proton pair</th>
<th>NMR $\sigma \times 10^3 /s^{-1}$</th>
<th>Molecular dynamics $r_{ij} /\text{Å}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r_{ij} /\text{Å}$</td>
<td>PARM22/SU01</td>
</tr>
<tr>
<td>H1&quot; H3'</td>
<td>123.7</td>
<td>2.29</td>
</tr>
<tr>
<td>H1&quot; H4'</td>
<td>12.75$^a$</td>
<td>3.34</td>
</tr>
<tr>
<td>H1&quot; H3&quot;</td>
<td>73.63</td>
<td>2.50$^b$</td>
</tr>
<tr>
<td>H1' H5'</td>
<td>90.53</td>
<td>2.40$^c$</td>
</tr>
<tr>
<td>H1' H4</td>
<td>171.4</td>
<td>2.1</td>
</tr>
<tr>
<td>NAc-Me&quot; H1&quot;</td>
<td>0.458</td>
<td>4.85$^d$</td>
</tr>
<tr>
<td>NAc-Me&quot; H6$_{pro-S}$</td>
<td>0.113</td>
<td>6.13</td>
</tr>
</tbody>
</table>

$^a$ Averaged cross-relaxation rate across glycosidic linkage.
$^b$ Reference distance in the NOESY experiment.
$^c$ Reference distance in the STEP-NOESY experiment.
$^d$ Reference distance in the CSSF-NOESY experiment.

Only three trans-glycosidic $^3J_{\text{CH}}$ could be extracted from the NMR experiments, and those are all within approximately 0.5 Hz of the coupling constants calculated in the MD simulations at 280 K, see Table 8.2.

Table 8.2 Trans-glycosidic $^3J_{\text{CH}}$ coupling constants derived from NMR experiments and MD simulations at 280 K. Coupling constants are given in Hz.

<table>
<thead>
<tr>
<th>Atom pair</th>
<th>$^3J_{\text{H-MB}}$</th>
<th>PARM22/SU01</th>
<th>CHARMM2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1&quot; C3'</td>
<td>4.21</td>
<td>3.89</td>
<td>3.79</td>
</tr>
<tr>
<td>C1&quot; H3'</td>
<td>4.73</td>
<td>3.64</td>
<td>4.47</td>
</tr>
<tr>
<td>H1' C4</td>
<td>3.78</td>
<td>3.25</td>
<td>3.49</td>
</tr>
<tr>
<td>C1' H4</td>
<td>n.d.$^a$</td>
<td>5.79</td>
<td>5.58</td>
</tr>
</tbody>
</table>

$^a$ n.d. = not determined.
Due to the estimated uncertainty of the J-HMBC experiment and the Karplus-type relationships used, deviations within ±0.75 Hz are considered to be in agreement with experimental data. The $^3J_{C1',H4}$ calculated from MD simulations are similar and the largest of the four trans-glycosidic couplings, also comparable to the previously measured $^3J_{C1,H4}$ coupling from methyl β-cellobioside, which has a coupling of 5.3 Hz in D$_2$O.\(^\text{86}\)

When plotting the torsion angles $\psi$ versus $\phi$, see Figure 8.4, it is evident that the β-GlcNAc-(1→3)-Gal linkage is more flexible than the β-Gal-(1→4)-GlcNAc-OMe linkage, as has been showed with other internal epitopes of the Le$^a$Le$^x$ hexasaccharide.\(^\text{87}\)

### 8.3 Conclusions

The trisaccharide 1 is a model system for an internal epitope of the hexasaccharide Le$^a$Le$^x$, and was investigated by NMR experiments to measure effective distances as well as trans-glycosidic heteronuclear couplings. Molecular dynamics simulations were performed in three different force fields and the results were in very good agreement as compared to experimental results. The simulations show that the trisaccharide is highly flexible in solution, especially so at the β-(1→3)-linkage, and that there is hydrogen bonding between the residues in the β-(1→4)-linkage as well as within the reducing end residue.
9 Concluding remarks and future prospects

There are about 170 (non-capsular) *E. coli* strains and far from all O-antigenic polysaccharides have been determined. In this thesis the primary structures of the repeating units of four *E. coli* strains have been determined, bringing us one step closer to fully mapping and understanding the role of the *E. coli* O-antigen.

The O-antigen repeating unit for *E. coli* O175:H28 was determined to be a linear pentasaccharide of 20 repeating units on average with a structurally similar repeating unit to that of *E. coli* O22 and O83, explaining the previously reported cross-reactivity. Several trans-glycosidic NOEs could be observed which, in combination with the corresponding $^3J_{CH}$ across the glycosidic linkages, make the O-polysaccharide interesting for conformational analysis using quantitative NMR measurements in combination with MD simulations. Also, it would be of interest to synthesize the disaccharide, $\alpha$-D-Manp-(1→3)-$\beta$-D-GalpNAc-(1→, which gives rise to the largest amount of NOE-correlations, to further investigate conformational properties at this linkage.

The structure of the O-polysaccharide repeating unit from *E. coli* O177 was determined as a linear tetrasaccharide of about 20 repeating units on average, with structural similarities to that of *E. coli* O15 and O18.

The O-polysaccharide repeating unit of *E. coli* O103 was shown to be highly similar to the repeating unit of the *S. enterica* O55 O-polysaccharide, both producing linear pentasaccharide repeating units differing only in the substituent of two residues. The gene cluster for *S. enterica* O55 was sequenced and gene manipulation resulted in an O-antigen repeating unit close to identical to that of *E. coli* O103, showing that the gene differing in the two strains encodes an acyl transferase. The high-level similarity between the gene clusters indicates that the *E. coli* O103 and *S. enterica* O55 originate from the same ancestor gene cluster. With the continuing work of mapping the *E. coli* O-antigen, similarities between other bacterial strains will surely be discovered. Further studies of the gene cluster of these bacterial strains should be performed to answer the question of what type of structural motifs has developed after the divergence and possibly also why the specific modifications has taken place.

The *E. coli* TD2158 O-polysaccharide was determined to be composed of hexasaccharide repeating units with an *E. coli* O18A1-type tetrasaccharide backbone, substituted with an $\alpha$-GlcP as well as either a $\beta$-D-GlcPNAc or a
β-D-Glcp residue. The bacteriophage HK620 tailspike protein was shown to digest the *E. coli* TD2158 O-polysaccharide specifically at the α-D-GlcpNAc-(1→2)-α-L-Rhap-linkage, leaving the β-D-Glcp-linked chain virtually undigested. The gene cluster coding for the *E. coli* TD2158 O-antigen has not yet been sequenced, and this would be a good starting point for future studies. Also, it would be of interest to investigate the binding process of the oligosaccharide containing the α-D-GlcpNAc-(1→2)-α-L-Rhap-linkage further, which would require that the bacteriophage digestion is performed with a polysaccharide from a homogenous O18A1-type producing strain.

The CPS of the probiotic LAB strain *Leuconostoc mesenteroides* subsp. *cremoris* PIA2 was determined to be composed of an unusual unbranched galactan polysaccharide. LAB producing CPS without unattached EPS has not been shown before this study. Maintaining the balance of the intestinal microflora is crucial, and ingestion of probiotics is suggested to prevent colonization of pathogenic strains. In addition, the physical properties of dairy products can be substantially improved using EPS-producing bacteria. Understanding the structure/function relationship of EPS in the intestinal tract would allow for tailoring of the polysaccharides to obtain modified fermented dairy products with the desired characteristics. With the knowledge of the type and structure of the CPS of *Leuconostoc mesenteroides* subsp. *cremoris* PIA2, the physiological role of the CPS should be further studied.

The conformational properties of the trisaccharide β-D-GlcpNAc-(1→3)-β-D-Galp-(1→4)-β-D-GlcpNAc-OMe, one of the internal epitopes of the hexasaccharide TACA Le°Le*, was investigated. The analysis showed that the trisaccharide is especially flexible at the β-(1→3)-linkage in solution, and that a great deal of hydrogen bonding increase the flexibility also at the β-(1→4)-linkage. The carbohydrates covering the cell surface of diseased and normal cells can differ considerably. Identification of specific carbohydrate markers of diseased cells would open the door to earlier diagnosis, prevention via vaccines and treatment via drugs that specifically target the interactions of the carbohydrates with their binding partner.
### Appendix

**Table A 1** $^1$H and $^{13}$C NMR chemical shifts of $\alpha$-D-Glcp-(1→1)-D-Gro in D$_2$O at 70°C referenced to internal TSP ($\delta_H$ 0.00) and to external dioxane in D$_2$O ($\delta_C$ 67.40).

<table>
<thead>
<tr>
<th>Residue</th>
<th>1 (ppm)</th>
<th>2 (ppm)</th>
<th>3 (ppm)</th>
<th>4 (ppm)</th>
<th>5 (ppm)</th>
<th>6 (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-D-Glcp-(1→1)-D-Gro</td>
<td>4.92 (3.8)$^a$</td>
<td>3.54</td>
<td>3.72</td>
<td>3.40</td>
<td>3.67</td>
<td>3.75, 3.84</td>
</tr>
<tr>
<td></td>
<td>99.54</td>
<td>72.37</td>
<td>74.00</td>
<td>70.63</td>
<td>72.77</td>
<td>61.59</td>
</tr>
<tr>
<td>$\rightarrow$1)-D-Gro</td>
<td>3.51, 3.79</td>
<td>3.94</td>
<td>3.59, 3.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.87</td>
<td>71.51</td>
<td>63.49</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ $J_{H1,H2}$ in parenthesis.

---

**Table A 2** $^1$H and $^{13}$C NMR chemical shifts of D-arabino-Hexp in D$_2$O at 70 °C referenced to internal TSP($\delta_H$ 0.00) and to external dioxane in D$_2$O ($\delta_C$ 67.40).

<table>
<thead>
<tr>
<th>Residue</th>
<th>1 (ppm)</th>
<th>2 (ppm)</th>
<th>3 (ppm)</th>
<th>4 (ppm)</th>
<th>5 (ppm)</th>
<th>6 (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-D-Ara-Hexp</td>
<td>5.37 (3.6)$^a$</td>
<td>1.71$^b$, 2.13$^c$</td>
<td>3.95</td>
<td>3.37</td>
<td>3.81</td>
<td>3.77, 3.85</td>
</tr>
<tr>
<td></td>
<td>92.19</td>
<td>38.13</td>
<td>68.85</td>
<td>72.27</td>
<td>72.93</td>
<td>61.82</td>
</tr>
<tr>
<td>$\beta$-D-Ara-Hexp</td>
<td>4.92 (9.8)$^a$</td>
<td>1.51$^b$, 2.27$^c$</td>
<td>3.71</td>
<td>3.29</td>
<td>3.38</td>
<td>3.74, 3.90</td>
</tr>
<tr>
<td></td>
<td>94.27</td>
<td>40.42</td>
<td>71.39</td>
<td>71.90</td>
<td>76.84</td>
<td>61.99</td>
</tr>
</tbody>
</table>

$^a$ $J_{H1,H2}$ in parenthesis.

$^b$ axial, $^c$ equatorial. The exchange rate at pD 6 is for H2$_\text{ax}$ $\gg$ H2$_\text{eq}$ 89.
11 Acknowledgements

The work here described was carried out between August, 2007, and May, 2012, at the Department of Organic Chemistry, Stockholm University. There are so many people I would like to thank for getting here. Thank you:

Professor Göran Widmalm – för att du antog utmaningen (mig!). Ditt engagemang som handledare har varit ovärderligt.

My co-workers in all of the projects I have participated in.

Professor Jan-Erling Bäckwall för visat intresse av min forskning.


Jerk, Robert och Carolina för era åsikter och uppmuntrande ord vid rättningen av denna avhandling.

Christoffer för ditt snygga jobb med Figure 6.4 och Robert för Figure 8.6.

TA-personalen som håller avdelningen flytande – ett stort tack till Kristina, Carin, Martin, Louise, Ola, Sigrid, Christina, Joakim och Pia.

Torbjörn Astlind och Helena Kovacs för all hjälp vid spektrometrarna.

Everyone at the Department of Organic Chemistry throughout the years, for the fun parties as well as the chemistry. Ett särskilt tack till Ylva, Karin och Nanna som satte mig i arbete utanför universitetet (Yoga! Tjurruet! Lidingsölloppet!).


My diploma workers Maria, Ylva and Ramona for allowing me to practice my teaching.


Laura och Daniel – tack för att ni ständigt ifrågasätter och stöttar mig. Ni har följt mig genom hela mitt liv, och vetat när jag behöver tas ner på jorden (även solen har sina fläckar)!

Mamma och pappa – tack för er värme. Ni hade nog inte tänkt att jag skulle hålla på riktigt så länge när ni väl övertalat mig att läsa vidare efter gymnasiet! Jag hade inte kommit hela vägen hit utan er.

Slutligen vill jag tacka min älskade Stan-Erik – för att du finns där och påminner mig om vad som är viktigt i livet. Utan dig hade resan varit både längre och svårare.
12 References


52
(70) Hassan, a N. *J. Dairy Sci.* **2008**, *91*, 1282-98.


