Glycomic studies of parasitic nematodes

A thesis submitted for the Degree of Doctor of Philosophy of

Imperial College London

Submitted by

Grigorij Sutov

Department of Life Sciences

Imperial College London

South Kensington Campus

SW7 2AZ

September 2015
Abstract

*Haemonchus contortus* and *Dictyocaulus viviparus* are parasitic nematodes which infect small ruminants and cattle, respectively worldwide causing huge economic losses. The current major control method is the usage of anthelmintic drugs. However, increasingly this strategy is failing due to an increase in parasite drug resistance. New control methods are therefore urgently needed and an attractive potential choice would be vaccination. However, all the attempts to create recombinant vaccines to date have failed therefore a potentially novel approach would be the development of carbohydrate-based vaccines. One of the major bottlenecks in this approach is the lack of structurally characterized parasitic nematode glycans.

In this thesis mass-spectrometry-based structural glycomic characterization of protein linked N- and O-glycans from *Haemonchus contortus* and *Dictyocaulus viviparus* adult and L3 stage parasites has been performed. The adult *Haemonchus contortus* glycomic analysis revealed previously unseen Gal-Fuc structures present in both N- and O-glycans. None of these structures were detected in L3 stage N- and O-glycan analysis. In comparison phosphorylated Galβ1-3(Galβ1-6)-GalNAc Core O-glycans were observed. The O-glycan analysis of L3 stage *Dictyocaulus viviparus* revealed the presence of Galβ1-3(Galβ1-6)-GalNAc Core glycans which can be further modified with fucose and α/β-galactoses. In contrast to L3 stage, the adult O-glycan analysis reveals mainly Core-2 based structures with the highest molecular weight structure forming the Lewis*α* antenna which is also abundantly observed in adult N-glycans.
Declaration

I hereby declare that this thesis is my own work and effort and that it has not been submitted to anywhere else for any other award. Where other sources of information have been used, they have been acknowledged.

Grigorij Sutov

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution-Non Commercial-No Derivatives license. Researches are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or distribution, researches must make clear to others the license terms of this work.
Acknowledgements

I would like to thank people who helped me to complete the thesis.

First of all I am particularly grateful to my supervisors Dr. Stuart Haslam and Prof. Anne Dell. Their constant support, discussions, useful insights were the engine which kept my thesis driving forward. Many thanks are given to Prof. Howard Morris who not only established the Biopolymer Mass Spectrometry Lab at Imperial College but also gave valuable insights into my work. I would also like to thank Dr. Maria Panico for fruitful discussions and smooth operation of the lab.

I also have to express big thanks to the past and present members of the lab who were around me every day, helped in the lab, shared a cup of coffee, held stimulating conversations. Aristotelis, Simon, Gang, Chos, Paul, Valeria, Poh-Choo, Paola, Laura, Mathew, Kevin, Tiandi, Nan, Dinah, Rebecca-this would not be possible without them.

I wish to acknowledge the support provided by my collaborators, Dr. Gerald Coles, University of Bristol for providing biological material and Dr. Roberto Castangia, Kratos Analytical, Shimadzu, Manchester for helping with MSn analysis.

I also want to express my warm thanks to my family, namely Mum, Dad, Maja, Liza, Katia, Seriozha and his family for their moral support. In addition I want to say big thanks to Sarah, Alex, Andrius, Karl, Laurynas, Valentin who shared my highs and pulled me up if I was low.

Finally I want to express my gratitudes to BBSRC for providing funding of my research.
Abbreviations

AAD  Aminoacetonitrile derivatives
AC   Alternating current
AMBIC Ammonium bicarbonate
APC  Antigen-presenting cell
Asn  Asparagine
Cer  Ceramide
CD   Cluster of differentiation
CHAPS 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate
CLR  C-type lectin receptor
Da   Dalton
DABP 3, 4-diaminobenzophenone
DC   Dendritic cell
DC   Direct current
DC-SIGN Dendritic cell-specific Intercellular adhesion molecule-3-grabbing non-integrin
DEAE 2-(Diethylamino)ethanol
DHB  2,5-dihydroxybenzoic acid
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DV</td>
<td><em>Dictyocaulus viviparus</em></td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EI</td>
<td>Electron ionization</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>ES</td>
<td>Excretory-secretory</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionisation</td>
</tr>
<tr>
<td>FAB-</td>
<td>Fast atom bombardment</td>
</tr>
<tr>
<td>Fuc</td>
<td>Fucose</td>
</tr>
<tr>
<td>Gal</td>
<td>Galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>N-acetylgalactosamine</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Gas chromatography coupled to mass spectrometry</td>
</tr>
<tr>
<td>Glc</td>
<td>Glucose</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetylgulosamine</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein-coupled receptor</td>
</tr>
<tr>
<td>GPI</td>
<td>Glycophosphatidylinositol</td>
</tr>
<tr>
<td>HA</td>
<td>Haemagglutinin</td>
</tr>
<tr>
<td>HC</td>
<td><em>Haemonchus contortus</em></td>
</tr>
</tbody>
</table>
Hib  *Haemophilus influenza* b type

Hex  Hexose

HexNAc  N-acetyllhexosamine

HF  Hydrofluoric acid

HPLC  High pressure liquid chromatography

IAA  Iodoacetic acid

IgA  Immunoglobulin A

IL  Interleukine

IRIV  Immune-potentiating reconstituted influenza virosome

KOH  Potassium hydroxide

KBH₄  Potassium borohydride

KLH  Keyhole limpet haemocyanin

LDN  GalNAcβ1-4GlcNAc

LDNF  GalNAcβ1-4(Fucα1-3)GlcNAc

Lewis³  Fucα1-3(Galβ1-4)GlcNAc

Lewis⁹  Fucα1-3(Fucα1-2Galβ1-4)GlcNAc

LGC  Ligand-gated channel

LN or LacNAc  Galβ1-4GlcNAc

LPG  Lipophosphoglycan
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man</td>
<td>Mannose</td>
</tr>
<tr>
<td>MGL</td>
<td>Macrophage galactose type C-type lectin</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MR</td>
<td>Mannose receptor</td>
</tr>
<tr>
<td>MALDI-</td>
<td>Matrix assisted laser desorption ionization</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MS&lt;sup&gt;n&lt;/sup&gt;</td>
<td>Mass spectrometry to the power of ‘n’</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem mass spectrometry</td>
</tr>
<tr>
<td>&lt;i&gt;m/z&lt;/i&gt;</td>
<td>Mass-to-charge ratio</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>NH&lt;sub&gt;4&lt;/sub&gt;OH</td>
<td>Ammonium hydroxide</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>OST</td>
<td>Oligosaccharyltransferase</td>
</tr>
<tr>
<td>OVA</td>
<td>Ovalbumin</td>
</tr>
<tr>
<td>PC</td>
<td>Phosphorylcholine</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PEtN</td>
<td>Phosphoethanolamine</td>
</tr>
<tr>
<td>PGC</td>
<td>Porous graphitized carbon</td>
</tr>
<tr>
<td>PMMA</td>
<td>Partially methylated alditol acetates</td>
</tr>
<tr>
<td>QIT</td>
<td>Quadrupole ion trap</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Q-TOF</td>
<td>Quadrupole time-of-flight</td>
</tr>
<tr>
<td>RF</td>
<td>Radio frequency</td>
</tr>
<tr>
<td>SEA</td>
<td>Schistosome egg antigen</td>
</tr>
<tr>
<td>Schisto Core</td>
<td>Galβ1-3(Galβ1-6)GalNAc</td>
</tr>
<tr>
<td>TACA</td>
<td>Tumour-associated carbohydrate antigens</td>
</tr>
<tr>
<td>TCR</td>
<td>T-cell receptor</td>
</tr>
<tr>
<td>Th1/Th2</td>
<td>T-helper 1/T-helper 2</td>
</tr>
<tr>
<td>TIC</td>
<td>Total ion count</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>TOF</td>
<td>Time of flight</td>
</tr>
<tr>
<td>TRIS</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>TT</td>
<td>Tetanus toxoid</td>
</tr>
<tr>
<td>UDP</td>
<td>Uridine diphosphate</td>
</tr>
</tbody>
</table>
Contents

Abstract 2
Declaration 3
Acknowledgements 4
Abbreviations 5
Contents 10
List of Figures 17
List of Tables 28
Chapter 1 Introduction 31
  1.1 Nematodes and their control methods 32
    1.1.1 Introduction to nematodes 32
    1.1.2 Introduction to *Haemonchus contortus* and its life cycle 34
    1.1.3 Introduction to *Dictyocaulus viviparus* and its life cycle 36
    1.1.4 Anthelmintics as a major control of nematode infection 38
    1.1.5 Currently available nematode vaccines 40
    1.1.6 Unsuccessful attempts to create recombinant vaccines against nematode infections 40
    1.1.7 The development of carbohydrate-based vaccines 41
  1.2 Structural and functional Glycomics in nematode species 46
    1.2.1 Introduction into glycobiology 46
    1.2.2 N-glycan biosynthesis in nematodes 47
      1.2.2.1 Overview of N-linked glycosylation in nematodes 47
      1.2.2.2 The biosynthesis of the lipid-linked Glc3Man9GlcNAc2 N-glycan precursor and its transfer to the protein 48
      1.2.2.3 The conversion of Glc3Man9GlcNAc2 N-glycan into truncated Man2GlcNAc3 structure based on model nematode *C. elegans* studies 50
      1.2.2.4 Chitobiase core modifications of Man2GlcNAc3 N-linked glycan in *C. elegans* model nematode 52
      1.2.2.5 Antennae modifications of N-linked glycans in *C. elegans* model nematode 54
1.2.3 N-linked glycosylation in *H. contortus* parasite 56
1.2.4 N-linked glycosylation in *D. viviparus* parasite 57
1.2.5 N-linked glycosylation in other nematode species 58
1.2.6 O- glycan biosynthesis in nematodes 60
   1.2.6.1 Overview of mucin-type O-linked glycosylation 60
   1.2.6.2 Mucin-type O-linked glycosylation in nematodes 61
1.2.7 Overview of glycolipids in nematodes 63
1.2.8 Functions of parasite glycans in shaping the immunity of the host 64
   1.2.8.1 Overview on parasite glycan-host immunity relationship 64
   1.2.8.2 Helminth glycan-host CLR interactions on dendritic cells 66
   1.2.8.3 Downstream events in dendritic cells upon helminth glycan-host CLR interactions 67
   1.2.8.4 Helminth Glycan-host soluble galectin interactions in the immune system 69
   1.2.8.5 Host immune system modulation by PC-containing glycans 69
1.3 Mass spectrometry as a tool for structural glycomics 71
   1.3.1 Overview and historical background of mass spectrometry: from electrons to glycoproteins 71
   1.3.2 Tandem mass spectrometry 73
   1.3.3 Other trends in modern mass-spectrometry 74
   1.3.4 Fast Atom Bombardment mass spectrometry as an early technique for glycan analysis 74
   1.3.5 AB 4800 MALDI-TOF-TOF instrument as a replacement for the FAB-MS sector mass spectrometer used in glycomics studies 76
      1.3.5.1 Overview 76
      1.3.5.2 Matrix Assisted Laser Desorption Ionisation 76
      1.3.5.3 Time-of-flight (TOF) Analyser 78
      1.3.5.4 Linear TOF analyser 79
      1.3.5.5 Reflectron TOF and delayed pulsed extraction 79
   1.3.6 MALDI-QIT-TOF instrument 80
      1.3.6.1 Overview of the MALDI-QIT-TOF instrument 80
      1.3.6.2 Principles of the QIT operation 81
   1.3.7 GC-MS instrument 81
      1.3.7.1 Overview 81
1.3.7.2 Electron Ionisation
1.3.7.3 Quadrupolar Analyzer
1.3.8 Derivatization of glycans
1.3.9 Overview of the strategy for interpreting MS spectra
1.3.10 Overview of glycan fragmentation pathways

1.4 Aims of the thesis

Chapter 2 Materials and Methods

2.1 Materials
2.1.1 Biological samples
2.1.2 General chemicals and reagents
2.1.3 Enzymes
2.1.4 Equipment and consumables

2.2 Methods
2.2.1 Overview
2.2.2 Worm homogenization
2.2.3 Reduction and carboxymethylation
2.2.4 Trypsin digestion and subsequent (glyco)peptide purification
2.2.5 PNGase F digestion
2.2.6 Sep-Pak separation of PNGase-F released N-glycans from remaining (glyco)peptides
2.2.7 PNGase A digestion and subsequent N-glycan separation from remaining (glyco)peptides
2.2.8 Reductive elimination of O-glycans and their subsequent purification
2.2.9 HF treatment of PNGase F released N-glycans
2.2.10 Exoglycosidase(s) treatment of N- and O-glycans
2.2.11 Permethylation procedure with subsequent glycan purification
2.2.12 MS analysis
2.2.13 Sample preparation for GC-MS linkage analysis
2.2.14 Data reproducibility

Chapter 3 Structural characterization of Haemonchus contortus N- and O-glycomes

3.1 Introduction
3.2 Structural characterization of the N-glycome of adult H. contortus
3.2.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from adult *H. contortus* 104
3.2.2 GC-MS linkage analysis of oligosaccharides released by PNGase F 108
3.2.3 MALDI-MS/MS analysis of permethylated PNGase-F released N-glycans 109
3.2.4 MALDI-QIT-TOF MS³ analysis of permethylated PNGase F released N-glycan 110
3.2.5 MALDI-MS analysis of permethylated PNGase F released, α-mannosidase digested N-glycans 111
3.2.6 The β1-4 specific galactosidase treatment coupled with MALDI-MS analysis of permethylated PNGase F released, α-mannosidase digested N-glycans 112
3.2.7 The β1-4 specific galactosidase treatment coupled with MALDI-MS-MS analysis of permethylated PNGase F released, α-mannosidase digested N-glycans 114
3.2.8 MALDI-MS-MS analyses of permethylated PNGase F released m/z 1661, 2010 N-glycans 118
3.2.9 MALDI-QIT-TOF MSⁿ analyses of permethylated PNGase F released m/z 2213 N-glycan 120
3.2.10 The summary of N-glycan structures released by PNGase F from adult *H. contortus* 123
3.2.11 MALDI-MS analysis of permethylated PNGase A released N-glycans from adult *H. contortus* 125
3.2.12 GC-MS linkage analysis of oligosaccharides released by PNGase A 129
3.2.13 MALDI-MS/MS analysis of permethylated PNGase A released N-glycans 130
3.2.14 MALDI-QIT-TOF MSⁿ analyses of permethylated PNGase A released m/z 1693 N-glycans 131
3.2.15 MALDI-QIT-TOF MS³ analysis of m/z 433 fragment ion from permethylated PNGase A released m/z 1519 N-glycan 133
3.2.16 MALDI-MS and MALDI-MS/MS analyses of permethylated PNGase F and PNGase A released, broad specificity β-galactosidase and broad specificity α-galactosidase digested N-glycans 135
3.2.17 MALDI-MS and MALDI-MS/MS analyses of permethylated PNGase A released, reducing end ring-opened N-glycans digested with the range of galactosidase enzymes 135
3.2.18 The summary of N-glycan structures released by PNGase A from adult *H. contortus* 135
3.2.19 MALDI-TOF MS analysis of α-mannosidase digested, hydrofluoric acid treated N-glycans released by PNGase F enzyme 137
3.2.20 MALDI-TOF-TOF MS/MS analysis of α-mannosidase digested, hydrofluoric acid treated N-glycans released by PNGase F enzyme 140
3.3 Structural characterization of the O-glycome of adult *H. contortus* 141
3.3.1 MALDI-MS analysis of permethylated O-glycans released by reductive elimination from adult *H. contortus* 141
3.3.2 GC-MS linkage analysis of partially methylated alditol acetates of O-glycans released by reductive elimination from adult *H. contortus* 143
3.3.3 MALDI-TOF-TOF analyses of permethylated, reductively eliminated lower molecular weight O-glycans 144
3.3.4 MALDI-QIT-TOF MS" analyses of permethylated, reductively eliminated *m/z* 1606 O-glycan 147
3.3.5 MALDI-TOF-TOF MS/MS analyses of permethylated, reductively eliminated other higher molecular weight O-glycans 151
3.3.6 MALDI-TOF MS analysis of permethylated reductively eliminated, broad specificity α1-3, 4, 6 galactosidase digested O-glycans 153
3.3.7 MALDI-MS and MALDI-MS/MS analyses of permethylated reductively eliminated, broad specificity β-galactosidase and β1-4-galactosidase digested O-glycans 155
3.3.8 The summary of reductively eliminated O-glycan structures from adult *H. contortus* 155

3.4 Structural characterization of the N-glycome of L3 stage *H. contortus* 157
3.4.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from L3 stage *H. contortus* 157
3.4.2 GC-MS linkage analysis of L3 stage *H. contortus* N-glycans released by PNGase F (35% acetonitrile fraction) 159
3.4.3 MALDI-MS/MS analyses of permethylated PNGase F released N-glycans from L3 stage *H. contortus* 160
3.4.4 MALDI-MS analysis of permethylated PNGaseA released N-glycans from L3 stage *H. contortus* 163
3.4.5 GC-MS linkage analysis of L3 *H. contortus* parasite N-glycans released by PNGase A (35% acetonitrile fraction) 164
3.4.6 MALDI-MS/MS analyses of permethylated PNGase A released N-glycans from L3 stage *H. contortus* 165

3.5 Structural characterization of the O-glycome of L3 stage *H. contortus* 168
3.5.1 MALDI-MS analysis of reductively eliminated, permethylated O-glycans from infective L3 stage *H. contortus* parasite 168
3.5.2 GC-MS linkage analysis of partially methylated alditol acetates obtained from L3 *H. contortus* O-glycans 169
3.5.3 MALDI-MS/MS analyses of reductively eliminated, permethylated O-glycans from L3 *H. contortus* 170

3.6 Structural conclusions of *H. contortus* parasite glycomes 175
3.6.1 Structural conclusions of adult N-glycans 175
3.6.2 Structural conclusions of adult O-glycans 177
3.6.3 Structural conclusions of L3 stage N- and O-glycans 179

Chapter 4 Structural characterization of *Dictyocaulus viviparus* N- and O-glycomes 180
4.1 Introduction

4.2 Structural characterization of the N-glycome of adult *D. viviparus*

4.2.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from adult *D. viviparus*

4.2.2 MALDI-MS/MS analysis of permethylated PNGase F released N-glycans from adult *D. viviparus*

4.2.3 MALDI-MS analysis of permethylated PNGase A released N-glycans from adult *D. viviparus*

4.3 Structural characterization of the O-glycome of adult *D. viviparus*

4.3.1 MALDI-TOF MS analysis of reductively eliminated, permethylated O-glycans from adult *D. viviparus*

4.3.2 GC-MS linkage analysis of adult *D. viviparus* O-glycans released by β-reductive elimination

4.3.3 MALDI-TOF-TOF MS/MS analysis of reductively eliminated, permethylated O-glycans from adult *D. viviparus*

4.3.4 The summary of reductively eliminated O-glycan structures from adult *D. viviparus*

4.4 Structural characterization of the N-glycome of L3 stage *D. viviparus*

4.4.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from L3 stage *D. viviparus*

4.4.2 GC-MS linkage analysis of L3 *D. viviparus* N-glycans released by PNGase F

4.4.3 MALDI-TOF-TOF MS/MS analysis of permethylated PNGase F released N-glycans from L3 stage *D. viviparus*

4.4.4 MALDI-MS analysis of permethylated PNGase A released N-glycans from L3 stage *D. viviparus*

4.5 Structural characterization of the O-glycome of L3 stage *D. viviparus*

4.5.1 MALDI-MS analysis of permethylated reductively eliminated O-glycans from L3 stage *D. viviparus*

4.5.2 GC-MS linkage analysis of partially permethylated alditol acetates of O-glycans released by reductive elimination from L3 stage *D. viviparus*

4.5.3 MALDI-MS/MS analyses of permethylated reductively eliminated m/z 534 and 738 O-glycans

4.5.4 MALDI-MS\(^6\) analyses of permethylated reductively eliminated m/z 1116 O-glycan

4.5.5 MALDI-MS/MS analysis of permethylated reductively eliminated m/z 1146 O-glycan

4.5.6 MALDI-MS/MS analyses of other permethylated reductively eliminated O-glycans

4.5.7 MALDI-MS and MALDI-MS/MS analyses of permethylated reductively eliminated, broad specificity β-galactosidase digested O-glycans
4.5.8 MALDI-MS analyses of permethylated reductively eliminated O-glycans which were subjected to consecutive broad specificity β-galactosidase and broad specificity α-galactosidase treatment

4.5.9 MALDI-TOF-TOF MS/MS analyses of permethylated reductively eliminated O-glycans which were subject to consecutive broad specificity β-galactosidase and broad specificity α-galactosidase treatment

4.5.10 MALDI-MSn analysis of permethylated reductively eliminated m/z 708 O-glycan after consecutive broad specificity β-galactosidase and broad specificity α-galactosidase treatment

4.5.11 MALDI-MS analysis of permethylated reductively eliminated, X. manihotis β-galactosidase digested O-glycans

4.5.12 MALDI-MSn analyses of permethylated reductively eliminated, X. manihotis β-galactosidase digested O-glycan at m/z 942

4.5.13 MALDI-MSn analyses of permethylated reductively eliminated, X. manihotis β-galactosidase digested O-glycan at m/z 1146

4.5.14 MALDI-MS analysis of permethylated reductively eliminated, α-galactosidase digested O-glycans

4.5.15 MALDI-MSn analyses of permethylated reductively eliminated, α-galactosidase digested O-glycan at m/z 942

4.5.16 MALDI-MSn analyses of permethylated reductively eliminated, α-galactosidase digested m/z 1146 O-glycan

4.5.17 MALDI-MS analysis of permethylated reductively eliminated, broad specificity α-galactosidase and subsequently broad specificity β1-3, 4, 6 galactosidase digested O-glycans

4.5.18 MALDI-MSn analyses of permethylated reductively eliminated, α-galactosidase digested m/z 1350 O-glycan

4.5.19 The summary of reductively eliminated O-glycan structures from L3 stage D. vivipar"us

4.6 Structural conclusions of D. vivipar"us parasite glycome

4.6.1 Structural conclusions of adult N and O-glycans

4.6.2 Structural conclusions of L3 stage N- and O-glycans

Future perspectives

References
List of Figures

Figure 1 The phylogenetic tree of Nematode species 33

Figure 2 The lifecycle of Haemonchus contortus parasite 36

Figure 3 The life cycle of Dictyocaulus viviparus 38

Figure 4 The Schistosome mansoni glycoconjugate vaccine-candidate construct 43

Figure 5 The processing of glycoconjugate vaccine by B-cell and its presentation to T-carb cells 44

Figure 6 Most common monosacharides found in nematodes 46

Figure 7 The biosynthesis of lipid-linked GlcNAc2Man9Glc3 N-glycan precursor and its transfer to the protein in the ER 48

Figure 8 The biosynthesis of truncated Man2GlcNAc from Glc3Man9GlcNAc2 N-glycan in the ER and Golgi compartments 51

Figure 9 The chitobiose core modification pathways in C. elegans 52

Figure 10 The N-glycan antennae modifications found in C. elegans organisms 54

Figure 11 The N-glycan core and antennae modifications found in H. contortus organisms 56

Figure 12 The N-glycan structures found in D. viviparus parasite 57

Figure 13 N-glycan core and antennae modifications found in different nematode species 58

Figure 14 Different O-glycan Core structures found in mammals 60

Figure 15 The O-glycans found in nematode parasites 61
Figure 16 The glycosphingolipids reported in various nematodes

Figure 17 Different dendritic cell C-type lectin-pathogen glycan interactions

Figure 18 Two different signalling cascades happening in dendritic cells upon Lewis\(^x\) and high-mannose N-glycans engaging with the DC-SIGN receptor

Figure 19 Fast Atom Bombardment (FAB) Ionization Source

Figure 20 Matrix-assisted-Laser-Desorbtion-Ionisation (MALDI) Ionization Source

Figure 21 The schematic representation of the MALDI AXIMA QIT instrument

Figure 22 The schematic diagram of the Electron Ionisation

Figure 23 The schematic diagram of the quadrupole analyser

Figure 24 The mechanism of the permethylation reaction

Figure 25 Symbolic nomenclature and permethylated masses of monosacharides

Figure 26 The glycan fragmentation pathways

Figure 27 The proposed mechanisms for the formation of glycan ‘C-type’ and ‘Z-type’ ions

Figure 28: The proposed reverse Diels-Alder cross-ring fragmentation mechanism

Figure 29 Overview of glycomic methods used in the PhD thesis

Figure 30 MALDI MS spectra of PNGase F released adult *H. controtus* N-glycans

Figure 31 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released *m/z* 1549 from adult *H. contortus* N-glycan
Figure 32 MALDI-QIT-TOF MS³ spectrum of m/z 678 fragment ions from permethylated, PNGase F released m/z 1549 N-glycan

Figure 33 MALDI-TOF MS spectrum of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (35% acetonitrile fraction)

Figure 34 MALDI-TOF MS spectra of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (35% acetonitrile fraction) prior (top) and after (bottom) the β1-4 galactosidase digestion

Figure 35 MALDI-TOF-TOF MS/MS spectra of A) permethylated PNGase F released, α-mannosidase trimmed m/z 1315 N-glycan and B) permethylated PNGase F released, α-mannosidase and subsequently β1-4 specific galactosidase trimmed m/z 1315 N-glycan

Figure 36 MALDI-TOF-TOF MS/MS spectra of A) permethylated PNGase F released, α-mannosidase trimmed m/z 1141 N-glycan and B) permethylated PNGase F released, α-mannosidase and subsequently β1-4 specific galactosidase trimmed m/z 1141 N-glycan

Figure 37 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 1661 adult *H. contortus* N-glycan

Figure 38 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 2010 adult *H. contortus* N-glycan

Figure 39 MALDI-QIT-TOF MS² spectrum of permethylated PNGase F released m/z 2213 N-glycan

Figure 40 MALDI-QIT-TOF MS³ spectrum of m/z 905 fragment ions from permethylated PNGase F released m/z 2213 N-glycan
Figure 41 MALDI-QIT-TOF MS$^4$ spectrum of m/z 433 fragment ion generated from precursor fragment ion m/z 905

Figure 42 MALDI MS spectra of permethylated N-glycans released from the *H. contortus* adult stage whole worm extracts by PNGase A enzyme

Figure 43. MALDI-TOF-TOF MS/MS spectrum of 50% acetonitrile permethylated PNGase A released m/z 1693 adult *H. contortus* N-glycan

Figure 44 MALDI-QIT-TOF MS$^2$ spectrum of 35% acetonitrile permethylated PNGase A released m/z 1693 adult *H. contortus* N-glycan

Figure 45 MALDI-QIT-TOF MS$^3$ spectrum of m/z 852 fragment ion from permethylated (35% acetonitrile fraction) PNGase A released m/z 1693 N-glycan

Figure 46 MALDI-QIT-TOF MS$^2$ spectrum of permethylated PNGase A released m/z 1519 N-glycan

Figure 47 MALDI-QIT-TOF MS$^3$ spectrum of m/z 433 fragment ion from permethylated PNGase A released m/z 1519 N-glycan

Figure 48 MALDI-TOF MS m/z 500-5015 range spectra of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (50% acetonitrile fraction) prior (top) and after (bottom) the hydrofluoric acid treatment

Figure 49 MALDI-TOF MS m/z 1900-4000 range spectra of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (50% acetonitrile fraction) prior (top) and after (bottom) the hydrofluoric acid treatment

Figure 50 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 3551 N-glycan digested with α-mannosidase and treated with hydrofluoric acid
Figure 51 MALDI MS spectrum of permethylated O-glycans released from the *H. contortus* adult stage whole worm extracts by reductive elimination (35% acetonitrile fraction)

Figure 52 MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 534 adult *H. contortus* O-glycan

Figure 53 MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 575 adult *H. contortus* O-glycan

Figure 54 MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 738 adult *H. contortus* O-glycan

Figure 55 MALDI-QIT-TOF MS² spectrum of reductively eliminated m/z 1606 O-glycan

Figure 56 MALDI-QIT-TOF MS³ spectrum of m/z 486 fragment ions from permethylated reductively eliminated m/z 1606 O-glycan

Figure 57 MALDI-QIT-TOF MS³ spectrum of m/z 433 fragment ion from permethylated reductively eliminated m/z 1606 O-glycan

Figure 58 showing the MALDI-TOF-TOF MS/MS spectrum of reductively eliminated m/z 1402 O-glycan

Figure 59 MALDI-TOF-TOF MS/MS spectrum of reductively eliminated m/z 1361 O-glycan

Figure 60 MALDI-TOF MS spectra of permethylated reductively eliminated O-glycans before (top) and after (bottom) the broad specificity α-galactosidase digestion

Figure 61 MALDI-TOF MS spectra of permethylated 35% acetonitrile (top) and 50% acetonitrile (bottom) N-glycans released from L3 *H. contortus* by PNGase F enzyme
Figure 62 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 1345 L3 infective stage *H. contortus* N-glycan

Figure 63 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 1141 L3 infective stage *H. contortus* N-glycan

Figure 64 MALDI-TOF MS spectrum of permethylated 35% acetonitrile N-glycans released from L3 infective stage *H. contortus* parasite by PNGase A enzyme

Figure 65 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase A released m/z 1519 from L3 *H. contortus* N-glycan

Figure 66 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase A released m/z 1489 from L3 *H. contortus* N-glycan

Figure 67 MALDI-TOF MS spectrum of permethylated reductively eliminated O-glycans from L3 *H. contortus* (35% acetonitrile fraction)

Figure 68 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 534 O-glycan from L3 stage *H. contortus*

Figure 69 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 575 O-glycan from L3 stage *H. contortus*

Figure 70 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 738 O-glycan from L3 *H. contortus*

Figure 71 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 832 O-glycan from L3 *H. contortus*

Figure 72 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 926 O-glycan from L3 stage *H. contortus*
**Figure 73:** proposed structural conclusions of novel N-glycan structures discovered in the *Haemonchus contortus* adult parasite

**Figure 74** proposed structural conclusions of O-glycan structures discovered in the *Haemonchus contortus* adult parasite

**Figure 75** MALDI-TOF MS spectrum of permethylated 50% acetonitrile fraction N-glycans released from *D. viviparus* adult stage whole worm extracts by PNGase F

**Figure 76** MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 3215 N-glycan

**Figure 77** MALDI MS spectrum of permethylated, 35% acetonitrile fraction O-glycans released from the adult *D. viviparus* adult stage whole worm extracts by reductive elimination

**Figure 78** MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 534 adult *D. viviparus* O-glycan

**Figure 79** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 575 adult *D. viviparus* O-glycan

**Figure 80** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 738 adult *D. viviparus* O-glycan

**Figure 81** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 779 adult *D. viviparus* O-glycan

**Figure 82** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 1157 adult *D. viviparus* O-glycan

**Figure 83** MALDI-TOF MS spectrum of permethylated 50% acetonitrile N-glycans released from *D. viviparus* L3 stage whole worm extracts by PNGase F

**Figure 84** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 2694 L3 stage *D. viviparus* N-glycan
**Figure 85** MALDI-TOF MS spectrum of permethylated reductively eliminated 35% acetonitrile fraction O-glycans from *D. viviparus* L3 stage parasite

**Figure 86** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 534 L3 stage *D. viviparus* O-glycan

**Figure 87** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 738 L3 stage *D. viviparus* O-glycan

**Figure 88** MALDI-QIT-TOF MS\(^2\) spectrum of permethylated reductively eliminated m/z 1116 O-glycan from *D. viviparus* L3 stage

**Figure 89** MALDI-QIT-TOF MS\(^3\) spectrum of m/z 637 fragment ions from permethylated reductively eliminated m/z 1116 O-glycan

**Figure 90** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 1146 from L3 *D. viviparus* O-glycan

**Figure 91** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 1524 O-glycan from the *D. viviparus* infective L3 stage parasite

**Figure 92** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 1759 O-glycan from the L3 stage *D. viviparus*

**Figure 93** MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with bovine testes broad specificity β1-3, 4, 6 galactosidase

**Figure 94** MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated m/z 1146 L3 *D. viviparus* O-glycans before and after the bovine testes broad specificity β1-3, 4, 6 galactosidase digestion

**Figure 95** MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with bovine testes broad specificity β1-3, 4, 6 galactosidase and broad specificity α-galactosidase enzymes
Figure 96 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated, β1-3, 4, 6 galactosidase digested m/z 738 L3 *D. viviparus* O-glycans before and after the green coffee beans broad specificity α1-3, 4, 6 galactosidase digestion.

Figure 97 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated, β1-3, 4, 6 galactosidase digested m/z 912 L3 *D. viviparus* O-glycans before and after green coffee bean broad specificity α1-3, 4, 6 galactosidase digestion.

Figure 98 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated, β1-3, 4, 6 galactosidase digested m/z 942 L3 *D. viviparus* O-glycans before and after the green coffee bean broad specificity α1-3, 4, 6 galactosidase digestion.

Figure 99 MALDI-QIT-TOF MS² spectrum of permethylated reductively eliminated, bovine testes β1-3, 4, 6 galactosidase and green coffee bean α1-3, 4, 6 galactosidase digested m/z 708 O-glycan from the *D. viviparus* L3 stage.

Figure 100 MALDI-QIT-TOF MS³ spectrum of m/z 433 fragment ion from permethylated reductively eliminated, bovine testes β1-3,4,6 galactosidase and green coffee bean α1-3,4,6 galactosidase digested m/z 708 O-glycan from *D. viviparus* L3 stage.

Figure 101 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with *X. manihotis* β-galactosidase.

Figure 102 MALDI-QIT-TOF MS³ spectrum of permethylated reductively eliminated, *X. manihotis* β-galactosidase digested m/z 942 O-glycan from the *D. viviparus* L3 stage.

Figure 103 MALDI-QIT-TOF MS³ spectrum of m/z 463 fragment ions from permethylated reductively eliminated, *X. manihotis* β-galactosidase digested m/z 942 O-glycan.

Figure 104 MALDI-QIT-TOF MS³ spectra are showing m/z 463 Glcα1-3Glc and Manα1-6Man fragment ions from permethylated known standard compounds.

Figure 105 MALDI-QIT-TOF MS³ spectrum of permethylated reductively eliminated, *X. manihotis* β-galactosidase digested m/z 1146 O-glycan from the *D. viviparus* L3 stage.
Figure 106 MALDI-QIT-TOF MS\textsuperscript{3} spectrum of m/z 667 fragment ions from permethylated reductively eliminated, \textit{X. manihotis} \(\beta\)-galactosidase digested m/z 1146 O-glycan

Figure 107 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with green coffee bean broad specificity \(\alpha1-3, 4, 6\) galactosidase

Figure 108 MALDI-QIT-TOF MS\textsuperscript{2} spectrum of permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 942 O-glycan from the \textit{D. viviparus} L3 stage parasite

Figure 109 MALDI-QIT-TOF MS\textsuperscript{3} spectrum of m/z 706 fragment ion from permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 1942 O-glycan

Figure 110 MALDI-QIT-TOF MS\textsuperscript{4} spectrum of m/z 463 fragment ion from m/z 706 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 1146 O-glycan

Figure 111 MALDI-QIT-TOF MS\textsuperscript{2} spectrum of permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 1146 O-glycan from the \textit{D. viviparus} infective L3 stage parasite

Figure 112 MALDI-QIT-TOF MS\textsuperscript{3} spectrum of m/z 910 fragment ion from permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 1146 O-glycan

Figure 113 MALDI-QIT-TOF MS\textsuperscript{4} spectrum of m/z 667 fragment ions from m/z 910 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 1146 O-glycan

Figure 114 MALDI-QIT-TOF MS\textsuperscript{4} spectrum of m/z 463 fragment ion from m/z 910 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested m/z 1146 O-glycan
Figure 115 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with broad specificity α-galactosidase and bovine testes broad specificity β1-3, 4, 6 galactosidase enzymes.

Figure 116 MALDI-QIT-TOF MS² spectrum of permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1350 O-glycan from *D. viviparus* L3 stage parasite.

Figure 117 MALDI-QIT-TOF MS³ spectrum of m/z 1114 fragment ion from permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1350 O-glycan.

Figure 118 MALDI-QIT-TOF MS⁴ spectrum of m/z 463 fragment ion from m/z 1114 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1350 O-glycan.

Figure 119 proposed structural conclusions of O-glycan structures discovered in *D. viviparus* adult and L3 stages.

Figure 120 The overview of the strategies for enzymatic synthesis of adult *H. contortus* N- and O-glycans.
List of Tables

Table 1 Summary of anthelmintics 39

Table 2 The C. elegans genes corresponding to eukaryotic analogues shown in the Figure 7 49

Table 3 Assignment and quantification of PNGase F released adult H. contortus N-glycans 106

Table 4 Assignment and quantification of PNGase F released adult H. contortus N-glycans 107

Table 5 GC-MS analysis of partially methylated alditol acetates obtained from H. contortus N-glycans released by PNGase 108

Table 6 showing the proposed N-glycan structures released by PNGase F from adult H. contortus 124

Table 7 The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (Figure 42, top) of permethylated (35% acetonitrile fraction), PNGase A released N-glycans of adult H. contortus 127

Table 8 The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (Figure 42, bottom) of permethylated (50% acetonitrile fraction), PNGase A released N-glycans of adult H. contortus 128

Table 9 GC-MS analysis of partially methylated alditol acetates obtained from H. contortus N-glycans released by PNGase A. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak 129
Table 10 showing the proposed N-glycan structures released by PNGase A from adult *H. contortus*.

Table 11 The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 51) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of adult *H. contortus*.

Table 12 GC-MS analysis of partially methylated alditol acetates obtained from reductively eliminated *H. contortus* O-glycans.

Table 13 showing the proposed O-glycan structures released by reductive elimination from adult *H. contortus*.

Table 14 GC-MS analysis of partially methylated alditol acetates obtained from L3 stage *H. contortus* N-glycans released by PNGase F.

Table 15 GC-MS analysis of partially methylated alditol acetates obtained from L3 *H. contortus* N-glycans released by PNGase A.

Table 16 GC-MS analysis of partially methylated alditol acetates obtained from L3 infective stage *H. contortus* O-glycans released by reductive elimination.

Table 17 The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 77) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of adult *D. viviparous*.

Table 18 GC-MS analysis of partially methylated alditol acetates obtained from adult *D. viviparous* O-glycans released by reductive elimination (35% acetonitrile fraction).

Table 19 showing the proposed O-glycan structures released by reductive elimination from adult *D. viviparous*.
Table 20 GC-MS analysis of partially methylated alditol acetates obtained from L3 stage *D. viviparus* N-glycans released by PNGase F (35% acetonitrile fraction)

Table 21 The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 85) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of L3 stage *D. viviparus*

Table 22 GC-MS analysis of partially methylated alditol acetates obtained from L3 stage *D. viviparus* O-glycans released by reductive elimination (35% acetonitrile fraction)

Table 23 showing the proposed O-glycan structures released by reductive elimination from L3 stage *D. viviparus*
Chapter 1

Introduction
1.1 Nematodes and their control methods

1.1.1 Introduction to nematodes

Nematodes or roundworms are extremely diverse multicellular invertebrate organisms belonging to the phylum Nematoda (see Figure 1). About 80 percent of all multicellular terrestrial animals are nematodes which establish them as the most abundant multicellular organisms on Earth (Platt, 1994). Their habitats include terrestrial, fresh water and marine environments and in the latter they comprise 90% of all multicellular organisms (Danovaro et al., 2008). In 2011 the nematode species Halicephalobus mephisto was reported to be found in mines in South Africa at depths of 1.3 km. The authors speculated that this finding ‘has important implications for the search for subsurface life on other planets in our Solar System’ (Borgonie et al., 2011).

Although species like Placentonema gigantissimum infecting sperm whales can be up to a meter long, the majority of nematodes are less than 1 mm in length. They feed on different sources of food with most of them being harmless to humans and higher order organisms. For example, free-living Caenorhabditis elegans is usually found in compost heaps where it feeds on bacteria and small eukaryotes (Felix and Braendle, 2010). This organism has also attracted a lot of scientific attention. Sydney Brenner, John Sulston and Robert Horvitz won the Nobel Prize in 2002 using Caenorhabditis elegans as a model organism for their research on organ development and apoptosis. Some nematodes are parasites infecting humans, animals and plants. For example, Oncocerca gibsoni infects cattle whereas the closely related species Oncocerca volvulus infects humans. Trichinella spiralis is found in a wider range of hosts, namely rodents, pigs, horses, humans. Closely related species Toxocara canis and Toxocara cati are known parasites of dogs and cats, respectively. Other species such as Haemonchus contortus and Dictyocaulus viviparus infect small ruminants and cattle and these two parasites are the nematodes of my research interest.
**Figure 1** The phylogenetic tree of Nematode species.
The figure has been taken from (Liu et al., 2013) and modified. The nematode species mentioned later in the thesis are circled in black. The close relationship of *Dictyocaulus viviparus* parasite to Trichostrongloidea, Strongyloidea and Ancylostomatoida families is based on (Gasser et al., 2012). Another parasite not shown but mentioned in my thesis, *Ostertagia ostertagi* is close to *Ancylostoma* and *Necator* species as based on (Abubucker et al., 2009).
1.1.2 Introduction to *Haemonchus contortus* and its life cycle

*Haemonchus contortus* is an economically important gastrointestinal parasite infecting small ruminants such as goats and sheep. By residing inside the animal and feeding on its blood, the parasite causes anaemia, weight loss, reduction of milk production and occasional death of the host thus causing huge economic losses worldwide. *H. contortus* is usually found in warmer countries such as Australia, New Zealand, South Africa and Southern U.S.A, however likely due to global climate change the incidence of haemonchosis in the UK is becoming more common. A study conducted in 2012 (Burgess et al., 2012) found *H. contortus* on about fifty percent of farms across the UK confirming its widespread presence in this country.

As originally described in 1915 by Dr Frank Veglia (Veglia, 1916), during its three week life cycle *H. contortus* undergoes different morphological stages (see Figure 2) and recent genomic and transcriptomic analyses provided additional molecular details of the changes happening at each developmental stage (Laing et al., 2013; Schwarz et al., 2013) (see Figure 2). When eggs within the faeces are excreted from the animal, the eggs are exposed to the harsh environment outside the animal. This is consistent with big changes of the transcriptome of *H. contortus* during egg to L1 stage transition (Laing et al., 2013; Schwarz et al., 2013). A total of 1621 genes are up-regulated in the L1 stage. Of those 641 code for various channels such as ligand-gated channels and G-protein-coupled receptors, 43 code for transcription factors, 100 code for kinases and 35 for phosphatases most likely reflecting the adaptation to changes in chemo-sensation, mechano-sensation, osmo-sensation and proprioception (Laing et al., 2013; Schwarz et al., 2013).

Upon larval transition from L1 to L2 and consequently to the L3 developmental stage, the down-regulation of earlier up-regulated genes is observed (Laing et al., 2013; Schwarz et al., 2013). 50 and 85 percent down-regulations of gene expression are observed upon L1 to L2 and L2 to L3 stages respectively (Schwarz et al., 2013) with many down-regulated genes related to motor activity and metabolic processes (Laing et al., 2013; Schwarz et al., 2013). These observations are consistent with ensheathment of L3 stage larva which are no longer able to feed and survive on accumulated resources at lower metabolic rates (Veglia, 1916). On the other hand genes involved in oxygen transport and haem binding as well as various oxidoreductases are up-regulated (Laing et al., 2013). The latter results are consistent with earlier findings of up-regulated levels of cytochrome P450 activity in larval, but not adult,
stages of *H. contortus* (Kotze, 1997). The increased activity of cytochrome P450 is most likely related to degradation of endogenous waste accumulated during morphological changes in the oxygen-rich environment.

Upon animal grazing on the grass, the infective stage L3 organism gets inside the animal and undergoes a new boost of gene up-regulation upon transition to the L4 stage (Laing et al., 2013; Schwarz et al., 2013). Many of the up-regulated genes are related to motor activity and structural changes such as collagen-containing cuticle development (Laing et al., 2013; Schwarz et al., 2013). In addition more than 120 peptidases such as metalloproteases, aspartic and cysteine proteases are up-regulated in the L4 stage as well as the adult stage (Schwarz et al., 2013) which can be explained by the need to digest proteins for the blood-feeding L4 and adult stage parasites. Other transcriptional changes upon L4 stage larva transition to the adult parasite include a set of gender-specific genes resulting in male and female parasites (Laing et al., 2013; Schwarz et al., 2013). Transcriptional changes to deal with reactive oxygen species upon L4 to adult transition are observed such as up-regulation of glutathione-S-transferase, Cu/Zn superoxide dismutase and catalase (Schwarz et al., 2013). However L4 to adult transition undergoes less transcriptional changes overall when compared to egg-L1 or L3-L4 transitions (Laing et al., 2013; Schwarz et al., 2013).
1.1.3 Introduction to *Dictyocaulus viviparus* and its life cycle

*Dictyocaulus viviparus*, commonly known as bovine lungworm, is an economically important parasitic nematode which infects cattle in temperate regions across the world. By infecting the lungs of the cattle *D. viviparus* causes parasitic bronchitis commonly known as ‘husk’ resulting in weight loss of the animal and occasional death thus bringing huge economic losses to the farming industry. It has a complex life cycle (see **Figure 3**) and analyses of the transcriptomes for some of the parasitic stages provided more insights into the molecular changes occurring throughout the life cycle (Cantacessi et al., 2011; Ranganathan et al., 2007; Strube et al., 2012).

The life cycle of *D. viviparus* begins when worms residing in the cattle lungs lay eggs into the bronchi. Eggs are coughed up by the animal and then swallowed into the gastrointestinal
tract where they develop into the L1 larval stage. The L1 stage together with faeces is excreted and undergoes transition into L2 and L3 stage larvae. The faeces where L1, L2 and L3 stage larvae reside support the growth of *Pilobolus* fungi. The L3 stage larvae infect these fungi and reside in the sporangium awaiting the maturation of spores. When fungal spores get dispersed the infective L3 larvae spread around. Cattle graze the grass and get infected with L3 stage larvae. Although no genomics/transcriptomics analyses are available for eggs, L1 and L2 stages, transcriptomic analyses of the infective L3 stage, when compared to adults, revealed up-regulation of astacin metalloproteases (Cantacessi et al., 2011) which are possibly involved in the exsheathment and free-living to parasitic stage transition of this nematode (Cantacessi et al., 2011). Compared to the later L5 stage, up-regulation of the cleavage stimulating factor, cstf-64 was reported in the L3 stage (Strube et al., 2012). The cstf-64 protein forms part of the Poly-(A⁺) machinery needed for poly-adenylation and 3-end cleavage of mammalian pre-mRNA (Takagaki et al., 1990) and was also shown to be involved in *C. elegans* gene silencing (Kim et al., 2005). Other genes such as those coding for collagen and serine-threonine kinases were expressed in both L3 and adult stage parasites (Cantacessi et al., 2011).

The ingested *D. viviparus* L3 stage penetrates the intestinal wall, reaches the lymph nodes and there matures into the L4 stage larvae. The L4 stage larvae use the bovine blood circulation system to reach the lungs where they mature into pre-adult L5 and eventually adult stage parasites. Unfortunately no genomics/transcriptomics data is available yet for L4 stage parasite. Pre-adult L5 stage transcriptomic analysis, compared to L3 transcriptomics, revealed up-regulation of genes belonging to SCP/TAPS family (Strube et al., 2012) which have been suggested to play a role in host-pathogen interactions and defence mechanisms (Cantacessi and Gasser, 2012). In addition, compared to L3 larvae, cysteine proteases and H-Gal-GP orthologs are also up-regulated in the L5 stage and their possible roles in nutrition, digestion, host-pathogen interactions have been proposed (Strube et al., 2012). The adult stage parasite transcriptome revealed the expression of common house-keeping genes such as ribosomal proteins, aldolases and kinases (Ranganathan et al., 2007). In addition the transcription of serine-, cysteine- and metallo-proteases has been demonstrated and their potential role in tissue migration and feeding has been proposed (Ranganathan et al., 2007). The expression of genes coding for protease inhibitors have been shown (Ranganathan et al., 2007), which may protect the parasite from host proteases thus helping to establish infection in the lungs (Knox, 2007).
1.1.4 Anthelmintics as a major control of nematode infection

Anthelmintics are small molecule-based chemicals which are the major choice in controlling nematode infections. Anthelmintics have a broad range of activity mechanisms and can be divided into the three major classes, namely benzimidazoles, imidathiazolates-tetrahydropyrimidines and macrocyclic lactones. Benzimidazoles bind to the microtubule-forming protein tubulin, and block the polymerization of one end of the microtubule whereas the other end of the microtubule gets depolymerized thus disrupting the cytoskeleton of nematode cells (Lacey, 1990). Imidathiazolates-tetrahydropyrimidines such as levamisole and pyrantel act as nicotinic acetylcholine receptor agonists and by stimulating these receptors they cause calcium ion influx inside muscle cells and resultant muscle spasms (Williamson et al., 2009). In addition levamisole has been shown to exert an immunomodulatory effect on the host immune system thus having an additional role as an adjuvant (Renoux, 1980). Macro cyclic lactones such as ivermectin disrupt the...
neurotransmission of the parasite by acting on glutamate-gated chloride channels which are found only in invertebrate organisms (Cully et al., 1994).

Although widely available and applicable to a wide range of parasitic nematodes and other species, the major concern with anthelmintics is the worldwide increase of parasite resistance to them (Kaplan, 2004; Taman and Azab, 2014), see Table 1. The benzimidazole drug thiabendazole was introduced into the market around 1961 (Kaplan, 2004). However a few years later the resistance of H. contortus was reported (Conway, 1964). The imidodiazolate-tetrahydropyrimidines drug levamisole was introduced into the market around 1970 (Kaplan, 2004) and in 1979 resistance to this and other classes of anthelmintics was reported in the nematodes Trichostrongylus colubriformis and Ostertagia circumcincta (Sangster et al., 1979). The avermectin-milbemycin drug ivermectin was introduced around 1981 and in 1988 H. contortus strains were reported with multi-drug resistance against ivermectin, closantel, rafoxanide and benzimidazole anthelmintics (van Wyk and Malan, 1988). Recently new class of anthelmintics against sheep and cattle parasites, the aminoacetonitrile derivatives (AAD), has been introduced (Scott et al., 2013). By acting on nematode-specific acetylcholine receptor subunits, AADs showed the ability to affect parasitic nematodes displaying multi-drug resistance (Kaminsky et al., 2008). However in 2013 resistance to this drug was also reported (Scott et al., 2013).

<table>
<thead>
<tr>
<th>Drug (class)</th>
<th>Year introduced</th>
<th>Year resistance reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thiabendazole (Benzimidazole)</td>
<td>1961 (Kaplan, 2004)</td>
<td>1964 (Conway, 1964)</td>
</tr>
<tr>
<td>Levamisole (Imidodiazolate-</td>
<td>1970 (Kaplan, 2004)</td>
<td>1979(Sangster et al., 1979)</td>
</tr>
<tr>
<td>tetrahydropyrimidines)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monepantel (aminoacetonitrile</td>
<td>2009 (Kaminsky et al., 2008)</td>
<td>2013 (Scott et al., 2013)</td>
</tr>
<tr>
<td>derivatives)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Summary of anthelmintics
Table shows representatives from each class of anthelmintics, year the drug has been introduced and year the resistance against the drug has been reported for the first time. The introduction of the drug may vary slightly for different countries.
1.1.5 Currently available nematode vaccines

In 1959 the vaccine against *D. viviparus* called Huskvac’ or ‘Dictol’ has been developed (Jarrett et al., 1959; Jarrett et al., 1960a, b) and for a very long time it was the only nematode vaccine commercially available. ‘Huskvac’ is based on attenuated *D. viviparus* larvae and is effective in reducing the parasite burden. However as most attenuated vaccines it has a range of drawbacks such as the need for the parasite donor, short shelf-life and batch-to-batch variations (Bain, 1999).

In October 2014 the second nematode vaccine called ‘Barbervax’ was commercially launched against *H. contortus* parasite. The vaccine is based on ‘hidden’ native antigens purified from infected sheep. Similarly to attenuated larva-based vaccines although effective it has number of drawbacks such as the need for animal donors and potential batch-to-batch variations. In addition there is no boosted response after challenge infection and hence the need for repeated vaccinations. For these reasons cheap, effective and well-defined synthetic molecular based vaccines for *H. contortus* and *D. viviparus* would be preferable.

1.1.6 Unsuccessful attempts to create recombinant vaccines against nematode infections

Vaccination trials against *H. contortus* identified a range of parasite proteases as major vaccine targets (Bethony et al., 2006). One such candidate, H11 is a membrane-associated glycoprotein from adult stage intestinal microvilli (Munn et al., 1993) having microsomal aminopeptidase characteristics (Smith et al., 1997). Vaccination of lambs with H11 resulted in a substantial decrease of parasite load (Munn et al., 1993). Another major vaccine candidate, H-gal-HP is a multi-protein complex of at least 12 proteins, mainly proteases found in the microvillar membrane of the parasite intestine (Smith et al., 1999). Vaccination of lambs using different components of H-gal-HP multi-protein complex showed up to 72% and up to 93% reduction of worm and egg burdens, respectively, in sheep (Smith et al., 2003). Recombinant Hc15/24 (Vervelde et al., 2002) and HCp26/23 proteins (Garcia-Coiradas et al., 2010), cysteine proteases (Redmond and Knox, 2006) and galectins (Yanming et al., 2007) have been expressed in *E. coli*. However no successful immunization has been demonstrated. Alternative expression systems capable of protein glycosylation have been attempted with proteases from the G-Gal-GP complex being expressed in an insect cell line.
(Cachat et al., 2010) and H11 protein being expressed both in insect cell line (Reszka et al., 2007) and C. elegans expression systems (Roberts et al., 2013; Zhou et al., 2014). However, no significant protection levels have been achieved. Excretory-secretory (ES) proteins from D. viviparus showed promising results in inducing significant levels of protection in a guinea pig model (McKeand et al., 1994a, b). However vaccination of calves with native ES products and a single recombinant component, acetylcholinesterase showed no significant protection against parasites in vaccinated animals (Matthews et al., 2001).

Amongst the potential reasons for the lack of success of recombinant protein-based vaccines is the requirement for minor unidentified components in native extracts as well as potential incorrect protein folding and post-translational modifications in recombinant material. The choice of adjuvant can also have a significant impact on the final worm burden reduction. Using the L3 stage-purified HcsL3 antigen together with pertussis (Jacobs et al., 1999) or Freund’s adjuvants (Turnbull et al., 1992) vaccination of sheep resulted in significant boost of antibody responses without significant protection. On the other hand, aluminium and DEAE-dextran adjuvants for the same HcsL3 antigen resulted in reduced cumulative faecal egg count and worm burden with the latter adjuvant being almost twice as effective as aluminium (Piedrafita et al., 2013). Interestingly the addition of Quil A adjuvant resulted in abrogation of protection in lambs (Piedrafita et al., 2013). Different outcomes of vaccination using the same antigen but different adjuvant could be partially explained by the fact that the QuilA adjuvant is known to promote a Th1 immune response (Sun et al., 2009). On the other hand, aluminium adjuvants are known to bias immune responses towards the Th2 response (Marrack et al., 2009) with eosinophil recruitment, which may be beneficial in clearing helminth infections (Meeusen and Balic, 2000). Interestingly the use of DEAE-dextran adjuvant also showed up regulation of IgE antibodies and eosinophils in vaccinated lambs suggesting a Th2-type immune response (Piedrafita et al., 2013).

**1.1.7 The development of carbohydrate-based vaccines**

Considering the unsuccessful attempts to create recombinant vaccines against nematode parasites discussed in section 1.1.6, new approaches are urgently needed and carbohydrate-based vaccines can be considered as an alternative to create a well-defined molecular approach-based vaccine (Astronomo and Burton, 2010). Carbohydrate-based human vaccines
have already been successfully utilized for several years. Currently there are licensed carbohydrate-based vaccines against *Haemophilus influenzae* (type b), *Neisseria meningitides* serotypes A, C, W135, Y, *Salmonella typhi, Staphylococcus aureus* and *Streptococcus pneumoniae* bacteria (Huang and Wu, 2010). However, bacteria tend to be easier to culture resulting in easier access of potential antigens, whereas parasites such as *H. contortus* and *D. viviparus* are not readily available as they require host donors to produce them. Therefore together with advancements in carbohydrate synthesis methodologies it is more attractive to generate synthetic carbohydrate-based vaccines (Astronomo and Burton, 2010).

Although an attractive target for vaccine development, carbohydrate molecules are usually considered as poor immunogens as they activate B-cells and trigger antibody production, but in most cases they are not presented on MHC II complex and according to previously held beliefs do not interact with T-cell receptors (TCR) on CD 4⁺ T-helper cells (Mond et al., 1995a; Mond et al., 1995b). The resultant T-cell-independent response is known to be strong but short-lived without immunological memory and without antibody class-switching mainly consisting of IgM antibodies (Mond et al., 1995a; Mond et al., 1995b). For these reasons for most carbohydrates to elicit the long-lived immunogenic responses, their conjugation to an immunogenic protein is required, ensuring the generation of immunological memory. Carrier proteins which are known to be highly immunogenic such as diphtheria toxin CRM197, keyhole limpet haemocyanin (KLH), tetanus toxoid (TT) and bovine serum albumin (BSA) have been used (Astronomo and Burton, 2010) (Nyame et al., 2004). The choice of linker molecule must be considered as it needs to provide appropriate chemistry to link the carbohydrate and protein, as well as being flexible and not triggering strong immune responses in the vaccine recipient. A schematic diagram of the synthetic glycoconjugate vaccine candidate against trematode *Schistosoma mansoni* developed by Cummings and colleagues (Nyame et al., 2004) is shown in Figure 4. In this case the modified asparagine-linked N-glycan with two GalNAcβ1-4GlcNAc (LDN) antennae (see section 1.2.1) has been reacted with 2-iminothiolane (Traut’s reagent) and subsequently with maleimide-activated carrier protein BSA (Nyame et al., 2004).
Recently, the cellular and molecular mechanisms of how glycoconjugate carbohydrate vaccines trigger the immune response have been revealed as shown in the Figure 5 (Avci et al., 2011). The glycoconjugate vaccines have been shown to be depolymerized in the antigen-presenting cell (APC) endosome (Avci et al., 2011). Superoxide and hydroxyl, but not hydroperoxide radicals have been demonstrated to be responsible for the carbohydrate portion degradation (Avci et al., 2011). Then in contrast to previous belief, Kasper and co-workers demonstrated that the carbohydrate portion of the glycoconjugate vaccine attached to the peptide can be presented onto the MHC II complex. They proposed that the peptide portion interacts with the MHC II molecule and the carbohydrate part interacts with the T-cell receptor (TCR) of CD 4+ T-helper cells resulting in IL-2 and IL-4 secretion and activation of antibody-producing B-cells with subsequent isotype switching and immunological memory (Avci et al., 2011). Furthermore, to support the carbohydrate-TCR interactions, they isolated T-helper cells which recognized carbohydrates only and termed them T<sub>carbs</sub> cells (Avci et al., 2012). In addition, bacterial zwitterionic capsular polysaccharides carrying alternating positive and negative charges (Cobb et al., 2004), CD1d molecule-bound glycolipids (Brigl and Brenner, 2004) and glycopeptides (Dzhambazov et al., 2005; Gad et al., 2002) have all been demonstrated to be able to be recognized by T-cells. These observations indicate that it is not the recognition of carbohydrates by T-cells, but their presentation onto MHC-like molecules, that could be a limiting factor in causing long-lived immunogenic responses.
Since carbohydrate synthesis in organisms is not template driven it is therefore technically challenging to produce large quantities of specific carbohydrates. This contrasts to the situation with DNA and proteins where amplification technologies such as PCR and cloning and expression systems can be utilized. Bearing in mind the complexity of carbohydrate chemistry, due to the very similar properties of different hydroxyl groups present in monosaccharides and the importance of linkage position between them, carbohydrate synthesis is still not an easy task. Various chemo- and/or enzymatic strategies have been employed together with genetically modified expression systems to produce desired glycans or glycoproteins for various applications (Mrazek et al., 2013). However, in vaccine development chemical synthesis still remains the method of choice.
The first fully synthetic carbohydrate-based vaccine was developed in 1988 in Cuba against \textit{Haemophilus influenza} type b pathogen, widely known as Hib vaccine it was involved in a national vaccination programme in that country. This vaccine is based on ribosylribitol oligomers attached to thiolated TT protein (Verez-Bencomo et al., 2004). Other examples of fully synthetic carbohydrate-based vaccines are against \textit{Streptococcus pneumonia} and \textit{Neisseria meningitides} bacteria (Astronomo and Burton, 2010). Although there are no vaccines against human parasitic infections and only a few vaccines are commercially available to tackle life stock parasites (Knox and Redmond, 2006) there is a range of attempts going on in this field and quite a few of them exploit carbohydrate epitopes.

Seeberger and colleagues used an automated solid-phase synthesis to create a Man4-glycosylphospatidylinossitol (GPI) core structure found in the malaria parasite \textit{Plasmodium falciparum} and conjugated it to ovalbumin (OVA) and KLH proteins (Hewitt et al., 2002; Schofield et al., 2002). Although mice immunized with the synthetic vaccine showed reduced mortality it did not prevent parasite proliferation. Further development of this anti-malarial vaccine is on-going. Another example of a potential future vaccine candidate is the synthetically synthesised Manα1-2Manα1-2(Galβ1-4)Man1 tetrasaccharide which caps the lipophosphoglycans (LPG) in \textit{Leishmania donovani} parasites. This synthetic epitope was linked to succinimide-activated influenza hemagglutinin (HA) and was loaded into immune-potentiating reconstituted influenza virosome (IRIV) constructs (Liu et al., 2006) which also act as an adjuvant (Gluck, 1992). Immunization studies showed that mice produced glycan specific IgG antibodies that recognize \textit{Leishmania} parasites demonstrating its potential as a vaccine candidate against this human and other mammalian parasite (Liu et al., 2006).
1.2 Structural and functional Glycomics in nematode species

1.2.1 Introduction into glycobiology

[Figure 6 Most common monosacharides found in nematodes. The symbolic nomenclature is used in the thesis as outlined by the Consortium for Functional Glycomics Nomenclature Committee (May 2004) (Varki et al., 2009). Full documentation available from: http://glycomics.scripps.edu/CFGnomenclature.pdf.

Carbohydrates are the most abundant class of biomolecules on Earth and for a long time have been mainly considered as energy and structural molecules. In the last few decades this perception has changed with the advent of the glycobiology field which studies the structure and function of sugars in cells and organisms. The research has shown that this class of molecules is extremely important not only providing structural support for cells but also maintaining the integral three-dimensional shape of proteins (Wormald and Dwek, 1999), performing protein folding quality control, intracellular trafficking (Aebi et al., 2010) and also mediating cell-cell communication in multicellular organisms (Cummings and Pierce, 2014).
Glycans are sugar molecules made of one or more monosaccharides (see Figure 6) which are attached to proteins or lipids forming glycoconjugates. In contrast to RNA transcription or protein translation the synthesis of glycans is not a template driven process and requires a range of enzymes called glycosyltransferases and glycosidases to add or remove monosaccharides from the growing oligosaccharide chain. Different glycosyltransferases or glycosidases are usually required to add or remove different monosaccharides in a linkage specific manner. Each monosaccharide can be added to different hydroxyl groups of the acceptor mono/oligosaccharide resulting in different glycans with different properties thus tremendously increasing the biodiversity of glycans and the range of functions they can perform.

There are two types of glycoconjugates in the cell. Glycoproteins can be N-glycosylated meaning the sugar portion is attached to the protein via a nitrogen atom or O- glycosylated meaning glycans are attached via an oxygen atom to the protein and glycolipids. Usually glycoproteins and glycolipids are found on the surface of the cell and form a dense layer called the glycocalyx. Glycoconjugates are also found in the extracellular matrix (ECM) and body fluids. In addition a specific form of protein glycosylation called O-GlcNAc-ylation is found in the cytoplasm and nucleus of cells (Hardiville and Hart, 2014).

1.2.2 N-glycan biosynthesis in nematodes

1.2.2.1 Overview of N-linked glycosylation in nematodes

N-linked glycosylation is a post-translational protein modification present in all eukaryotes. The early stages of this process include the synthesis of a lipid-linked oligosaccharide precursor which is subsequently transferred to a protein. These stages are conserved amongst eukaryotic organisms with the majority of the experimental elucidation coming from yeast studies. Later stages of N-glycan biosynthesis, which involve both hydrolysis and addition of new monosaccharides, can be substantially different amongst species resulting in different glycan structures. The novel features of nematode N-linked glycosylation include N-glycans with unusual chitobiose cores and non-reducing antennae modifications often not seen in other organisms. In this chapter I will describe in detail N-glycan biosynthesis in the non-parasitic Caenorhabditis elegans nematode as this model organism is the most heavily researched and therefore best characterised nematode.
1.2.2.2 The biosynthesis of the lipid-linked Glc3Man9GlcNAc2 N-glycan precursor and its transfer to the protein

![Diagram of the biosynthesis of the lipid-linked Glc3Man9GlcNAc2 N-glycan precursor and its transfer to the protein in the ER.](image)

**Figure 7** The biosynthesis of lipid-linked GlcNAc2Man9Glc3 N-glycan precursor and its transfer to the protein in the ER. The *C. elegans* transcripts corresponding to eukaryotic genes used in this diagram can be found in the Table 2.

The biosynthesis of the lipid-linked Glc3Man9GlcNAc2 N-glycan precursor with the subsequent transfer of the N-glycan onto the protein (see Figure 7) is a conserved process amongst eukaryotes. Understanding of it came mainly from studies on yeast protein glycosylation (Aebi, 2013), however corresponding genes have been found in the nematode model organism *C. elegans* (Akiyoshi et al., 2015; Schachter, 2004).

The biosynthesis of N-glycans starts on the cytoplasmic side of the Endoplasmic Reticulum (ER) membrane. The Man9GlcNAc2 precursor is assembled on the lipid dolichol made of condensed five-carbon isoprene units. The monosaccharide donors are UDP-GlcNAc and GDP-Man. Subsequently the Man9GlcNAc2-dolichol precursor is translocated across the ER membrane. Although the identity of the flippase protein remains unknown, in 2008 Sanyal and colleagues provided the biochemical basis for Man9GlcNAc2-dolichol flippase. They suggested it sediments at 4S, is trypsin-sensitive and operates in an ATP-independent manner.
by facilitated diffusion (Sanyal et al., 2008). Later they showed that this protein candidate is very specific for the lipid-linked Man5GlcNAc2 oligosaccharide. It requires dolichol made of α-saturated isoprene (Sanyal and Menon, 2010), an exact number of mannoses present in a specific configuration (Sanyal and Menon, 2009, 2010) and specific oligosaccharide molecule enantiomer (Sanyal and Menon, 2010). After the Man5GlcNAc2-dolichol precursor has been flipped across the ER membrane, addition of sugar residues continues inside the ER with the oligosaccharide chain becoming Glc3Man9GlcNAc2 Interestingly the RNAi silencing of most genes responsible for lipid-linked oligosaccharide synthesis in the cytoplasmic, but not luminal side of the ER such as human alg7, alg13, alg14, alg2 and alg11 orthologs in C. elegans (see Figure 7) cause severe phenotype changes (Akiyoshi et al., 2015) confirming the importance of these genes in worm development.

<table>
<thead>
<tr>
<th>Eukaryotic gene</th>
<th>C. elegans transcript</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>alg7</td>
<td>T08D2.2, Y60A3A.14</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>alg13</td>
<td>R10D12.12</td>
<td>(RNAi)</td>
</tr>
<tr>
<td>alg14</td>
<td>M02B7.4</td>
<td>(RNAi)</td>
</tr>
<tr>
<td>alg1</td>
<td>T26A5.4</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>alg2</td>
<td>T23F2.1, F09E5.2</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>alg11</td>
<td>B0361.8</td>
<td>(RNAi)</td>
</tr>
<tr>
<td>alg3</td>
<td>K09E4.2</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>alg9</td>
<td>C14A4.3</td>
<td>(Schachter2004),(RNAi)</td>
</tr>
<tr>
<td>alg12</td>
<td>ZC513.5</td>
<td>(RNAi)</td>
</tr>
<tr>
<td>alg6</td>
<td>C08B11.8</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>alg8</td>
<td>C08H9.3</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>alg10</td>
<td>T24D1.4</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
<tr>
<td>OST subunits-coding genes</td>
<td>T22D1.4, M01A10.3, ZK686.3, T09A5.11, F57B10.10, T12A2.2,</td>
<td>(Schachter2004), (RNAi)</td>
</tr>
</tbody>
</table>

Table 2 The C. elegans genes corresponding to eukaryotic analogues shown in the Figure 7
The synthesised dolichol-linked Glc₃Man₉GlcNAc₂ oligosaccharide is then transferred onto the protein in the lumen of the ER by the enzyme oligosaccharyltransferase (OST) (Knauer and Lehle, 1999). For the oligosaccharide chain to be added onto the protein it needs to meet specific requirements. Firstly the N-glycan is added onto asparagine residues in the protein sequence Asn-X-Ser or Asn-X-Thr where X can be any amino acid except proline. There are rare cases when the protein is glycosylated at Asn-X-Cys sequence (Chi et al., 2010). It is important to note that not all Asn-X-Ser or Asn-X-Thr sequences get glycosylated. In terms of protein 3D shape the N-glycans are usually found on the surface of the protein making good energetic rational for relatively hydrophilic protein modification and not to be in the hydrophobic interior of the protein. Finally, as the addition of the N-glycan chain occurs in the lumen of the ER it ensures that only the extracellular portions of plasma membrane proteins and secretory proteins acquire N-glycan modifications. The RNAi silencing of gene orthologs coding for OST in *C. elegans* resulted in a germline phenotype and small body size of worms (Akiyoshi et al., 2015) indicating the importance of this protein in organism development.

### 1.2.2.3 The conversion of Glc₃Man₉GlcNAc₂ N-glycan into truncated Man₅GlcNAc₂ structure based on model nematode *C. elegans* studies

Once the Glc₃Man₉GlcNAc₂ oligosaccharide is transferred to the protein the trimming of the hexoses begins in the ER and Golgi apparatus. Initially *C. elegans* α1-2 ER glucosidase I MOGS-1 trims the terminal glucose and subsequently *C. elegans* α1-3 ER glucosidase II AAGR-3 removes the other two glucoses. The functions of MOGS-1 and AAGR-3 enzymes are proposed based on sequence homology and have not been cloned and characterized (Schachter, 2004).

In 2012 α-mannosidases homologous to the yeast *S. cerevisiae* class I mannosidases were discovered in *C. elegans* which act on a Man₉GlcNAc₂ N-glycan, namely MANS-1, MANS-2, MANS-3 and MANS-4 (Wilson, 2012). ER-resident MANS-3 and MANS-4 from *C. elegans* α-mannosidases remove one α-mannose from the ‘middlemost’ branch of the Man₉GlcNAc₂ N-glycan (Wilson, 2012). Golgi-resident *C. elegans* MANS-1 and MANS-2 enzymes efficiently remove up to 3 α-mannose residues, two from the ‘lower’ α1-3 branch and one from the ‘upper’ branch (Wilson, 2012). Together these four enzymes result in the conversion of N-linked Man₉GlcNAc₂ glycan to the Man₅GlcNAc₂ structure.
The next step in C. elegans N-glycan biosynthesis is the addition of GlcNAc to the ‘lower’ α1-3 branch of the Man₅GlcNAc₂ substrate by the Golgi-resident Gly-13 gene-coded GlcNAc transferase which is a mammalian GlcNAcT I homologue (Chen et al., 2003). There are two other enzymes coded by gly-12 and gly-14 genes which are able to carry out the same reaction in C. elegans. However, only gly-13 gene knockout worms are arrested in L1 developmental stage making GLY-13 the major ‘lower’ α1-3 branch GlcNAc transferase in these nematodes (Chen et al., 2003).

After addition of GlcNAc to the ‘lower’ α1-3 branch of the Man₅GlcNAc₂ N-glycan, the Golgi resident C. elegans mannosidase II coded by aman-2 gene removes both terminal α1-3 and α1-6 mannoses (Paschinger et al., 2006). Subsequently the putative C. elegans α-mannosidase (which to date has not been characterized) removes the α1-6 mannose resulting in a Man₂GlcNAc₃ structure which can undergo other structural modifications as described further. All these N-glycan modifications from Glc₃Man₉GlcNAc₂ to truncated Man₂GlcNAc₃ N-linked glycan structure described in this section are depicted in Figure 8.

**Figure 8** The biosynthesis of truncated Man₂GlcNAc from Glc₃Man₉GlcNAc₂ N-glycan in the ER and Golgi compartments. Those C. elegans proteins which have been cloned and characterized are put into circles.
1.2.2.4 Chitobiose core modifications of Man$_2$GlcNAc$_3$ N-linked glycan in *C. elegans* model nematode

**Figure 9** The chitobiose core modification pathways in *C. elegans*. The addition of α1-2 galactose to the distal GlcNAc-linked fucose by putative α1-2 galactosyltransferase is not shown. The figure is based on (Yan et al., 2013).

The Chitobiose core of *C. elegans* N-linked glycans can be modified with up to three fucoses attached to the 3-positions of proximal and distal GlcNAc and the 6-position of proximal GlcNAc. Three different fucosyltransferases are responsible for these modifications. The fucoses in turn can be capped with galactoses (Hanneman et al., 2006; Titz et al., 2009; Yan et al., 2012). The biosynthetic machinery for these modifications in *C. elegans* has been partially elucidated (see Figure 9).

The first fucosyltransferase characterized in 2004 by I.B.H. Wilson and colleagues was FUT-1 which adds fucose to the 3-position of the proximal GlcNAc (Titz et al., 2009). The action of the enzyme was inhibited by the presence of non-reducing end terminal GlcNAc (Titz et al., 2009) which was later re-confirmed by the same group in 2013 (Yan et al., 2013). N-linked glycans carrying the 6-linked fucose on the proximal GlcNAc without non-reducing terminal GlcNAc can also act as substrates for the FUT-1 enzyme (Yan et al., 2013). The
authors also proposed that FUT-1 enzyme should work on N-glycans bearing the 3-linked fucose on the distal GlcNAc although it has not been proven directly (Yan et al., 2013).

Another C. elegans FUT-8 fucosyltransferase transfers fucose to the 6-position of the proximal GlcNAc (Paschinger et al., 2005). In contrast to FUT-1 enzyme it requires the non-reducing terminal GlcNAc to be present in order for the proximal GlcNAc 6-position fucosylation to occur, hence its requirement for the prior action of GlcNAc transferase I carried out mainly by the GLY 13 enzyme in C. elegans as mentioned earlier (Chen et al., 2003). In addition FUT-8 enzymatic action is inhibited by the presence of α1-3 fucose on either the proximal or distal GlcNAc (Paschinger et al., 2005; Yan et al., 2013).

The last of the three C. elegans chitobiose core fucosyltransferases characterized was the FUT-6 enzyme which fucosylates the distal GlcNAc at the 3-position (Yan et al., 2013). In terms of its substrate specificity it has been determined that the non-reducing terminal GlcNAc can be present or absent for the reaction to occur (Yan et al., 2013). It also acts on FUT-8 reaction products carrying fucose 6-linked to the proximal GlcNAc (Yan et al., 2013). The authors also predicted the enzyme’s ability to fucosylate FUT-1 reaction products carrying α1-3-linked fucose to the proximal GlcNAc although more evidence is needed to support this (Yan et al., 2013). Interestingly the FUT-6 enzyme is blocked by the presence of α1-6, but not α1-3-linked mannose on the trimannosyl core (Yan et al., 2013), hence its pre-requisite for the action of the putative α1-6 mannosidase.

All aforementioned fucoses in C. elegans can be capped with galactose (Hanneman et al., 2006; Titz et al., 2009; Yan et al., 2012). The galactose capping the fucose 6-linked to the proximal GlcNAc has been proven to be β1-4 linked (Hanneman et al., 2006). The GALT-1 galactosyltransferase carrying out this reaction in C. elegans has been characterized by I.B.H. Wilson and colleagues who showed its structural substrate preferences for the non-reducing terminal GlcNAc although in vitro the reaction can be carried out but at a slow rate (Titz et al., 2009). Later on in 2013 they showed that, similarly to the FUT-8 enzyme, the GALT-1 enzyme is inhibited by the presence of α1-3 fucose on either proximal or distal GlcNAc generated by the FUT-1 and FUT-6 enzymes respectively (Yan et al., 2013).

The galactose capping the α1-3 fucose linked to the proximal GlcNAc has been proposed to be also β1-4 linked based on MSn evidence (Hanneman et al., 2006). In contrast α1-3 fucose linked to the distal GlcNAc can be modified by the α1-2 linked galactose (Yan et al., 2012),
however the galactosylation machinery carrying out these two reactions has not been characterized yet.

1.2.2.5 Antennae modifications of N-linked glycans in *C. elegans* model nematode

**Figure 10** The N-glycan antennae modifications found in *C. elegans* organisms

Based on the current knowledge of the N-linked glycan biosynthetic pathway there is no rigorous evidence suggesting that these organisms can synthesize elongated complex N-glycans as commonly seen in mammals. The non-reducing end termini of *C. elegans* N-linked glycans can be modified with single non-reducing end terminal GlcNAc residues and corresponding GlcNAc transferases have been suggested. The mammalian GnT V analogue GLY 2 in *C. elegans* was proposed to add β1-6 linked GlcNAc to the 6-linked mannose of the trimannosyl core (Warren et al., 2002). The mammalian GnT II analogue GLY 20 was suggested to add β1-2 linked GlcNAc to the same ‘upper’ 6-linked mannose residue (Chen et al., 2002). The earlier mentioned *C. elegans* enzymes GLY -12/13/14 which are analogues to mammalian GNT I add β1-2 linked GlcNAc to the ‘lower’ 3-linked trimannosyl core mannose with the main enzyme being GLY-13 (Warren et al., 2002) (Chen et al., 2003).
The attached non-reducing end GlcNAc residues in C. elegans can be modified further by the nematode-specific post-translational modification which involves attachment of a phosphorylcholine (PC) group to the GlcNAc residue 6-position (Haslam et al., 2002). The biosynthetic details of this reaction are not completely understood however phosphorylcholine as a PC unit donor and fukutin-related family protein as potential PC transferase have been suggested (Harnett et al., 2010).

Another antennae modification present in C. elegans is capping of the trimannosyl core with galactose which can be further modified at the 2-position with fucose (Haslam et al., 2002). Some evidence also suggests fucose being attached to mannose residues (Altmann et al., 2001). Interestingly, C. elegans glycome sequencing revealed over 20 putative α1, 2-fucosyltransferases and one of them has been characterized (Zheng et al., 2002) revealing its unique substrate requirements. This C. elegans enzyme named by authors CE2FT-1 (Zheng et al., 2002) prefers terminal galactose with the penultimate residue being neither GlcNAc nor glucose thus making it a potential α1, 2-fucosyltransferase which modifies the terminal galactose in the C. elegans N-glycome. All the possible antennae modifications described for C. elegans above are summarized in the Figure 10.
1.2.3 N-linked glycosylation in *H. contortus* parasite

![Diagram of N-glycan core and antennae modifications](image)

**Figure 11** The N-glycan core and antennae modifications found in *H. contortus* organisms. High-mannose structures present in many different organisms are not shown.

The most rigorous studies on *H. contortus* N-glycosylation using mass-spectrometry-based techniques were conducted by Haslam and colleagues revealing a high degree of chitobiose core fucosylation (Haslam et al., 1996) as seen in the Figure 11. The proximal GlcNAc was shown to be able to carry up to two fucoses which are 3- and 6-linked and the distal GlcNAc is modified with an, at that time, previously unseen fucosylation at the 3-position (Haslam et al., 1996). This was subsequently shown to be stage specific as it is absent in the L3 stage of *H. contortus* (Haslam et al., 1998). The data generated from FAB-MS studies also suggested the presence of high mannose structures and complex N-glycans bearing GalNAcβ1-4GlcNAc (LDN) and GalNAcβ1-4(Fucα1-3)GlcNAc (LDNF) antennae (see Figure 11) (Haslam et al., 1996). The presence of the latter structure was also proposed by (Geldhof et al., 2005; van Stijn et al., 2010b; Vervelde et al., 2003) based on antibody responses to *H. contortus* infection and Western Blot studies. In addition vaccination studies revealed the presence of antibodies against the Galα1-3GalNAc epitope and extracts from adult *H. contortus* were shown to contain a galactosyltransferase that was able to synthesize this epitope (van Stijn et al., 2010b). Although the biosynthetic machinery carrying out these N-glycan modifications remains unknown the recent genome sequencing and generation of the extensive transcriptomics datasets for *H. contortus* (Laing et al., 2013) gives an optimistic perspective for cloning and characterizing the glycosyltransferases and glycosidases responsible for the glycome of this parasitic nematode.
1.2.4 N-linked glycosylation in *D. viviparus* parasite

The only structural studies on the glycome of *D. viviparus* were performed by Haslam and colleagues using FAB-MS instrumentation in 2000 (Haslam et al., 2000). The major N-glycan families seen were high mannose structures with general composition Hex$_{5-9}$HexNAc$_2$, truncated cores with and without fucose matching the composition Hex$_{2-4}$Fuc$_{0-1}$HexNAc$_2$, and complex bi, tri- and tetra-antennary structures with general composition of Hex$_{3-7}$Fuc$_{0-5}$HexNAc$_{3-6}$. The fragment ion analysis of these structures confirmed the presence of up to four Lewis$^x$ epitopes on these glycans. The structures are summarised in Figure 12.

Unfortunately the genome of *D. viviparus* has not been sequenced yet and no details about biosynthetic machinery of glycans are available, although the similarity of parasite N-glycans to mammalian structures infers that enzymes homologous to mammalian could operate in this nematode.
1.2.5 N-linked glycosylation in other nematode species

As mentioned earlier, the main features of nematode N-glycosylation in comparison to mammalian are antennae and chitobiose core modifications. In addition to earlier described C. elegans and H. contortus organisms the porcine roundworm A. suum was shown to have 3, 6-difucosylated proximal GlcNAc. Later the presence of Galβ-4Fuc on the 6-position of proximal GlcNAc and Galα1-2Fuc on the 3-position of the distal GlcNAc was demonstrated in this nematode (Yan et al., 2012).

The attachment of the phosphorylcholine (PC) moiety to the 6-position of non-reducing end GlcNAc in C. elegans has been described earlier (Haslam et al., 2002). This type of post-translational modification has been also found in the intracellular mammalian skeletal muscle parasite T. spiralis (Morelle et al., 2000), the bovine filarial parasite O. gibsoni (Haslam et al., 1999), the human filarial worm O. volvulus (Haslam et al., 1999), the porcine roundworm A. suum (Poltl et al., 2007) and the rodent filarial nematode A. viteae (Haslam et al., 1997).
Another N-glycan antennae modification not seen in mammals but present in *T. spiralis* parasite is capping of LDNF antennae with the rare dideoxy sugar tyvelose (Reason et al., 1994). The N-glycans of this nematode contain up to 4 LDNF epitopes and surprisingly virtually all of them are capped with this sugar (Reason et al., 1994). Other antennae present in nematodes include the Galα1-3Gal sequence in white-tailed deer parasite *P. tenuis* (Duffy et al., 2006) which is also found in its host, and the LacNAc structure in abomasal cattle nematode *O. ostertagi* (Meyvis et al., 2008) and likely in *A. vitae* (Haslam et al., 1997). The latter parasite is also likely to have LDN and LDNF antennae modifications as well, based on characteristic FAB-MS fragment ions and GC-MS data (Haslam et al., 1997). In addition chito-oligomers made of poly-GlcNAc with a capping HexNAc have been reported in three filarial nematodes (Haslam et al., 1999). Finally the most conserved and widespread high-mannose and pauci-mannose N-glycans are seen amongst a wide range of nematodes *H. contortus* (Haslam et al., 1996), *D. viviparus* (Haslam et al., 2000), *T. spiralis* (Morelle et al., 2000), *A. vitae* (Haslam et al., 1997). The summary of nematode N-linked glycosylation can be seen in the Figure 13.
1.2.6 O-glycan biosynthesis in nematodes

1.2.6.1 Overview of mucin-type O-linked glycosylation

Figure 14 Different O-glycan Core structures found in mammals

Mucin type O-linked glycosylation is the process of adding glycans to the hydroxyl group of Serine and Threonine residues in a protein. In contrast to N-glycosylation, O-glycosylation does not involve pre-synthesised and en-block glycan transfer. Instead all monosaccharides present in the structures are added one by one to the O-glycan chain. In addition O-glycans do not have a consensus sequence within the protein. However, as O-glycosylation is a post-folding event only Ser and Thr residues on the surface of a protein can be O-glycosylated. This makes intrinsically disordered regions such as linkers, coils and turns to be usually favoured which is supported by the computational O-glycosylation sites prediction strategies (Hansen et al., 1995; Nishikawa et al., 2010). This type of post-translational modification happens in the Golgi apparatus and in the same way as N-glycosylation is a non-template driven process. Glycosyltransferases add one by one sugar residues to the nascent O-linked glycans resulting in at least, in mammals, 8 different potential Core structures (Wopereis et al., 2006) as seen in Figure 14 with Core 1 and Core 2 structures being most common. The Core structures can in turn be further modified with other residues such as sialic acid residues in humans.
1.2.6.2 Mucin-type O-linked glycosylation in nematodes

In contrast to N-glycans and their biosynthetic pathways, O-glycosylation is much less characterised in nematodes even in the model organism *C. elegans*. In 1998 Hagen and Nehrke cloned and expressed 11 UDP-N-acetyl-D-galactosamine:Polypeptide-N-Acetylgalactosaminytransferases which are homologous to human pp-GalNAc-Ts (EC 2.4.1.41). This enzyme family initiates the first step in mucin type O-glycan biosynthesis by adding α-GalNAc to the protein Ser/Thr hydroxyl group (Hagen and Nehrke, 1998). The core 1 O-glycan T-synthase, which initiates the formation of Core-1 structures by adding a β-galactose onto the 3-position of GalNAc, has been cloned and characterized from *C. elegans* (Ju et al., 2006). Interestingly, in contrast to mammalian core 1 O-glycan T-synthase, the *C. elegans* analogue does not require the Cosmc chaperone for correct folding, although the fact that expression of functional enzyme is achieved in insect, not mammalian cell lines, might indicate the requirement for some unidentified invertebrate-specific factors (Ju et al., 2006).

Later another galactosyltransferase generating Core 1 structures in *C. elegans* was proposed (Palaima et al., 2010) based on knock-out studies although the gene has not been cloned and characterized.

**Figure 15** The O-glycans found in nematode parasites
The amount of characterized enzymes to date potentially involved in mucin-type O-glycan biosynthesis in the model nematode *C. elegans* does not reflect the complexity of O-glycan structures in this organism which were most rigorously characterized by Guerardel and colleagues employing both MS and NMR strategies (Guerardel et al., 2001). The study revealed an abundance of unusual Core 1 structures to which glucose can be attached to the 6-position of the reducing-end GalNAc (Guerardel et al., 2001). These structures can be further modified with glucose residues in various positions together with glucuronic acid (Guerardel et al., 2001). To add to this, β-galactose can be 6-linked to the GalNAc residue of the Core 1 structure forming the Galβ1-6(Galβ1-3)GalNAc Schisto Core structure. The Galβ1-6(Galβ1-3)GalNAc structure has been coined ‘Schisto Core’ by me for the purpose of this thesis as it has been previously characterised in the parasitic trematode *S. mansoni* (Huang et al., 2001). Finally one O-glycan in *C. elegans* has been found linked to the protein via a GlcNAc residue (Guerardel et al., 2001).

The literature information about O-glycosylation in other nematodes is rather scarce. O-glycans from the dog-infecting nematode *T. canis* have been rigorously characterized and two structures have been proposed (Khoo et al., 1991). Both of them bear the Core-1 structure with 2-methyl-α-fucose capping galactose at the 2-position. An additional methyl group at the 4-position of the Core 1 galactose can be present or absent (Khoo et al., 1991). The closely related cat-infecting *T. cati* nematode was also proposed to have similar O-glycan structures with dominant structure being a dimethylated, not monomethylated trisaccharide (Khoo et al., 1991). Another nematode that has had its O-glycans characterized is the deer parasite *P. tenuis*. MS-structural studies of its O-glycans proposed Core 1 and α1-2 fucosylated Core 1 structures to be present (Duffy et al., 2006). The O-glycans found in nematodes can be seen in the Figure 15.
1.2.7 Overview of glycolipids in nematodes

The second major class of glycoconjugates found in cells are glycolipids. These molecules contain the glycan portion attached to the lipids embedded in the plasma membrane. Together with N- and O-glycans attached to the transmembrane proteins, glycolipids constitute the layer of sugars around the cell called the glycocalyx. Glycolipids can be divided into two broad classes, glycosphingolipids and glycosphingolipids based on the type of lipid the sugar portion is attached to. The presence of glycosphingolipids together with its biosynthetic machinery has been reported in *C. elegans* where the phosphatidyl glycerol molecule is modified with inositol to which the Manα1-2Manα1-6-[PEtN-2]Manα1-4GlcNα1-6- sugar moiety is attached forming the so-called GPI anchor for proteins (Murata et al., 2012).

The second class of glycolipids in nematodes is based on glycosphingolipids in which the glycan is attached to a lipid called ceramide made of fatty acids and the long-chain amino alcohol sphingosine. All reported nematode glycosphingolipids (see Figure 16) have glucose as the first residue which in mammals is added by the action of the Glu transferase CGluT
the ER lumen and the structure is extended in the Golgi compartment by the sequential action of other glycosyltransferases (Kapitonov and Yu, 1997; Schaeren-Wiemers et al., 1995; Schulte and Stoffel, 1993). The structures of glycosphingolipids have been reported for C. elegans (Gerdt et al., 1997; Griffitts et al., 2005), O. volvulus (Wuhrer et al., 2000), A. suum (Friedl et al., 2003; Lochnit et al., 1998; van Riet et al., 2006), A. lumbricoides (van Riet et al., 2006) and A. viteae (Houston et al., 2002) nematodes. All of the reported nematode glycosphingolipids are based on the Galα1-3GalNAcβ1-4GlcNAcβ1-3Manβ1-4Glcβ-Cer structure or its precursors. Attachment of a phosphorylcholine (PC) moiety to the 6-position of the GlcNAc has been observed in all four nematodes with phosphorylethanolamine being attached to the same position in A. suum (Lochnit et al., 1998; van Riet et al., 2006). A. suum and O. volvulus can also have PC attached to the 6-position of the GalNAc residue (Lochnit et al., 1998; Wuhrer et al., 2000). A. lumbricoides glycosphingolipids have been shown to have α-galactose modified with β-galactoses at 3- and 6- positions (van Riet et al., 2006) and the closely related A. suum most likely has the same terminal glycosphingolipid modifications which can be further modified with Hex, Fuc and HexNAc residues (Friedl et al., 2003). In C. elegans the non-reducing end has been shown to be modified with other galactose residues which in turn are modified with fucose and 2-O-methylated fucose residues (Griffitts et al., 2005).

1.2.8 Functions of parasite glycans in shaping the immunity of the host

1.2.8.1 Overview on parasite glycan-host immunity relationship

The hallmark of nematode and general helminth infections is the ability to alternatively activate antigen presenting cells (APCs) such as macrophages and dendritic cells (DCs) and also to induce Tregs and subsequently shift the immune response towards an anti-inflammatory Th2 (T-helper cell 2) type response. In contrast, classical bacterial infection, as an example, causes polarization of the immune system towards pro-inflammatory Th1 cell responses. This modulation of the immune system by helminths seems to be beneficial for both the parasite and the host since the absence of acute pro-inflammatory immune responses results in less damage to the host as well as allowing better survival of the parasite.
For a long time the exact molecular details of the Th2 immune response polarization by helminths were unknown. In 1999 Okano et al. showed that altering glycan structures on soluble schistosome egg antigens (SEA) with sodium metaperiodate abolished SEA’s ability to cause CD4+ Th2 responses in mice (Okano et al., 1999). Subsequently Tawill et al. showed that sodium periodate treatment also abolished the Th2 immune response in both parasitic Brugia malayi and non-parasitic C. elegans nematodes (Tawill et al., 2004). In this way two independent elegant studies revealed that helminth glycans must play a functional role in skewing host immune responses towards anti-inflammatory Th2. However the exact role of the glycans in the process was not revealed in these studies.

Later a substantial body of evidence, mainly through work on the parasitic trematode S. mansoni suggested critical roles of C-type lectin receptors (CLRs) and Toll-like receptors (TLR), collectively known as pattern recognition receptors (PRRs), in binding parasitic glycans. This subsequently leads to alternative activation of APCs such as DCs and macrophages and causes an anti-inflammatory Th2 immune response (Prasanphanich et al., 2013; Tundup et al., 2012; van Die and Cummings, 2010). The helminth’s strategy of survival by modulating host immune response via glycan-CLRs/TLRs interactions was coined ‘glycan gimmickry’ (van Die and Cummings, 2010).
1.2.8.2 Helminth glycan-host CLR interactions on dendritic cells

Dendritic cells are antigen presenting cells (APCs) which link the innate immune response to the adaptive and thus orchestrate the latter immune branch in the organism (Steinman, 2012). Upon phagocytosis of the pathogen, DCs migrate into lymph nodes where they present antigen to naïve T cells thus polarizing their development into specific subsets. Although other APCs such as macrophages also instruct T cells in lymph nodes, DCs are considered to be the most important in this process since they alone were able to induce Th2 differentiation in mice upon *S. mansoni* infection (MacDonald et al., 2001).

As well as Toll-like receptors (TLR) involved in pathogen recognition by DCs, previously mentioned CLRs are also crucial for DC activation and T cell subset expansion (van Die and Cummings, 2010). The types of CLRs present on DCs are Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin (DC-SIGN), macrophage galactose-type C-type lectin (MGL) as well as the mannose receptor (MR). All these CLRs...
have been shown to be able to interact with a wide range of glycan structures. DC-SIGN was shown to interact with Le^x, Le^y, LDNF and high mannose structures (Frison et al., 2003; Guo et al., 2004; van Liempt et al., 2006). The MGL receptor recognizes glycans bearing terminal α− or β- GalNAc residues such as LDN and the Tn antigen (Suzuki et al., 1996; Tsuiji et al., 2002; van Vliet et al., 2005). The MR is known to bind various mannose-containing glycans (Martinez-Pomares et al., 2001). The summary of lectins found in nematodes and other helminths and their suggested glycan binding partners are depicted in Figure 17.

1.2.8.3 Downstream events in dendritic cells upon helminth glycan-host CLR interactions

![Diagram showing two different signalling cascades happening in dendritic cells upon Lewis^x and high-mannose N-glycans engaging with the DC-SIGN receptor.](image)

Based on S. mansoni soluble egg antigen (SEA) studies, the glycans that bind to the C-type lectins DC-SIGN, MGL and MR are internalized by dendritic cells and targeted to MHC II^+ lysosomes with subsequent modulation of immune response towards Th2 type thus confirming the endocytotic nature of these C-type lectins (van Liempt et al., 2007).
The signalling cascades inside DC upon CLR activation have not been clearly defined, however emerging data points to the complexity of events happening inside the cell. For example, Gringhuis and colleagues showed that activation of DC-SIGN on DCs by mannose-containing glycans from *M. tuberculosis* or HIV-1 caused recruitment of effector proteins LARG and RhoA to the already existing, resting-state DC-SIGN signalosome which consists of scaffold proteins LSP1, KSR1 and CNK and Raf-1 (Gringhuis et al., 2009). The recruited effectors activated Raf-1 by phosphorylating it with subsequent p65 acetylation (Gringhuis et al., 2009). As a result, there was an up-regulation of interleukins Il-10, Il-12, Il-6 indicating a Th1 pro-inflammatory response (Gringhuis et al., 2009).

Surprisingly the same study showed that stimulation of DC-SIGN on DCs by the Le^{a} containing bacteria *H. pylori* which shares similar glycan structures with *S. mansoni* and is also known to be able to modulate immune response towards Th2 type (Bergman et al., 2004) results in opposite effects. The KSR-1-CNKRaf-1 protein triad dissociates from already existing, resting-state DC-SIGN signalosome causing up-regulation of IL-10, but down-regulation of IL-12 and Il-6 and blocking the pro-inflammatory Th1 response thus promoting Th2 cell expansion instead (Bergman et al., 2004).

In addition, glycolipids from adult worm of *S. mansoni* were shown to be able to induce pro-inflammatory Th1 response in human DCs and both TLR4 and DC-SIGN were needed as shown by blocking either receptor with monoclonal antibodies (van Stijn et al., 2010a). These results indicate that signalling through DC-SIGN in DCs does not necessarily polarize immune response to Th2 as well as suggesting cross-talk between TLR and CLR signalling pathways (van Stijn et al., 2010a). The downstream signalling pathways activated by glycan-lectin interactions in parasites, to the extent they are known, are shown in the Figure 18.
1.2.8.4 Helminth Glycan-host soluble galectin interactions in the immune system

Apart from CLRs present on the surface of APCs, helminth glycans are also able to interact with soluble extracellular lectins such as galectins. These receptors are able to recognize terminal galactose as well as terminal GalNAc. Galectin-3 has been demonstrated to be able to bind both neoglyconjugates carrying LDN as well as *S. mansoni* soluble egg antigens (van den Berg et al., 2004), species known to contain terminal GalNAc as well as galactose (Jang-Lee et al., 2007; Meevissen et al., 2010). In addition high levels of galectin-3 were localised in liver granulomas of *S. mansoni*-infected hamsters, together with localization of LDN and galectin-3 on parasite eggshells (van den Berg et al., 2004). To add to this study, galectin-3 was shown to be able to mediate the phagocytosis of LDN-glycans by macrophages (van den Berg et al., 2004). Another study showed that LDN antigens can cause Th2 immune response-associated formation of liver granulomas in mice with concomitant up-regulation of galectin-3 (Van de Vijver et al., 2006). The role of galectin-3 in Th2 immune response-mediated granuloma formation was supported by the knock-out of galectin-3 in mice which resulted in down-regulation of granuloma formation (Breuilh et al., 2007; Oliveira et al., 2007). All these data support the potential role of parasite glycan-host galectin interactions in modulating the host immune system towards a Th2 response which is known to happen upon infection of nematodes and other helminths.

1.2.8.5 Host immune system modulation by PC-containing glycans

N-glycans from many nematode species, namely *T. spiralis* (Morelle et al., 2000), *O. gibsoni* (Haslam et al., 1999), *O. volvulus* (Haslam et al., 1999), *A. suum* (Poltl et al., 2007), *A. viteae* (Haslam et al., 1997), and *C. elegans* (Haslam et al., 2002) have been shown to contain the phosphorylcholine (PC) moiety which is known to be able to modulate the host immune response (Helmby, 2009; Pineda et al., 2014). A lot of data about the effects of nematode-derived phosphorylcholine came from studies on the ES-62 protein from *A. viteae*, which was demonstrated to contain N-glycans bearing PC-modified terminal GlcNAc residues (Harnett et al., 2003). Later by attaching PC to ovalbumin (OVA) (Harnett et al., 2008) it was demonstrated that the majority of ES-62 immunomodulatory properties can be ascribed to the PC moiety. To my knowledge there are no reports of PC-containing glycans interacting with lectins. However, nematode-derived PC-containing ES-62 glycoprotein was demonstrated to interact with TLR-4 on bone-marrow derived mast cells (BMMC) causing degradation of
protein kinase C (PKC)-α needed for the mast cell activation (Melendez et al., 2007). ES-62 may be targeting other PKC isoforms such as PKC-ε and PKC-θ with the overall result being downregulation of pro-inflammatory cytokines IL-6 and TNF-α (Bell et al., 2015). Similar pathway was demonstrated to be present in other types of mast cells, namely peritoneal-derived mast cells (PDMC) and connective tissue-derived mast cells (CTMC) (Ball et al., 2013). In addition nematode-derived PC-containing ES-62 glycoprotein was demonstrated to interact with TLR-4 on APC as well as disrupting BCR and TCR signalling, resulting in mast cell desensitization and blockage of Th17 cells (Rzepecka et al., 2013). The ES-62-derived PC moiety was demonstrated to be able to mimic helminth infection in mice, causing Th2-shifted immune responses (Houston et al., 2000). Similarly in the chronic inflammatory arthritis (CIA) disease model, ES-62 was shown to supress pro-inflammatory Th1 responses (McInnes et al., 2003). Surprisingly in an airway inflammation disease model, ES-62 caused up-regulation of interferon-gamma (IFN-γ) production (Rzepecka et al., 2013), which is one of the Th1 response indicators, thus demonstrating that shifting of the immune response by PC towards Th1 or Th2 may depend on the inflammatory environment already present in the host. As a consequence of these studies small drug-like compounds based on PC moiety of ES-62 were developed and showed promising results in treating autoimmune diseases such as systemic lupus erythematosus (Rodgers et al., 2015), oxazolone-induced ear inflammation (Al-Riyami et al., 2015) and rheumatoid arthritis (Pineda et al., 2015).
1.3 Mass spectrometry as a tool for structural glycomics

1.3.1 Overview and historical background of mass spectrometry: from electrons to glycoproteins

Mass spectrometers are instruments which measure the mass/charge ratio ($m/z$ values) of gaseous ionised molecules. The common features of mass spectrometers are the ionisation chamber where a sample is ionised, the analyser where gaseous sample ions are separated from each other according to their $m/z$ and the detector where the separated ions are measured and quantitated. Nowadays mass spectrometry instruments are commonly found in many biology and biochemistry laboratories where they are applied in proteomics, lipidomics, metabolomics and as in my case glycomics research.

The first MS instruments appeared more in physics laboratories and it took years for the transition from physics applications to biological applications to occur. The instrument which was later considered as the first mass spectrometer was built around the beginning of the 20th century in the prestigious Cambridge’s Cavendish Laboratory led by J.J. Thomson (Griffiths, 2008). Subsequently he measured the mass-to-charge ration of the electron, $e/m$ for which in 1906 he was awarded the Nobel Prize. Later his work was extended by his student F. Aston who applied the first MS instrument built by J.J. Thomson to measure the masses of charged atoms. This led to the discovery that many elements such as neon and chlorine are made of different isotopes. For this work in 1922 Aston himself was awarded the Nobel Prize in Chemistry.

During World War II mass spectrometry was successfully applied to isolate and study the radioactive U-235 isotope for nuclear weapon development (Nier 1989). At the same time mass spectrometry started to find applications in other fields. Alfred Nier was one of the pioneers in bringing mass spectrometry to a wider scientific community. He performed studies such as the determination of the age of the Earth by the Pb-207/Pb-206 ratio for geologists and studies of C-13 isotope which was later extensively used for metabolic studies by biologists (Nier 1989).

The early ionisation method of the molecules used by J.J. Thompson was called electric discharge and in the post-war period was replaced by the Electron Impact Ionisation (later called the Electron Ionisation or EI) (McLafferty, 2011). The EI ionisation method (see section 1.3.7.2) was softer and did not induce as extensive fragmentation of parental ions.
Consequently it allowed the detection of molecular ions as well as fragment ions. This prompted the usage of mass spectrometry to analyse not only hydrocarbons but also for the first time small biomolecules found in natural sources (Biemann, 2002).

Scientists started to collect large volumes of spectra and utilize them for the characterisation of fragmentation reactions. As an example, one of the pioneers in this field Fred Mc Lafferty whilst at Dow Chemical Company collected about 4000 spectra before coming up with his famous ‘Mc Lafferty rearrangement’ for the β-cleavage of keto-group containing compounds (McLafferty 1959). Another pioneer Klaus Biemann at the Massachusetts Institute of Technology started to apply mass spectrometry to study alkaloids and to figure out their fragmentation patterns (Biemann, 2002). He also established the early rules for peptide fragmentation thus laying the foundation for current proteomics research (Biemann, 2014). In the United Kingdom the key player in developing Mass Spectrometry applications to biomolecules was Howard Morris who showed that mass spectrometry could be applied to problems such as the study of the Vitamin D metabolism with the consequent establishment of treatment for kidney failure patients (Lawson et al., 1971).

At the beginning of the proteomics era in late 1960-s two important inventions appeared in the mass spectrometry field. The first one was the chemical ionisation method which was a softer ionisation method than the Electron Ionisation (see section 1.3.7.2) and induced little fragmentation in the peptide ions. The second invention was gas chromatography-mass spectrometry (GC-MS) instruments (see section 1.3.7) which allowed the coupling of mass spectrometry with gas chromatography and allowed the separation of peptides online with the subsequent introduction into the MS instrument without the need to have a pure sample (Gohlke 1959; Ryhage 1964). Nevertheless the main obstacle at that time was the fact that by nature peptides were difficult to ionise hence there was a need for their derivatization. Klaus Biemann developed the poly-amino alcohol derivatization approach for peptides using GC-MS instruments (Nau et al., 1973). In a competitive strategy Howard Morris adopted the sugar permethylation reaction initially discovered by Hakomori (Hakomori, 1964) and together with the N-acetylation of the N-termini applied it to create N-acetylated permethyl peptide derivatives (Morris et al., 1971). This work for the first time demonstrated the application of mass spectrometry to analyse mixtures of peptides as at that time only the analysis of single peptides was considered feasible (Morris et al., 1971).
In the late 1970s Fast Atom Bombardment (FAB) (see section 1.3.4) was invented by Michael Barber at the University of Manchester Institute of Science and Technology (Barber, Bordoli et al. 1981) and quickly revolutionised the field of MS-based biopolymer analysis. It did not require the tedious sample derivatization and allowed the analysis of native oligosaccharides and peptides (Morris et al., 1981). Although the technique was limited to the analysis of molecules of a few thousand Daltons (Barber, Bordoli et al. 1981), at that time it was a remarkable achievement. Later the FAB ionisation method (see section 1.3.4) got replaced by softer ionisation techniques such as electrospray ionisation (ESI) (Fenn et al., 1989) and matrix-assisted laser desorption ionisation (MALDI) (Karas and Hillenkamp, 1988). These two soft ionisation methods allowed the analysis of much bigger biomolecules. MALDI (see section 1.3.5) has been extensively used for glycan analysis during my PhD studies.

1.3.2 Tandem mass spectrometry

One of the later important advancements in the field of mass spectrometry was the creation of tandem, or MS/MS, mass spectrometry in the period 1971-1973 by J.H. Beynon (Beynon, Cooks et al. 1973). As the name suggests, this type of experiment involves two stages of analysis. In the first stage molecular ions from the sample are produced and analysed. In the second individual ions with specific m/z value from the first experiment are selected, fragmented and further analysed. The first tandem mass spectrometer had two sector analysers (Beynon, Cooks et al. 1973) and this type of approach is called tandem in space as two separate analysers are used. Later other tandem in space instruments were developed such as TOF-TOF. One way to achieve fragmentation of the selected molecular ions is CID (collision-induced dissociation). In the CID experiment a stream of neutral gas such as argon is collided with the selected ions inside a collision cell. In this way the internal energies of the selected molecular ions are increased causing them to fragment. An alternative to the tandem in space approach is tandem in time instrumentation where only one analyser, usually an ion trap, is used to perform different stages of MS analysis at different times. The ionised molecules from the sample are trapped inside the trap with the desired ion of a certain m/z value being kept in the trap whilst other ions are released. Then CID is performed and resulting fragmented ions are expelled into an analyser. If needed, fragmentation of a fragment ion can be performed allowing MS^n experiments to be performed. Throughout my
thesis I have used tandem in space MALDI-TOF-TOF (see section 1.3.5) as well as tandem in time MALDI-QIT-TOF (see section 1.3.6) instruments.

1.3.3 Other trends in modern mass-spectrometry

The invention of soft ionisation techniques, namely MALDI (see section 1.3.5.2) (Karas, Bachmann et al. 1985) and ESI (Fenn et al., 1989) revolutionized the area of mass spectrometry allowing large biomolecules to be analysed. In the 21st century one of the major trends in mass spectrometry advancements is the rise of hybrid instruments where different analysers are combined together. One such an instrument having Q-TOF configuration was conceived at Imperial College by Howard Morris and built by Micromass Ltd (Morris, Dell et al. 1996). Another example of a hybrid instrument is the MALDI-QIT-TOF (see section 1.3.6) developed by Shimadzu and this instrument has been used throughout my PhD studies (Ding, Kawatoch et al. 1999).

1.3.4 Fast Atom Bombardment mass spectrometry as an early technique for glycan analysis

In 1980 at the Symposium on Soft Ionization Biological Mass Spectrometry held at Imperial College London the new ionisation technique later coined Fast Atom Bombardment (FAB) and invented by Barber at University of Manchester Institute of Science and Technology was introduced to the scientific community (Barber, Bordoli et al. 1981). This ionisation method involves sample deposition on a solid plate often dissolved in a matrix such as glycerol (Dell, 1987). An accelerated ion or atom beam is directed at the sample and the sample molecules are ionised via ion sputtering (Barber, Bordoli et al. 1981 as seen in the Figure 19.
This type of ionisation was used for older days glycomics analysis where the accelerated high-energy beam of neutral atoms or ions from the laser is focused on the sample and matrix mix deposited on the probe making sample to ionise. FAB brought big excitement into the scientific community and quickly revolutionised the field of the mass spectrometry-based biopolymer analysis. As compared to earlier used MS methods it offered a lot of advantages with the most obvious one being the ability to ionise large underivatized macromolecules (Barber, Bordoli et al. 1981). In addition the ionisation process generated both molecular and fragment ions thus providing a high level of structural information about the analyte molecule (Barber, Bordoli et al. 1981). Although nowadays spectra of more than ten thousands Da molecular weight molecules are obtained using different technologies (Angelini et al., 2010) in the early eighties FAB ionisation coupled with magnetic sector analyser rendered it possible to obtain the spectra of molecules of a few thousand Da molecular weight, which was a great achievement at that time (Barber, Bordoli et al. 1981).
As the scientific community quickly adopted the new methodology for glycoconjugates, FAB spectra were successfully generated from much larger oligosaccharides. Notably the structure of the sialylated fucosyl lactosaminoglycan from human neutrophil granulocytes was proposed (Fukuda et al., 1984). Although the molecular ion peak of the glycan was not observed due to the \( m/z \) limitations of the instrument, the FAB-generated fragment ion data suggested tetra-antennary N-glycans with poly-lactosaminyl side chains, some of which were decorated by terminal sialic acid and fucose \( \alpha 1-3 \) linked to GlcNAc residues (Fukuda et al., 1984). FAB-MS instruments were also used to characterize the glycomes of \textit{H. contortus} (Haslam et al., 1996; Haslam et al., 1998) and \textit{D. viviparus} (Haslam et al., 2000), the parasites of my PhD research interest.

1.3.5 AB 4800 MALDI-TOF-TOF instrument as a replacement for the FAB-MS sector mass spectrometer used in glycomics studies

1.3.5.1 Overview

Over the past decade or so MALDI-based instruments have replaced FAB-MS instruments and have established themselves as an indispensable tool for quick and sensitive glycan profiling. The main advantages of MALDI are a soft ionization leaving large biomolecules intact, a relative tolerance to salt and other impurities and a pulsed nature of ionization with small consumptions of the sample during the analysis. Combining MALDI with TOF-TOF analysers, as seen in the Applied Biosystems AB 4800 MALDI-TOF-TOF instrument, allows scientists to perform not only MS profiling of glycans but also their MS/MS fragmentation. In contrast to FAB-MS sector instruments, where generated spectra had a mixture of both molecular and fragment ions, MALDI-TOF-TOF instruments allows the selection of specific molecular ions for further MS/MS fragmentation. With the advent of improved MALDI-based techniques the FAB-MS data of the glycans for \textit{H. contortus} and \textit{D. viviparus} has been revisited.

1.3.5.2 Matrix Assisted Laser Desorption Ionisation

The projects carried out during my PhD studies mainly utilised matrix assisted laser desorption ionisation (MALDI) instruments. To compare to older FAB-mass sector machines these are more modern and powerful type of instrumentation were a MALDI ionisation
source (see Figure 20) is coupled to the time-of-flight analyser in a MALDI-TOF MS and the related tandem MALDI-TOF/TOF MS instruments. A laser pulse is applied to the sample, which is dissolved in the low-molecular weight matrix, capable of absorbing ultraviolet light at the same wavelength as the incident laser. This results in a softer ionisation method, thus rendering MALDI a more sensitive type of ionisation with reduced chemical noise background compared to the older FAB technology, though at a cost of losing the majority of fragment ions. Nevertheless some ions can undergo fragmentation inside the MALDI source and after leaving it in a process called post-source decay.

The discovery of MALDI can be traced back to 1985 when Michael Karas and Franz Hillenkamp realised that when a mixture of alanine and tryptophan is analysed by Laser Desorption Mass Spectrometry the signal for alanine is much stronger than when it is ionised alone (Karas, Bachmann et al. 1985). In other words, according to Hillenkamp ‘aliphatic amino acid was riding piggyback on tryptophan’ (Griffiths, 2008) and ‘tryptophan thus must be regarded as an absorbing matrix resulting in molecular ion formation of the nonabsorbing alanine’ and the process was coined ‘Matrix-Assisted Laser Desorption’ (Karas, Bachmann et al. 1985). In addition part of the Nobel Prize for Chemistry in 2002 ‘for their development of soft desorption ionisation methods of mass spectrometric analyses of biological molecules’ was awarded to Koichi Tanaka. He showed that the Laser Ionisation TOF mass spectrometer could in principle be used to analyse the biopolymers like proteins up to \( m/z \) 100000 when they are deposited on a solid surface using suspension of the fine metal powder and glycerol (Tanaka, Waki et al. 1988) although his method did not ionise molecules as well as MALDI invented by Karas and Hillenkamp.

As well as proteomics MALDI quickly found application in the field of the glycomics. The superiority of MALDI over older FAB-type instrumentation was shown by Babu and colleagues (Babu et al., 2009a) when analysing neutrophil N- and O-glycans using MALDI-TOF and MALDI-TOF-TOF techniques. They managed to show the presence of glycans up to \( m/z \) 6500 whereas the previous glycomics data from neutrophil granulocytes using FAB-MS instrument (Fukuda et al., 1984) showed the highest molecular weight species with \( m/z \) 1897. Furthermore using MALDI techniques Haslam and colleagues (North et al., 2010a) reported poly-LacNAc-containing N-glycans with up to 13000 \( m/z \) previously unseen in Chinese Hamster Ovary cell lines (North et al., 2010a) and mouse lung tissue (Bern et al., 2013) clearly demonstrating the ultra-sensitivity of the MALDI instrument at high molecular weight.
The concept of the Time-of-flight (TOF) analyser was introduced for the first time by Stephens in 1946 (Wolff, Stephens 1953). In 1955 Wiley and McLaren published the design of the linear TOF (Wiley, McLaren 1955) which later became commercialised. In the 1980s interest in TOF geometry instruments was renewed since they fitted very well with laser desorption ionisation and the creation of MALDI-TOF instruments opened up new research horizons in biopolymer mass spectrometry.
1.3.5.4 Linear TOF analyser

In the linear TOF analyser the ions expelled from the ionisation source are accelerated by the same voltage potential resulting in ions having equal potential energy which depends on applied voltage and charge of the particle. When ions are introduced into the analyser the potential energy is converted into kinetic energy which is proportional to the mass and the velocity squared of the particle. It is important to note that since the voltage applied to the particles is the same, the kinetic energy is also the same for different particles. From there it follows that particles with different mass to charge ratio will have different speeds in the analyser and the heavier and less charged the particle is, the longer it takes to reach the detector.

1.3.5.5 Reflectron TOF and delayed pulsed extraction

The linear TOF geometry offers very high upper mass limits as well as sensitivity. However one of its drawbacks is poor mass resolution. The reflectron TOF and delayed pulsed extraction were created to overcome these issues.

The reflectron TOF concept was suggested by Mamyrin in 1973 (Mamyrin, Karataev et al. 1972) and consists of a series of grids and ring electrodes which serve as retarding ion mirrors by acting electrostatically on the ions in the TOF analyser. They not only increase the path of flight of an ion, but also correct the dispersion of the kinetic energy of the particles by letting ions with slightly higher kinetic energy but the same \( m/z \) value to penetrate deeper into the reflector and catch up with the ions having the same \( m/z \) value but slightly smaller kinetic energy.

Another improvement of the linear TOF analyser was created by Wiley and McLaren, and called delayed pulsed extraction (Wiley, McLaren 1955). The idea consists of letting ions formed in the ionisation source enter the field-free region only after a nano- to microsecond delay. This is achieved by the application of a short pulse of voltage in contrast to the continuous ion extraction where ions formed in the ionisation source are extracted immediately into the analyser by means of the applied voltage. This allows the correction of the kinetic energy dispersion for ions with the same \( m/z \) values leading to improved resolution of the spectrum.
1.3.6 MALDI-QIT-TOF instrument

1.3.6.1 Overview of the MALDI-QIT-TOF instrument

The Shimadzu KRATOS AXIMA-QIT is unique in the market of instruments in that it combines a MALDI source with a quadrupole ion trap (QIT) (see Figure 21) as opposed to the more conventional electrospray ionisation (ESI)-QIT combination. Since during the ionisation process high molecular weight molecules generate ions with high kinetic energy, early attempts to combine MALDI ion source with QIT were not particularly successful due to the difficulty of introducing and keeping high m/z molecular ions inside the trap (He and Lubman, 1997) (Kaiser, Cooks 1991). The Shimadzu scientists improved the method by introducing the near orthogonal laser irradiation of the sample, by keeping the radio frequency (RF) voltage field switched off when ions enter the trap and subsequently switching it on once the ions are inside and by applying the retarding field in the QIT to reduce the kinetic energy of an ion inside the trap. With all these improvements they managed to create a sensitive MALDI-QIT-TOF instrument which has high trapping efficiency of high molecular weight ions (Ding, Kawatoch et al. 1999). The soft MALDI ionisation method is extremely suited for analysing glycans and the presence of the QIT, in contrast to the other MALDI instrument used during my PhD studies, the AB 4800 MALDI-TOF-TOF, offers the possibility to perform MS^n analysis revealing extra details of glycan structure and linkage between monosaccharides.

![Figure 21](image_url) The schematic representation of the MALDI AXIMA QIT instrument.
For principles of MALDI and TOF operation, see sections 1.3.5.2 and 1.3.5.3.

1.3.6.2 Principles of the QIT operation

Initially invented by Paul and Steinwedel (Paul, Steinwedel 1960) in 1960 and modified into a useful and commercially available mass spectrometer by Stafford and colleagues in 1980s (Stafford, Kelley 1984), the quadrupole ion trap (QIT) shares similarities with the quadrupolar analyser (see section 1.3.7.3). Quadrupole analysers use voltages in two dimensions to change the trajectory of the ions whereas the QITs have voltages in all three dimensions (Paul, Steinwedel 1960). The 3D voltage environment is achieved by applying the alternating current (AC) to end cap electrodes and direct current (DC) to a ring electrode. The essential difference between the QIT and quadrupole analyser is that in a quadrupole analyser ions with a stable trajectory will reach the detector whereas in the mass-selective instability mode operating QIT ions with an unstable trajectory are trapped inside (Stafford, Kelley 1984). Later by scanning the radio frequency (RF) voltage ions are emitted from the trap into another analyser such as TOF or the detector.

1.3.7 GC-MS instrument

1.3.7.1 Overview

The Perkin Elmer Clarus 500 and Bruker SCION SQ 456-GC are the GC-MS (gas chromatography- mass spectrometry) instruments used throughout my PhD studies and comprise a gas-chromatography column linked to an Electron Ionisation -Quadrupole analyser (EI-Q) mass spectrometer. The main use of these instruments throughout my studies was to perform linkage analyses of monosaccharides derived from permethylated glycans that were hydrolysed, reduced to generate free hydroxyl groups that are labelled with acetyl groups. The resultant partially methylated alditol acetates of different monosaccharides in most cases have different retention times on the GC column. In addition simultaneously acquired mass spectra provide the characteristic fragment ions allowing the confirmation of the linkage of the monosaccharide. Since mass spectrometry alone cannot distinguish different monosaccharides isomers of the same mass, such as mannose, glucose or galactose,
the GC-MS instrument is a useful complementary tool to prove the presence of the specifically linked monosaccharides in the sample.

1.3.7.2 Electron Ionisation

The Electron Ionisation source was invented by Dempster in 1917 (Dempster 1918) and improved by Bleakney (Bleakney 1929) and Nier (Nier, 1947). The principle of this ionisation technique is depicted in the Figure 22. The heated filament gives off electrons which are accelerated towards the anode trap. The gaseous sample molecules are introduced into the ionisation source where they interact with the accelerated electrons.

A consequence of the electron ionisation process is that there remains a lot of excess energy which results in extensive fragmentation of the analyte molecule and thus is considered a ‘harder’ ionisation technique. Although molecular ions are often not seen in the spectrum the fragment ions generated by extensive fragmentation are very useful in the structural elucidation of the analyte.

Figure 22 The schematic diagram of the Electron Ionisation.
1.3.7.3 Quadrupolar Analyzer

The quadrupolar analyser was devised by Paul and Steinweden at Bonn University in 1953 (Paul, Steinweden 1953). It consists of four parallel rods to which alternating positive and negative potentials are applied. The forces induced by oscillating electric fields affect the trajectory of ions which are travelling between the rods. The ions which hit the rods are discharged and hence do not reach the detector. The ions with a certain m/z value have a stable trajectory. By continuously varying the applied voltage, the range of m/z values can be scanned allowing ions with a certain m/z value only to reach the detector. The schematic diagram of the quadrupolar analyser is shown in the Figure 23 below.

Figure 23 The schematic diagram of the quadrupole analyser

1.3.8 Derivatization of glycans

Glycans being relatively polar molecules pose a greater challenge for ionization compared to less polar peptides. The decreased ionization efficiency results in a more difficult analysis of this class of biomolecules especially when dealing with minor components of a complex heterogeneous mixture. To overcome this problem the permethylation of hydroxyl and amido groups in glycan molecules was introduced which not only increases the ionisation efficiency but also directs the fragmentation of the glycan molecular ions in a less random, more predicted fashion.
Initial attempts to perform the permethylation reaction were carried out by Haworth (Haworth 1915) and Purdie (Purdie, Irvine 1903). The first method employed sodium or potassium hydroxide acting both as a base and a solvent with dimethyl sulphate acting as a methyl donor (Haworth 1915). The second method employed methyl iodide as a solvent and a methylation agent together with silver oxide acting as a base (Purdie, Irvine 1903). The major limitation of these methods was the need to repeat the experiment many times to achieve complete methylation with reaction times extending over a month or longer (Haworth 1915; Purdie, Irvine 1903).

In 1964 Hakomori proposed the alternative permethylation method which involved DMSO acting as a solvent, methyl iodide as a methylating agent and dimsyl anion as a base (Hakomori, 1964). Although it did not offer high yields this reaction could be completed within few days rather than months. Later Ciucanu and Kerek realised that hydroxide ions could remove protons more efficiently from the glycan hydroxyl groups than dimsyl anions. In 1984 they published a modified permethylation protocol where, in contrast to the Hakomori method, dimsyl anions are replaced with hydroxide ions (Ciucanu, Kerek 1984), see Figure 24. This method offers excellent permethylation yields, can be achieved within minutes and was used to permethylate glycans in my experiments.

![Figure 24](image)

**Figure 24** The mechanism of the permethylation reaction
The mechanism is proposed by Ciucanu and Kerek (Ciucanu, Kerek 1984). Hydroxide anion acts as a base by abstracting the proton from the sugar hydroxyl group. The resultant sugar alkoxide anion denoted as ‘R-O ‘ subsequently reacts with the alkylating reagent methyl iodide.

1.3.9 Overview of the strategy for interpreting MS spectra

In order to assign the composition of molecular ions we need to consider the masses of permethylated sugar monomers listed in Figure 25 (a), the masses of reducing and non-
reducing ends of permethylated structures (Figure 25 (b)) and the mass of the sodium ion used to ionise the molecule. The general assignment formula can be expressed as: \( X = \Sigma (M) + 46 + 23 \) where \( X \) is the \( m/z \) value of the peak, \( \Sigma (M) \) corresponds to the sum of permethylated residue masses present in the glycan, 46 is the permethylated mass of the reducing and non-reducing ends and 23 is the mass of a sodium ion.

As MS analyses give us only the composition of the glycans, knowledge of biosynthetic pathways is typically used as a powerful method for determining potential structures. To get more structural information, MS/MS analysis of peaks of interest is performed, whereby molecular ions of the interest are selected, broken down into smaller fragments and their masses measured.

![Figure 25](https://example.com/figure25.png)

**Figure 25** Symbolic nomenclature and permethylated masses of monosaccharides


b) The mass of a permethylated sugar is calculated by adding the sum of the increment (or residue) masses of the sugars (a) to the sum of the masses of the reducing and non-reducing ends. Figure is adopted from (North et al., 2010b) and is the courtesy of Dr. Simon North.
1.3.10 Overview of glycan fragmentation pathways

An overview of the glycan fragmentations happening upon MS/MS and MS^n experiments is outlined in Figure 26 and the fragments are named according to the nomenclature proposed by Domon and Costello (Domon, Costello 1988).

Figure 26 The glycan fragmentation pathways
The nomenclature is proposed by Domon and Costello (Domon, Costello 1988).

One of the common classes of fragments observed during tandem mass spectrometry are ‘C-type’ ions (Domon, Costello 1988). They form when the glycosidic bond oxygen abstracts the proton from one of the carbons forming a 4-membered intermediate and the charge remains on the non-reducing end fragment (see Figure 27). The analogous fragmentation pathway happens upon the ‘Y-type’ formation where the glycosidic bond oxygen abstracts the proton from the non-reducing end sugar, carbon two. If the glycosidic bond is cleaved on the reducing-end side with the charge staying on the same side, ‘Z-type’ ions are formed. They are favoured when the reducing-end monosaccharide is a 3-linked HexNAc since the resultant double bond between carbons two and three is in resonance with the nitrogen lone pair and the acetyl group (see Figure 27). These fragments are extremely useful in identifying 3-linked GlcNAc and GalNAc residues. ‘B-type’ ions are formed based on a
similar principle where the double bond between carbon one and two is resonating with the nitrogen lone pair and the acetyl group. This information allows us to deduce that the resultant ‘B-type’ ion most likely has a reducing-end GlcNAc or GalNAc.

Cross-ring cleavages (see Figure 28) usually require more energy than glycosidic bond cleavages and are typically observed when fragmenting small molecular weight ions such as in MS<sup>n</sup> experiments. The example of the cross-ring fragment formation mechanism is depicted in Figure 28 where reverse Diels-Alder fragmentation is taking place in the already formed ‘B-type’ ion (Ashline et al., 2005). The presence of the double bond between carbon one and two in the parental ion was proposed to favour the reverse Diels-Alder fragmentation pathway (Ashline et al., 2005) although in its absence analogous cross-ring fragmentations can also happen (Domon, Costello 1988). In MS<sup>n</sup> experiments the cross-ring fragments can be very informative in determining linkage information between monosaccharides.

**Figure 27** The proposed mechanisms for the formation of glycan ‘C-type’ and ‘Z-type’ ions. The nomenclature is proposed by Domon and Costello (Domon, Costello 1988).
Figure 28: The proposed reverse Diels-Alder cross-ring fragmentation mechanism

The formation of the examplar \textsuperscript{3,5}A-type and other ions was described by (Ashline et al., 2005).
1.4 Aims of the thesis

The aims of my thesis are to employ ultra-sensitive MALDI-TOF-TOF and MALDI-QIT-TOF-based techniques as well as GC-MS to structurally characterize *H. contortus* and *D. viviparus* glycomes. The specific aims are:

1. To revisit adult and L3 stage *H. contortus* N- glycans which have been investigated earlier with FAB-MS instrumentation.
2. For the first time to analyse adult and L3 stage *H. contortus* O- glycans.
3. To revisit adult *D. viviparus* N- glycans previously characterized using FAB-MS instrumentation.
4. For the first time to analyse adult *D. viviparus* O- glycans as well as L3 stage N- and O- glycans.
Chapter 2

Materials and Methods
2.1 Materials

2.1.1 Biological samples

Adult and L3 stage _Haemonchus contortus_ as well as adult and L3 stage _Dictyocaulus viviparus_ nematodes were kindly provided by Dr. G. Coles, University of Bristol, School of Veterinary Sciences. The _D. viviparus_ and _H. contortus_ L3 stage larvae were collected by Baermann apparatus, purified by sucrose gradient and thoroughly washed with tap water. The adult parasitic worms were hand picked from humanely slaughtered animals ensuring that worms are free of any remaining abomasal or lung debris. The picked worms were washed with Earle’s balanced salt solution.

2.1.2 General chemicals and reagents

Iodoacetic acid (IAA), sodium acetate, potassium borohydride (KBH₄), 2-amino-2-hydroxymethyl-propane-1,3-diol (TRIS), α-cyano-4-hydroxycinnamic acid, 2,5-dihydroxybenzoic acid (DHB) were purchased from Sigma-Aldrich (Poole, UK). Acetonitrile, ammonia solution, methanol, chloroform, dimethylsulphoxide (DMSO), propan-1-ol, ammonium hydrogen bicarbonate (AMBIC), sodium hydroxide, acetic acid, trifluoroacetic acid (TFA) were purchased from Romil (Cambridge, UK). Dithiothreitol (DTT) and 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate (CHAPS) were acquired from Roche Applied Science (East Sussex, UK). Methyl iodide was purchased from Alfa Aesar (Lancaster, UK). Ethylenediaminetetraacetic acid (EDTA) and formic acid were obtained from Fluka (Poole, UK). 18 MΩ·cm³ distilled (MiliQ®) water was produced by the Purite Neptune water purification system from Purite Ltd (Oxfordshire, UK).

2.1.3 Enzymes

Trypsin, Jack bean α-mannosidase and β1-4 specific galactosidase from _Streptococcus pneumonia_ were obtained from Sigma-Aldrich (Poole, UK). Recombinant β1-4 specific galactosidase was kindly provided by Prof. Iain Wilson, University of Natural resources and Life Sciences, Vienna. Recombinant PNGase F and PNGase A from sweet almond meal were purchased from Roche Applied Sciences (East Sussex, UK.). Green coffee bean α(1-3,4,6)-
galactosidase, bovine testes β(1-3,4,6) galactosidase, recombinant β(1-3,6) galactosidase (from *Xanthomonas manihotis*, expressed in *E. coli*) were purchased from ProZyme, CA, USA.

2.1.4 Equipment and consumables

Oasis® HLB and Sep-Pac C18 Classic cartridges were purchased from Waters Ltd (Hertfordshire, UK). Thermo Savant ModuloD freeze dryer, SnakeSkin® Pleated Dialysis Tubing, 70000 MWCO, 22mm dry diameter and Thermo Savant SPD121P Speed Vac concentrator were obtained from Thermo Scientific (Basingstoke, UK). Lo-bind® Eppendorfs tubes were purchased from Sigma-Aldrich (Poole, UK). Multi-tube vortexer, Vibra-Cell™ homogenizer, Falcon tubes, glass culture tubes (13x100 mm Corning, 7.5 ml), screw caps (disposable phenolic GPI 13-415 threaded with PTFE liner) and glass Pasteur pipettes were from VWR International Ltd (Lutterworth, UK.)
2.2 Methods

2.2.1 Overview

The methods of N- and O-glycan preparations and analyses by MALDI instrumentation are based on previous report (North et al., 2010b) and are summarized in the Figure 29. Briefly worms were homogenized and extracted (glyco)proteins were digested into smaller (glyco)peptides by trypsin. PNGase-F enzyme was then used to remove N-glycans from glycopeptides. The remaining (glyco)peptides were then subject to PNGase A treatment. This enzyme, in contrast to PNGase F, is capable of releasing N-glycans carrying α1-3-linked fucose on the proximal GlcNAc. This assists with establishing the position of fucoses present. The O-glycans are then released from remaining (glyco)peptides by reductive elimination. In order to obtain more structural information the released N- and/or O-glycans were subject to exoglycosidase and/or HF treatment. Finally the free glycans were permethylated and analysed by MALDI instruments. Compositional glycan profiling was obtained by MALDI-TOF MS instruments. Additional levels of structural information were obtained by MALDI-
TOF-TOF MS/MS and MALDI-QIT-TOF MS<sup>n</sup> instrumentations. Finally the permethylated glycans were subject to GC-MS linkage analyses which allows us to establish the monosaccharide configurations and linkages.

2.2.2 Worm homogenization

To clean the homogenizer the metal tip was placed into a beaker with 33% water, 33% methanol, 33% formic acid (v/v/v) solution and kept in the sonicator for 10 minutes. Then the metal tip was put back onto the homogenizer and cleaning continued by turning the homogenizer on for 1 minute in each of the following solutions: 1) 33% water, 33% methanol, 33% formic acid solution 2) methanol 3) 50% methanol, 50% chloroform solution 4) methanol 5) ultrapure water. Finally the homogenizer was kept on for one minute in an empty Falcon tube to get rid of all the remaining droplets.

For sample homogenization on average 0.2g of wet of material (L3 or adult <i>H. contortus/ D. viviparus</i> worms) was homogenized by adding 5 ml of ice cold homogenisation buffer (25 mM TRIS, 150 mM NaCl, 5 mM EDTA, and 1% CHAPS, pH 7.4 adjusted with dilute acetic acid) to the sample and by activating the homogenizer for 10 seconds three times with 15 seconds intervals on ice in between. Subsequently the homogenate was transferred to the SnakeSkin® pleated dialysis tubing, sealed with clips and put into the beaker with 4.5 litres of dialysis buffer (50 mM ammonium bicarbonate, pH 7.6) and kept for 48h at 4 degrees. Over that time the dialysis buffer was changed eight-ten times.

2.2.3 Reduction and carboxymethylation

A TRIS-acetate buffer, 0.6 M pH 8.5 was degassed with a stream of nitrogen for 30 minutes. To the dialysed and lyophilized samples 1 ml of degassed 2 mg/ml DTT in TRIS buffer solution was added. The reaction was incubated for 60 min at 37 °C. The protection of disulphide bridges was performed by adding 1ml 12 µg/ml iodoacetic acid (IAA) in Tris buffer and incubating for 90min at room temperature in the dark. Then the reduced and carboxymethylated homogenate was transferred into the SnakeSkin® pleated dialysis tubing, sealed with clips and put into the beaker with 4.5 litres of dialysis buffer (50 mM ammonium
bicarbonate, pH 7.6) and kept for 24h at 4 °C. Over this period the dialysis buffer was changed five-six times.

2.2.4 Trypsin digestion and subsequent (glyco)peptide purification

To the dialysed and lyophilized sample 300 µl of 1mg/ml trypsin in 50 mM AMBIC buffer, pH 8.4 solution was added. The reaction was incubated at 37 °C for 14h. The trypsin was deactivated by adding one drop of acetic acid to the solution. Subsequently the purification of (glyco)peptides using C18 chromatography was performed. An Oasis® HLB Plus cartridge was attached to a 20 ml plastic syringe and conditioned successively with 5 ml of methanol, 5 ml of 5% acetic acid, 5 ml of propan-1-ol, 15 ml of 5% acetic acid. Then the sample was diluted in 5% acetic acid and loaded onto the cartridge and washed with 20 ml of 5% acetic acid. The bound (glyco)peptides were gradually eluted from the C18 column by passing successively 4 ml of 20% propan-1-ol in 5% acetic acid, 4 ml of 40% propan-1-ol in 5% acetic acid and 4 ml of 100% propan-1-ol. All three propanol-based eluent fractions were collected and their volume was reduced in a SpeedVac®. Finally three fractions were combined and lyophilized.

2.2.5 PNGase F digestion

The lyophilised sample was dissolved in 150 µl of 50 mM, pH 8.4 AMBIC buffer. 5 U (5 µl) of PNGase F enzyme solution were added and the reaction was carried out for 24 hours at 37 °C. About the half-way though the reaction course another 5 U (5 µl) of PNGase F enzyme solution was added. Then the solution was frozen and lyophilized.

2.2.6 Sep-Pak separation of PNGase-F released N-glycans from remaining (glyco)peptides

A Sep-Pak C18 Classic cartridge was attached to the glass syringe and washed successively with 5 ml of methanol, 5 ml of 5% acetic acid, 5 ml of propan-1-ol and 15 ml of 5% acetic acid. The sample was dissolved in 200 µl of 5% acetic acid and loaded onto the cartridge. Then 5ml of 5% acetic acid, 4 ml of 20% propanol-1-ol in 5% acetic acid, 4 ml of 40% propanol-1-ol in 5% acetic acid, 4 ml of 100% propanol-1-ol were successively passed through the cartridge and all the fractions were collected and the volumes reduced in the
SpeedVac®. The N-glycans-containing 5% acetic acid fraction was then lyophilised and subject to exo-glycosidase treatment(s) (see section 2.2.10), HF treatment (see section 2.2.9) or permethylation reaction (see section 2.2.11). The (glyco)peptides-containing 20%, 40% and 100% propanol fractions were combined, had their volumes reduced and were lyophilized prior to PNGase A digestion.

2.2.7 PNGase A digestion and subsequent N-glycan separation from remaining (glyco)peptides

The combined and lyophilised 20%, 40% and 100% propanol fractions after PNGase F digestion (see section 2.2.6) were dissolved in 150 µl of 50 mM ammonium acetate buffer, pH 5.0. Then 5 mU (10 µl) of PNGase A enzyme solution were added and the reaction was carried out for 24 hours at 37 ºC. About half-way though the reaction course another 5 mU (10 µl) of PNGase A enzyme solution was added. Then the reaction was frozen and lyophilized. The PNGase A- released N-glycans were separated from remaining (glyco)peptides as described in section 2.2.6. The N-glycans-containing 5% acetic acid fraction was then lyophilised and subject to exo-glycosidase treatment(s) (see section 2.2.10) or permethylation reaction (see section 2.2.11). The (glyco)peptides-containing 20%, 40% and 100% propanol fractions were combined, had their volumes reduced and were lyophilized prior to reductive elimination as described further in section 2.2.8.

2.2.8 Reductive elimination of O-glycans and their subsequent purification

The combined and lyophilised 20%, 40% and 100% propanol fractions after PNGase A digestion (see section 2.2.7) were subject to reductive elimination by adding 400 µl of 1 M KBH₄ in 0.1 M KOH solution and incubating for 16h at 45 ºC. The reaction was stopped by adding 5 drops of acetic acid until fizzing stopped. Then the sample was loaded onto a desalting column loaded with Dowex 50W-X8(H) beads which was pre-washed by passing 15 ml of 5% acetic acid. After the sample was loaded, the column was eluted with 5 ml of 5% acetic acid. The desalted sample was collected and the volume reduced in the SpeedVac® and lyophilized. Excess of borates were removed by co-evaporation with 10% (v/v) acetic acid in methanol under a stream of nitrogen at room temperature.
2.2.9 HF treatment of PNGase F released N-glycans

PNGase F released N-glycans were lyophilised in a plastic Eppendorf tube and subsequently 50 µl of 48% hydrofluoric acid was added using a plastic micropipette tip. The sample was incubated on ice at 4 °C for 48h. The reaction was stopped and the sample dried under a gentle stream of nitrogen.

2.2.10 Exoglycosidase(s) treatment of N- and O-glycans

Free glycans were dissolved in 200 µl of 50 mM sodium acetate buffer, pH 5 and 1.5 U of Jack bean α-mannosidase from *Canavalia ensiformis* (Sigma, M7257, EC 3.2.1.24) were added to the sample over a total incubation time of 48h at 37 °C.

For β1-4 specific galactosidase digestion 30 mU of *Streptococcus pneumoniae* enzyme (Sigma-Aldrich, G0413) or 32 µg of recombinant β1-4 specific galactosidase kindly provided by Iain Wilson, University of Natural resources and Life Sciences, Vienna were added to the sample and incubated for 30h at 37 °C in 50 mM ammonium acetate buffer, pH 5.75.

For α-galactosidase digestion total of 5 U of Glyko® α(1-3,4,6)-galactosidase from green coffee bean (Prozyme, GKX-5007) were added to the sample over 48h and incubated at 37 °C in 50 mM ammonium acetate buffer, pH 6.0.

For broad specificity β- galactosidase Digestion- total of 0.5 U of Glyko® β(1-3,4)-galactosidase from bovine testes (Prozyme, GKX-5013) were added to the sample over 48h and incubated at 37 °C in 50 mM ammonium acetate buffer, pH 4.6.

For β1-3, 6 galactosidase digestion total of 0.2 U of of Glyko® β(1-3,6)-galactosidase (recombinant from *Xanthomonas manihotis*, expressed in E.coli.; Prozyme, GK80120) were added to the sample over 48h and incubated at 37 °C in 50 mM ammonium acetate buffer, pH 5.0.

Finally the N- or O- glycans were separated from the added enzyme by the Sep-Pak C18 chromatography as described in the section 2.2.6.
2.2.11 Permethylation procedure with subsequent glycan purification

5 pellets of sodium hydroxide per sample were ground in a dry mortar with pestle. 2 ml of dimethylsulphoxide (DMSO) were added to the ground powder and the slurry transferred to a glass tube containing N- or O-glycans. Subsequently 0.6 ml of methyl iodide were added and the sample was agitated for 20 min at room temperature. The reaction was terminated by the addition of few ml of ultra-pure water until the fizzing stopped. Then 1 ml of chloroform was added to the sample and the chloroform layer was washed three times with 3-4 ml of ultrapure water. After extraction the chloroform layer was dried under the nitrogen and re-dissolved in 200 µl of 50%/50% methanol/water solution.

A Sep-Pak C18 Classic cartridge was attached to a glass syringe and sequentially washed with 5 ml of methanol, 5 ml of water, 5 ml of acetonitrile and 15 ml of water. To purify the permethylated glycans the re-dissolved sample was loaded onto the cartridge and the column was washed with 5 ml of water. The bound permethylated glycans were sequentially eluted with 15%, 35%, 50%, 75% acetonitrile in water solutions. The volume of the fractions was reduced in a SpeedVac® with the subsequent freezing and lyophilisation of the sample for glycomics analysis (see next section 2.2.12). All fractions were collected however usually the glycans were detected in the 35% and 50% acetonitrile fractions.

2.2.12 MS analysis

The lyophilised permethylated glycans were dissolved in 10 µl of methanol. 1 µl of the solution was then mixed with 1 µl of 10mg/ml DHB in 70% (v/v) methanol-water solution for ABI Voyager-DE STR MALDI-TOF analysis. For the AB 4800 MALDI-TOF-TOF and AXIMA Resonance MALDI-QIT-TOF instrument (Kratos, Shimadzu, Manchester) instruments 1 µl of the sample solution was mixed with 1 µl of the 20 mg/ml DABP in 75% (v/v) acetonitrile-water solution. The mixture was then spotted onto the MALDI-plate and dried at room temperature. Usually for MALDI-MS experiments ABI Voyager-DE STR MALDI-TOF and AB 4800 MALDI-TOF-TOF were used and most of the MALDI-TOF-TOF experiments were performed by the AB 4800 MALDI-TOF-TOF instrument. The MS
experiments were performed on a AXIMA Resonance MALDI-QIT-TOF instrument (Kratos, Shimadzu, Manchester, courtesy of Dr. Roberto Castangia). All mass spectra were acquired in the positive ion mode.

2.2.13 Sample preparation for GC-MS linkage analysis

After MALDI analyses the permethylated samples were hydrolysed by adding 200 µl 2 M trifluoroacetic acid (TFA) and incubating for 2h at 121 ºC. Subsequently into a cooled and nitrogen-dried sample 200 µl of 10 mg/mL NaBD₄ in 2 M NH₄OH solution was added and the reaction was carried out for 2h at room temperature. The reaction was terminated by adding 5 drops of acetic acid and allowing it to dry (not completely) under a stream of nitrogen. The excess of borates was removed by co-evaporation with 10% (v/v) acetic acid in methanol under a stream of nitrogen. Then 200 µl of acetic anhydride was added and incubated for 1h at 100 ºC with subsequent drying under a stream of nitrogen. The sample was then purified by adding 1 mL of chloroform and washed three times with a few ml of water discarding the upper water layer. Subsequently the remaining chloroform was dried down under a stream of nitrogen. The resultant partially methylated alditol acetates (PMAA) were dissolved in 5-50 µl of hexane and injected manually into Bruker SCION SQ 456-GC or PERKIN Elmer Clarus 500 instruments.

2.2.14 Data reproducibility

MALDI analyses of PNGase F, PNGase A released N-glycans, reductively eliminated O-glycans of adult and L3 stage H. contortus were repeated three times (except PNGase F released adult N-glycans 50 % acetonitrile fraction which was analysed twice). β1-4 specific glycosidase digestions on PNGase-F and PNGase- A released adult H. contortus N-glycans were repeated twice. Hydrofluoric acid treatment (HF) on PNGase F released adult H. contortus N-glycans was performed twice. The rest of the experiments for Chapter 3 were performed once.

MALDI analyses of adult D. viviparus O-glycans were performed two times. MALDI analyses of PNGase F released L3 stage D. viviparus N-glycans were performed twice. Reductively eliminated L3 stage D. viviparus O-glycans were analysed three times. The rest of the experiments for Chapter 4 were performed once.
For calculations of the mean relative abundance value, the formula: \( \text{mean} = \frac{\text{RAs1} + \text{RAs2} + \text{RAs3}}{3} \) was used where RA stands for relative abundance; s1, s2, s3 stands for samples one, two and three. For variance calculations the formula: \( \text{variance} = \frac{(\text{RAs1}-\text{mean})^2 + (\text{RAs2}-\text{mean})^2 + (\text{RAs3}-\text{mean})^2}{3} \) was used. For standard deviation (SD) calculations the formula: \( \text{SD} = \sqrt{\text{variance}} \) was used. For standard error of the mean (SEM) the formula: \( \text{SEM} = \frac{\text{SD}}{\sqrt{3}} \) was used.
Chapter 3

Structural characterization of

*Haemonchus contortus* N- and O-glycomes
3.1 Introduction

As introduced in section 1.1, *Haemonchus contortus* is an economically important gastrointestinal parasite infecting small ruminants such as goats and sheep. By feeding on animal blood the parasite causes anaemia, weight loss, reduction in milk production and occasional death of the host thus causing huge economic losses worldwide. Current control methods, anthelmintics are often failing due to increased resistance worldwide including the new anthelmintic class aminoacetonitrile derivatives (Scott et al., 2013). For this reason an alternative control strategy is urgently needed and the development of an effective and cheap vaccine is desired. Early immunization studies showed native excretory/secretory (ES) and gut-derived glycoproteins to be promising vaccine candidates. However, all attempts to create a recombinant glycoprotein vaccine to date have failed (Knox et al., 2003; Roberts et al., 2013). An alternative solution is the development of a carbohydrate-based vaccine (Astronomo and Burton, 2010). However, one of the bottlenecks in their development is the limited number of well-defined carbohydrate structures available from the parasite.

As discussed in section 1.2.3 in 1996 Haslam and colleagues performed the most rigorous structural studies on *H. contortus* N-glycans using FAB-MS instrumentation (Haslam et al., 1996). They revealed the previously unseen distal GlcNAc core fucosylation at the 3-position (Haslam et al., 1996). Interestingly this chitobiose core modification was not detected in the L3 stage (Haslam et al., 1998). In addition other structural features were revealed such as fucosylation of proximal GlcNAc at the 3- and 6- positions, presence of GalNAcβ1–4GlcNAc (LDN), GalNAcβ1–4(Fuc α1–3)GlcNAc (LDNF) antennae and high mannose structures. The presence of LDNF structure was also proposed by other researchers based on antibody response to *H. contortus* infection and Western Blot and glycan array studies (Geldhof et al., 2005; van Stijn et al., 2010b; Vervelde et al., 2003). In addition one of the groups proposed the Galα1-3GalNAc epitope based on Western blot glycan array studies using antibodies from sheep immunized with *H. contortus* ES products (van Stijn et al., 2010b). No structural data of O-glycans for either adult or L3 stage *H. contortus* has been reported to date.
As discussed in section 1.3, since the original studies on *H. contortus* N-glycans in 1996 (Haslam et al., 1996) the FAB-MS instrumentation has been largely superseded by the instruments with MALDI ionisation (Karas, Bachmann et al. 1985) which offers better signal to noise-ratio and higher sensitivity above \( m/z \) 3000 rendering the detection of minor structures more feasible (Dell and Morris, 2001). The superiority of MALDI-MS over FAB-MS instruments has been elegantly demonstrated by Babu and colleagues (Babu, North et al. 2009) and North and colleagues (North et al., 2010a) who reported previously unseen Chinese Hamster Ovary cell N-glycans up to \( m/z \) 13000 demonstrating the ultra-sensitivity of MALDI instrument at higher molecular weight.

In the studies described in this chapter my objective was to utilize next-generation MALDI-TOF-TOF and MALDI-QIT-TOF-based techniques to revisit the N-glycomes of adult and L3 stage *H. contortus*. Since no O-glycans data has been reported to date, in addition the O-glycomes of adult and L3 stage have also been characterised using the same MALDI-MS technologies aiming to discover new glycan structures which can be potentially used for carbohydrate-based vaccine development against this parasite.
3.2 Structural characterization of the N-glycome of adult *H. contortus*

3.2.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from adult *H. contortus*

In order to investigate the N-glycome of *H. contortus* by using MALDI-TOF MS instrumentation the glycoproteins extracted from adult worms were subjected to PNGase F digestion and the released N-glycans were permethylated. N-glycans were recovered in both the 35% acetonitrile and 50 % acetonitrile C18 chromatography fractions. The first one contained the largest diversity of N-glycans as seen in the Figure 30, top and Table 3 whereas the 50% acetonitrile fraction was mainly predominated by abundant Hex5,6HexNAc2 high-mannose structures as seen in the Figure 30, bottom and Table 4. The most abundant group of N-glycans observed in the MALDI-MS spectrum from the 35% acetonitrile shown in Figure 30, top and Table 3 are paucimannose (or truncated) structures with or without fucose and having the general composition of Hex2,4Fuc0,2HexNAc2. The second less abundant group of N-glycans corresponds to complex structures with or without fucose and general composition of Hex2,4Fuc0,2HexNAc3,5. Most of the molecular ion peaks except those at m/z 1386, 1590, 1661, 1906, 2039, 2254 were observed earlier using the FAB-MS technology (Haslam et al., 1996). At this point it has to be emphasized that the MALDI-TOF MS data provides the compositional data only and in order to obtain more structural information MALDI-TOF-TOF fragmentations of molecular ions at m/z 1141, 1315, 1345, 1386, 1416, 1519, 1549, 1590, 1661, 1835, 2009, 2039, 2213 seen in the Figure 30 were performed.
Fig. 30 MALDI MS spectra of PNGase F released adult *H. contortus* N-glycans

The 35% acetonitrile fraction is at top and 50% acetonitrile fraction is at the bottom. The structures marked with asterisk* were later subject to MALDI-TOF-TOF analyses. The blue dashed lines denote the positions of the minor peaks. The higher molecular weight region of the upper spectrum marked with the black horizontal line has been magnified vertically 12 times. It is important to note that MALDI-TOF-MS experiments provide compositional data only shown in Tables 3 and 4 and the proposed structures are based on other experiments described further in the chapter.
<table>
<thead>
<tr>
<th>m/z value</th>
<th>composition</th>
<th>RA s1</th>
<th>RA s2</th>
<th>RA s3</th>
<th>Mean</th>
<th>Variance</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>967</td>
<td>HexNAc2Hex2</td>
<td>0.24</td>
<td>0.16</td>
<td>0.40</td>
<td>0.27</td>
<td>0.0099</td>
<td>0.099</td>
<td>0.058</td>
</tr>
<tr>
<td>1141</td>
<td>HexNAc2Hex2Fuc1</td>
<td>0.58</td>
<td>0.36</td>
<td>0.52</td>
<td>0.49</td>
<td>0.0086</td>
<td>0.092</td>
<td>0.054</td>
</tr>
<tr>
<td>1171</td>
<td>HexNAc2Hex3</td>
<td>0.90</td>
<td>0.78</td>
<td>1</td>
<td>0.89</td>
<td>0.0081</td>
<td>0.089</td>
<td>0.052</td>
</tr>
<tr>
<td>1315</td>
<td>HexNAc2Hex2Fuc2</td>
<td>0.26</td>
<td>0.15</td>
<td>0.10</td>
<td>0.17</td>
<td>0.0045</td>
<td>0.066</td>
<td>0.039</td>
</tr>
<tr>
<td>1345</td>
<td>HexNAc2Hex3Fuc2</td>
<td>1</td>
<td>1</td>
<td>0.96</td>
<td>0.99</td>
<td>0.00036</td>
<td>0.018</td>
<td>0.011</td>
</tr>
<tr>
<td>1375</td>
<td>HexNAc2Hex4</td>
<td>0.32</td>
<td>0.24</td>
<td>0.24</td>
<td>0.27</td>
<td>0.0014</td>
<td>0.037</td>
<td>0.022</td>
</tr>
<tr>
<td>1386</td>
<td>HexNAc3Hex2Fuc1</td>
<td>0.025</td>
<td>0.026</td>
<td>0.022</td>
<td>0.024</td>
<td>2.8E-06</td>
<td>0.0017</td>
<td>0.0010</td>
</tr>
<tr>
<td>1416</td>
<td>HexNAc3Hex3</td>
<td>0.098</td>
<td>0.073</td>
<td>0.12</td>
<td>0.097</td>
<td>0.00037</td>
<td>0.019</td>
<td>0.011</td>
</tr>
<tr>
<td>1519</td>
<td>HexNAc2Hex3Fuc2</td>
<td>0.17</td>
<td>0.11</td>
<td>0.066</td>
<td>0.12</td>
<td>0.0018</td>
<td>0.042</td>
<td>0.025</td>
</tr>
<tr>
<td>1549</td>
<td>HexNAc2Hex4Fuc1</td>
<td>0.099</td>
<td>0.07</td>
<td>0.037</td>
<td>0.069</td>
<td>0.00064</td>
<td>0.025</td>
<td>0.015</td>
</tr>
<tr>
<td>1579</td>
<td>HexNAc2Hex5</td>
<td>0.74</td>
<td>0.68</td>
<td>0.22</td>
<td>0.55</td>
<td>0.054</td>
<td>0.23</td>
<td>0.13</td>
</tr>
<tr>
<td>1590</td>
<td>HexNAc3Hex3Fuc1</td>
<td>0.064</td>
<td>0.049</td>
<td>0.054</td>
<td>0.056</td>
<td>3.8E-05</td>
<td>0.0062</td>
<td>0.0036</td>
</tr>
<tr>
<td>1661</td>
<td>HexNAc4Hex3</td>
<td>0.037</td>
<td>0.032</td>
<td>0.022</td>
<td>0.030</td>
<td>3.8E-05</td>
<td>0.0062</td>
<td>0.0036</td>
</tr>
<tr>
<td>1783</td>
<td>HexNAc2Hex6</td>
<td>0.037</td>
<td>0.033</td>
<td>0.011</td>
<td>0.027</td>
<td>0.00013</td>
<td>0.011</td>
<td>0.0066</td>
</tr>
<tr>
<td>1835</td>
<td>HexNAc4Hex3Fuc1</td>
<td>0.039</td>
<td>0.034</td>
<td>0.017</td>
<td>0.030</td>
<td>8.8E-05</td>
<td>0.0094</td>
<td>0.0054</td>
</tr>
<tr>
<td>1906</td>
<td>HexNAc5Hex3</td>
<td>0.0041</td>
<td>0.0084</td>
<td>0.0046</td>
<td>0.0057</td>
<td>3.7E-06</td>
<td>0.0019</td>
<td>0.0011</td>
</tr>
<tr>
<td>2009</td>
<td>HexNAc4Hex3Fuc2</td>
<td>0.024</td>
<td>0.023</td>
<td>0.0046</td>
<td>0.017</td>
<td>8.0E-05</td>
<td>0.0089</td>
<td>0.0051</td>
</tr>
<tr>
<td>2039</td>
<td>HexNAc4Hex4Fuc1</td>
<td>0.0073</td>
<td>0.011</td>
<td>0.0046</td>
<td>0.0076</td>
<td>6.9E-06</td>
<td>0.0026</td>
<td>0.0015</td>
</tr>
<tr>
<td>2080</td>
<td>HexNAc5Hex3Fuc1</td>
<td>0.0032</td>
<td>0.0084</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2213</td>
<td>HexNAc4Hex4Fuc2</td>
<td>0.0067</td>
<td>0.012</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2254</td>
<td>HexNAc5Hex3Fuc2</td>
<td>0.0028</td>
<td>0.0063</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 3. Assignment and quantification of PNGase F released adult *H. contratus* N-glycans

The molecular ions were observed in 35% acetonitrile MALDI-TOF-MS spectrum shown in Figure 30, top. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD-standard deviation, SEM-standard error of the mean, N/D-not detected, N/A-non applicable, s-sample. For formulas and calculations of the mean, variance, SD, SEM see section ‘2.2.14’.
<table>
<thead>
<tr>
<th>m/z value</th>
<th>composition</th>
<th>RA s1</th>
<th>RA s2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1141</td>
<td>HexNAc₂Hex₂Fuc₁</td>
<td>0.017</td>
<td>N/D</td>
</tr>
<tr>
<td>1171</td>
<td>HexNAc₂Hex₃</td>
<td>0.042</td>
<td>0.084</td>
</tr>
<tr>
<td>1315</td>
<td>HexNAc₂Hex₂Fuc₂</td>
<td>0.058</td>
<td>0.074</td>
</tr>
<tr>
<td>1345</td>
<td>HexNAc₂Hex₃Fuc₂</td>
<td>0.59</td>
<td>0.34</td>
</tr>
<tr>
<td>1375</td>
<td>HexNAc₂Hex₄</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>1416</td>
<td>HexNAc₂Hex₃</td>
<td>0.0039</td>
<td>N/D</td>
</tr>
<tr>
<td>1519</td>
<td>HexNAc₂Hex₃Fuc₂</td>
<td>0.70</td>
<td>0.26</td>
</tr>
<tr>
<td>1549</td>
<td>HexNAc₂Hex₄Fuc₁</td>
<td>0.041</td>
<td>0.057</td>
</tr>
<tr>
<td>1579</td>
<td>HexNAc₂Hex₅</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1590</td>
<td>HexNAc₂Hex₃Fuc₁</td>
<td>0.023</td>
<td>N/D</td>
</tr>
<tr>
<td>1783</td>
<td>HexNAc₂Hex₆</td>
<td>0.48</td>
<td>0.32</td>
</tr>
<tr>
<td>1835</td>
<td>HexNAc₂Hex₃Fuc₁</td>
<td>0.011</td>
<td>0.032</td>
</tr>
<tr>
<td>1987</td>
<td>HexNAc₂Hex₇</td>
<td>0.52</td>
<td>0.32</td>
</tr>
<tr>
<td>2009</td>
<td>HexNAc₂Hex₄Fuc₂</td>
<td>0.011</td>
<td>N/D</td>
</tr>
<tr>
<td>2191</td>
<td>HexNAc₂Hex₈</td>
<td>0.52</td>
<td>0.25</td>
</tr>
<tr>
<td>2213</td>
<td>HexNAc₂Hex₄Fuc₂</td>
<td>0.038</td>
<td>0.030</td>
</tr>
<tr>
<td>2395</td>
<td>HexNAc₂Hex₅Fuc₂</td>
<td>0.26</td>
<td>0.13</td>
</tr>
</tbody>
</table>

**Table 4.** Assignment and quantification of PNGase F released adult *H. contortus* N-glycans. The molecular ions were observed in 50% acetonitrile MALDI-TOF-MS spectrum shown in Figure 30, **bottom**. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA stands for relative abundance, N/D for not detected, s for sample.
### 3.2.2 GC-MS linkage analysis of oligosaccharides released by PNGase F

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.81</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.47</td>
</tr>
<tr>
<td>13.80</td>
<td>118, 143, 203</td>
<td>4-linked fucose</td>
<td>0.01</td>
</tr>
<tr>
<td>14.03</td>
<td>87, 89, 115, 130, 131, 175, 190</td>
<td>2-linked fucose</td>
<td>0.01</td>
</tr>
<tr>
<td>14.28</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal mannose</td>
<td>1.00</td>
</tr>
<tr>
<td>14.52</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal galactose</td>
<td>0.12</td>
</tr>
<tr>
<td>15.36</td>
<td>129, 130, 161, 190</td>
<td>2-linked mannose</td>
<td>0.25</td>
</tr>
<tr>
<td>15.56</td>
<td>101, 118, 129, 161, 234</td>
<td>3-linked mannose</td>
<td>0.11</td>
</tr>
<tr>
<td>15.76</td>
<td>102, 118, 129, 189, 233</td>
<td>6-linked mannose</td>
<td>0.04</td>
</tr>
<tr>
<td>16.87</td>
<td>129, 130, 189, 190</td>
<td>2,6-linked mannose</td>
<td>0.01</td>
</tr>
<tr>
<td>17.08</td>
<td>118, 129, 189, 234</td>
<td>3,6-linked mannose</td>
<td>0.57</td>
</tr>
<tr>
<td>17.74</td>
<td>117, 159, 205</td>
<td>Terminal GlcNAc</td>
<td>0.04</td>
</tr>
<tr>
<td>18.25</td>
<td>117, 159, 205</td>
<td>Terminal GalNAc</td>
<td>0.04</td>
</tr>
<tr>
<td>18.63</td>
<td>117, 159, 233</td>
<td>4-linked GlcNAc</td>
<td>0.46</td>
</tr>
<tr>
<td>19.45</td>
<td>117, 159</td>
<td>3,4-linked GlcNAc</td>
<td>0.03</td>
</tr>
<tr>
<td>19.92</td>
<td>117, 159, 261</td>
<td>4,6-linked GlcNAc</td>
<td>0.05</td>
</tr>
</tbody>
</table>

**Table 5.** GC-MS analysis of partially methylated alditol acetates obtained from *H. contortus* N-glycans released by PNGase F. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.

The GC-MS linkage data analyses from PNGase F released N-glycans are summarized in Table 5. The elution peaks at 14.52 and 13.80 show the presence of terminal galactose and 4-linked fucose in *H. contortus* supporting the presence of the Galβ1-4Fuc moiety (see sections 3.2.3, 3.2.4 and 3.2.7). The presence of 2-linked fucose is consistent with the proposed Galα1-2Fuc structure 3-linked to distal GlcNAc (see sections 3.2.13 and 3.2.15) and the novel modified LDNF structure where αGal is 2-linked to fucose (see section 3.2.9). The presence of terminal, 3-linked, 6-linked, 2-linked, 3, 6-linked mannoses are consistent with the presence of paucimannosidic and high mannose structures. The 2, 6-linked and 2-linked mannoses suggest the potential attachment sites for non-reducing end GlcNAcs in complex or hybrid glycans whose presence is also supported by the terminal GlcNAc. 4-linked GlcNAc
is consistent with the distal chitobiose core GlcNAc as well as the LDN structure (see section 3.2.8). The presence of 3, 4-linked GlcNAc agrees with fucosylation at the 3-position of the chitobiose core distal GlcNAc as well as LDNF structures (see section 3.2.8). 4, 6-linked GlcNAcitol is consistent with fucosylation of reducing-end GlcNAc at 6-position in PNGase-F released N-glycans. The presence of terminal GalNAc supports the presence of LDN and Galα1-2Fuc-modified LDN structures (see sections 3.2.8 and 3.2.9, respectively).

3.2.3 MALDI-MS/MS analysis of permethylated PNGase-F released N-glycans

Exemplar MS/MS data are described for the molecular ion at \( m/z \) 1549 with the composition of FucHexHexNAc\(_2\) (see Figure 31). Unexpectedly the most abundant fragment ions at \( m/z \) 894 and 678 showed the presence of the Hex-Fuc structure attached to the Asn-linked proximal GlcNAc. This chitobiose core modification was not previously detected by the FAB-MS instrumentation (Haslam et al., 1996). In addition no peak at \( m/z \) 1139 was observed, which would correspond to the Z-type rather than Y-type loss (Domon, Costello et
al. 1988) of a Hex-Fuc structure, which is favourable when fucose is 3-linked to the HexNAc. The non-reducing end of this structural isomer consisting of three hexoses, most likely mannoses based on a previous publication (Haslam et al., 1996), is supported by the fragment ion at m/z 667. All these data therefore suggest that the Asn-linked GlcNAc is modified at the 6-position with the Hex-Fuc structure which has not previously been observed in *H. contortus*. In addition the fragment ions at m/z 474 and 1098 correspond to the fucosylated proximal GlcNAc. Together with the fragment ion at m/z 871 corresponding to the four non-reducing end hexoses, these fragment ions indicate that the minor structural isomer with all four hexoses is at the non-reducing end of the glycan, as would have been predicted from previous analysis of *H. contortus* N-glycosylation (Haslam et al., 1996). The Hex-Fuc modification of the proximal GlcNAc was also observed upon the MALDI-TOF-TOF MS/MS fragmentation of the N-glycans at m/z 1519, 1549 and 2213.

### 3.2.4 MALDI-QIT-TOF MS$^3$ analysis of permethylated PNGase F released N-glycan

![Figure 32](image_url)

**Figure 32** MALDI-QIT-TOF MS$^3$ spectrum of m/z 678 fragment ions from permethylated, PNGase F released m/z 1549 N-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments.
To obtain additional linkage information about the novel Hex-Fuc structure the fragment ion at \( m/z \) 678 from the permethylated PNGase F released \( m/z \) 1549 N-glycan (see Figure 31) was subject to MS\(^3\) analysis on a MALDI-QIT-TOF instrument (see Figure 32). The fragment ions at \( m/z \) 241, 460 and 259 confirmed the presence of a terminal hexose and fragment ions at \( m/z \) 433 and 286- internally linked fucose. The cross-ring fragments at \( m/z \) 299 and 342 indicated that the hexose is 4-linked to the fucose. The cross-ring fragments at \( m/z \) 475, 489 and 505 once again confirmed that the Hex-Fuc structure is 6-linked to the Asn-linked GlcNAc.

3.2.5 MALDI-MS analysis of permethylated PNGase F released, \( \alpha \)-mannosidase digested N-glycans

![Figure 33 MALDI-TOF MS spectrum of permethylated, \( \alpha \)-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (35% acetonitrile fraction). The structures proposed are based on further MALDI-TOF-TOF MS/MS fragmentations (fragmented peaks are marked with an asterisk*), previous publication (Haslam et al., 1996) and the knowledge of biosynthetic pathways in other organisms.](image-url)
An α-mannosidase digest was performed on PNGase F released N-glycans to trim the α-mannose residues on the non-reducing end of the N-glycans, thus reducing the structural heterogeneity (see Figure 33). The spectrum is dominated by molecular ions at *m/z* 763-HexHexNAc₂, *m/z* 937-Fuc₁HexHexNAc₂, *m/z* 967-Hex₂HexNAc₂, *m/z* 1141-Fuc₁Hex₂HexNAc₂ and *m/z* 1171-Hex₃HexNAc₂. Di-fucosylated molecular ions were also observed at *m/z* 1315-Fuc₂Hex₂HexNAc₂, *m/z* 1560-Fuc₂Hex₃HexNAc₃ and *m/z* 2010-Fuc₂Hex₄HexNAc₄.

The remaining Hex-Fuc chitobiose core modifications seen after the MALDI-TOF-TOF MS/MS analyses (for MS/MS of *m/z* 1141 and 1315 peaks see Figures 35 A) and 36 A)) also diminish the likelihood of them being Man-Fuc structures as seen in the opportunistic amoebal parasite *Acanthamoeba* (Schiller et al., 2012).

3.2.6 The β1-4 specific galactosidase treatment coupled with MALDI-MS analysis of permethylated PNGase F released, α-mannosidase digested N-glycans

The PNGase F released, α-mannosidase digested N-glycans were subject to β1-4 specific galactosidase digest in order to define the hexoses in Hex-Fuc structure with the resultant MALDI-TOF MS spectrum seen in the Figure 34. As seen by the relative abundance of molecular ion peaks alone, initial MALDI-MS comparison did not reveal any significant changes when compared to the MS spectrum of PNGase F released, α-mannosidase digested, N-glycans. This can be explained by the presence of several structural isomers constituting the MALDI-MS molecular ion peaks. For this reason MALDI-TOF-TOF MS/MS analysis was undertaken to compare molecular ions before and after exoglycosidase digestion.
Figure 34 MALDI-TOF MS spectra of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (35% acetonitrile fraction) prior (top) and after (bottom) the β1-4 galactosidase digestion. ‘m’ denotes the minor isomer present. The structures proposed are based on further MALDI-TOF-TOF MS/MS fragmentations (fragmented peaks are marked with an asterisk*), previous publication (Haslam et al., 1996) and the knowledge of biosynthetic pathways in other organisms.
3.2.7 The β1-4 specific galactosidase treatment coupled with MALDI-MS-MS analysis of permethylated PNGase F released, α-mannosidase digested N-glycans

The MALDI-TOF-TOF MS/MS analysis of the \( m/z \) 1315 molecular ion prior to enzymatic treatment revealed the presence of two structural isomers as annotated in Figure 35 A) where the Hex-Fuc structure is linked either to the 3-position of the internal GlcNAc or the 6-position of the Asn-linked GlcNAc residues. The first structural isomer is best represented by the fragment ions at \( m/z \) 474 and 864 corresponding to the fucosylated proximal GlcNAc and its loss from the molecular ion. This structural isomer is also supported by the ‘Z-type’ (Domon, Costello 1988) fragment ion at \( m/z \) 905 indicating the elimination, not cleavage of the Hex-Fuc structure confirming it being 3-linked. The second structural isomer is best represented by the fragment ions at \( m/z \) 678 and 660 corresponding to the Hex-Fuc modified proximal GlcNAc structure and its loss from the molecular ion. This structure is also supported by the abundant ‘Z-type’ (Domon, Costello 1988) fragment ion at \( m/z \) 1109 corresponding to the elimination of the fucose 3-linked to the distal GlcNAc.

Upon β1-4 specific galactosidase treatment as seen in Figure 35 B) we observe the virtually complete disappearance of the fragment ions at \( m/z \) 660 and 678 suggesting the disappearance of the structural isomer bearing Hex-Fuc unit 6-linked to the Asn-linked GlcNAc residue (seen on the right in Figure 35 A)). These results are also supported by the disappearance of the fragment ion at \( m/z \) 1109 suggesting no fucose linked to the 3-position of the GlcNAc present in the molecular ion. All these data put together define the presence of Galβ1-4-Fuc 6-linked to the Asn-linked GlcNAc residue.
Figure 35 MALDI-TOF-TOF MS/MS spectra of permethylated PNGase F released, α-mannosidase trimmed m/z 1315 N-glycan (top) and permethylated PNGase F released, α-mannosidase and subsequently β1-4 specific galactosidase trimmed m/z 1315 N-glycan (bottom). The blue arrowed horizontal lines represent the loss of indicated structural fragments from the parental molecular ion.
The exemplar MALDI-TOF-TOF MS/MS is also shown for the molecular ions at m/z 1141 prior and after the β1-4 specific galactosidase digestion. After the α-mannosidase digestion alone three structural isomers represent the molecular ion at m/z 1141 as seen in the Figure 36 A). The first structural isomer has the Hex-Fuc structure on the 3-position of the distal GlcNAc (see Figure 36 A) top left) and is best represented by the ‘Z- type’ (Domon, Costello 1988) fragment ions at m/z 864 and 731 corresponding to the eliminations of the proximal GlcNAc and Hex-Fuc structures respectively. The second structural isomer is core fucosylated on the 6-position of the proximal GlcNAc (see Figure 36 A) top right). Its presence is supported by the fragment ions at m/z 474, 690 and 463. The third structural isomer has the Hex-Fuc modification on the 6-position of the proximal GlcNAc (see Figure 36 A) bottom left) and is best represented by the fragment ions at m/z 678 and 486 corresponding to the Hex-Fuc-GlcNAc fragment ion and its loss from the parental molecular ion.

Upon the β1-4 specific galactosidase treatment (see Figure 36 B)) the disappearance of fragment ions at m/z 678 and 486 was observed. These results indicate the disappearance of the third structural isomer with the Hex-Fuc proximal GlcNAc modification thus confirming this structure having the Galβ1-4Fuc modification of the proximal GlcNAc.
Figure 36 MALDI-TOF-TOF MS/MS spectra of permethylated PNGase F released, α-mannosidase trimmed m/z 1141 N-glycan (top) and permethylated PNGase F released, α-mannosidase and subsequently β1-4 specific galactosidase trimmed m/z 1141 N-glycan (bottom). The blue arrowed horizontal lines represent the loss of indicated structural fragments from the parental molecular ion.
3.2.8 MALDI-MS-MS analyses of permethylated PNGase F released m/z 1661, 2010 N-glycans

To investigate whether two non-reducing end HexNAc’s of the molecular ion at m/z 1661 are joined together forming putative LDN antennae or form two separate putative terminal GlcNAc’s, this molecular ion was subjected to MALDI-TOF-TOF analysis. Due to the low abundance of the molecular ion, the spectrum seen in the Figure 37 is not of high quality and not all of the expected fragment ions are observed. On the other hand the fragment ions at m/z 527 and 1157 compositionally match the LDN structure. The fragment ion at m/z 953 corresponding to the loss of putative GalNAc-GlcNAc-Man together with the fragment ion at m/z 676 corresponding to the additional loss of reducing-end GlcNAc confirm the proposed non-reducing end GalNAc-GlcNAc structure. The fragment ions at m/z 1384 and 1402 represent the loss of reducing-end GlcNAc and non-reducing end GalNAc accordingly and are consistent with the proposed structure seen in the Figure 37. Although MALDI-TOF-TOF data cannot distinguish different HexNAc’s, the presence of the terminal GalNAc and 4-
linked GlcNAc seen in the GC-MS data (see section 3.2.2) makes HexNAc-HexNAc structure observed very likely to be LDN.

To investigate whether the molecular ion at m/z 2010 contains both fucoses on the chitobiose core or is fucosylated on the non-reducing end antennae, the MALDI-TOF-TOF analysis of this molecular ion was performed as seen in the Figure 38. As in the case of the previously described m/z 1661 molecular ion, due to low abundance of the parental ion, the spectrum produced is not of the highest quality. On the other hand the abundant fragment ions at m/z 701 and 1332 correspond to HexNAc-Fuc fragment ion and its loss from the molecular ion respectively. The fragment ion at m/z 1751 corresponds to the loss of the non-reducing end HexNAc. Together these data suggest the presence of the HexNAc-(Fuc)HexNAc non-reducing end antennae where fucose is linked to the innermost HexNAc. Together with the presence of the terminal GalNAc and 3, 4-linked GlcNAc observed in the GC-MS experiment (see section 3.2.2), these data strongly suggest the presence of the LDNF structure. The
fragment ions at m/z 474 and 1559 correspond to the fucosylated reducing-end GlcNAc and its loss, respectively and are consistent with the proximal GlcNAc fucosylation at 6-position as seen in the proposed structure in the Figure 38.

### 3.2.9 MALDI-QIT-TOF MS² analyses of permethylated PNGase F released m/z 2213 N-glycan

In order to get more structural information about the higher molecular weight structures the permethylated PNGase F released N-glycan at m/z 2213 was subject to MALDI-QIT-TOF MS² analysis (see Figure 39) and indicated the presence of two structural isomers. One of them has HexNAc₂Fuc non-reducing terminus antenna as indicated by the fragment ions at m/z 701 and 1535. The presence of Z-type, not Y-type loss (Domon, Costello et al. 1988) of fucose with resulting fragment ion at m/z 2007 indicates fucose being 3-linked suggesting the presence of fucosylated LacdiNAc. The proposed fucosylated LacdiNAc structure is also supported by the presence of the terminal GalNAc and 3,4-linked GlcNAc as seen in the GC-
MS linkage analysis (see section 2.2.2) although the latter structure is also coming from the 3-linked chitobiose core distal GlcNAc.

The MALDI-QIT-TOF MS\(^2\) analysis of \(m/z\) 2213 N-glycan also revealed the presence of a structural isomer bearing the non-reducing terminal antenna with a composition of HexNAc\(_2\)FucHex as seen by fragment ions at \(m/z\) 905 and 1331 (see Figure 39). The Z type fragment ion (Domon, Costello et al. 1988) at \(m/z\) 1803 suggests the presence of 3-linked Hex-Fuc epitope. The absence of the structural isomer having Hex-Fuc 3-linked to core GlcNAc suggests the presence of 3-linked Hex-Fuc within \(m/z\) 905 HexNAc\(_2\)FucHex non-reducing terminus antenna. The fragment ion at \(m/z\) 905 has been also observed in PNGase F released 50% acetonitrile fraction N-glycan molecular ions at \(m/z\) 2040, 2419, 2459 (data not shown).

![Figure 40](image.png)

**Figure 40** MALDI-QIT-TOF MS\(^3\) spectrum of \(m/z\) 905 fragment ions from permethylated PNGase F released \(m/z\) 2213 N-glycan. Dashed arrowed lines on the proposed structures represent glycosidic bond cleavages and cross-ring fragments.

To get more structural information, and to confirm the presence of the Hex-Fuc structure in the HexNAc\(_2\)FucHex fragment ion, the fragment ion at \(m/z\) 905 was further fragmented by the MALDI-QIT-TOF MS\(^3\) analysis as seen in the **Figure 40**. The fragment ions at \(m/z\) 433...
and 415 confirmed the presence of Hex-Fuc structure. The fragment ions at $m/z$ 282, 300 and 646 indicate that the Hex-Fuc structure is attached to the innermost HexNAc of the HexNAc$_2$FucHex fragment at $m/z$ 905.

![Figure 41](image)

**Figure 41** MALDI-QIT-TOF MS$^4$ spectrum of $m/z$ 433 fragment ion generated from precursor fragment ion $m/z$ 905. The latter was obtained from permethylated PNGase F released $m/z$ 2213 N-glycan. Dashed arrowed lines on the proposed structures represent glycosidic bond cleavages and cross-ring fragments.

To get the potential Hex-Fuc linkage information the fragment ion at $m/z$ 433 was further disassembled by MALDI-QIT-TOF MS$^4$ analysis (see **Figure 41**). The fragment ions at $m/z$ 215, 241 and 259 indicate the presence of terminal hexose. The cross-ring fragments at $m/z$ 301, 315, 343 suggest the hexose is 2-linked to the internal fucose, thus consistent with the previously unseen modified LDNF structure were fucose is capped at 2-position with hexose. The hexose is assumed to be the galactose monosaccharide based on the presence of the Galα1-2Fuc structures in other N-glycans and O-glycans of the *H. contortus* as will be described later in sections 3.2.15 and 3.3.4 respectively.
3.2.10 The summary of N-glycan structures released by PNGase F from adult *H. contortus*

Putting all the data together acquired so far in the thesis about PNGase F released N-glycans from adult *H. contortus*, the proposed structures with the corresponding m/z values can be seen below in the Table 6. The linkages for the proposed Gal-Fuc structures, the major focus of the section 3.2 have been denoted in the structures as well.

<table>
<thead>
<tr>
<th>m/z value</th>
<th>Proposed structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>967</td>
<td><img src="image" alt="Structure 1" /></td>
</tr>
<tr>
<td>1141</td>
<td><img src="image" alt="Structure 2" /></td>
</tr>
<tr>
<td>1171</td>
<td><img src="image" alt="Structure 3" /></td>
</tr>
<tr>
<td>1315</td>
<td><img src="image" alt="Structure 4" /></td>
</tr>
<tr>
<td>1345</td>
<td><img src="image" alt="Structure 5" /></td>
</tr>
<tr>
<td>1375</td>
<td><img src="image" alt="Structure 6" /></td>
</tr>
<tr>
<td>1386</td>
<td><img src="image" alt="Structure 7" /></td>
</tr>
<tr>
<td>1416</td>
<td><img src="image" alt="Structure 8" /></td>
</tr>
<tr>
<td>1519</td>
<td><img src="image" alt="Structure 9" /></td>
</tr>
<tr>
<td>1549</td>
<td><img src="image" alt="Structure 10" /></td>
</tr>
</tbody>
</table>
Table 6 showing the proposed N-glycan structures released by PNGase F from adult *H. contortus*.
3.2.11 MALDI-MS analysis of permethylated PNGase A released N-glycans from adult *H. contortus*

N-glycans resistant to digestion with PNGase F were subjected to PNGase A digestion. This enzyme, in contrast to PNGase F, cleaves N-glycans bearing a 3-linked fucose attached to the Asn-linked GlcNAc residue. The MALDI-MS spectra from 35% and 50% acetonitrile fraction shown in Figure 42 revealed that the major family of N-glycans corresponds to heavily fucosylated paucimannose structures with the general composition of Hex$_2$-4Fuc$_1$-$\beta$HexNAc$_2$. Paucimannosidic and high mannose structures with general composition Hex$_3$-$\alpha$HexNAc$_2$ are also observed due to their incomplete removal by the PNGase F enzyme.

The MALDI-TOF MS spectrum in Figure 42 and Tables 7 and 8 display proposed structures based on further MS/MS analyses. Thus detailed MALDI-TOF-TOF MS/MS analyses of peaks at m/z 1315, 1345, 1489, 1519, 1723, 1897 have been performed. The exemplary data for the peaks at m/z 1693 and 1519 will be described in the following sections 3.2.13, 3.2.1 and 3.2.15.
Figure 42 MALDI MS spectra of permethylated N-glycans released from the *H. contortus* adult stage whole worm extracts by PNGase A enzyme (35% acetonitrile fraction at top and 50% acetonitrile fraction at the bottom). The structures marked with asterisk* were later subject to MALDI-TOF-TOF analyses. For compositional assignment of each peak see Tables 7 and 8 below.
<table>
<thead>
<tr>
<th>m/z value</th>
<th>composition</th>
<th>RA s1</th>
<th>RA s2</th>
<th>RA s3</th>
<th>Mean</th>
<th>Variance</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1315</td>
<td>HexNAc2Hex2Fuc2</td>
<td>0.39</td>
<td>0.40</td>
<td>0.25</td>
<td>0.35</td>
<td>0.0047</td>
<td>0.068</td>
<td>0.040</td>
</tr>
<tr>
<td>1345</td>
<td>HexNAc2Hex3Fuc1</td>
<td>0.42</td>
<td>0.49</td>
<td>0.26</td>
<td>0.39</td>
<td>0.0093</td>
<td>0.096</td>
<td>0.056</td>
</tr>
<tr>
<td>1375</td>
<td>HexNAc2Hex4</td>
<td>0.18</td>
<td>0.21</td>
<td>0.16</td>
<td>0.18</td>
<td>0.00042</td>
<td>0.021</td>
<td>0.012</td>
</tr>
<tr>
<td>1489</td>
<td>HexNAc2Hex3Fuc2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1519</td>
<td>HexNAc2Hex3Fuc2</td>
<td>0.85</td>
<td>0.85</td>
<td>0.69</td>
<td>0.80</td>
<td>0.0057</td>
<td>0.075</td>
<td>0.044</td>
</tr>
<tr>
<td>1579</td>
<td>HexNAc2Hex5</td>
<td>0.31</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
<td>2.2E-05</td>
<td>0.0047</td>
<td>0.0027</td>
</tr>
<tr>
<td>1693</td>
<td>HexNAc2Hex3Fuc3</td>
<td>0.30</td>
<td>0.23</td>
<td>0.24</td>
<td>0.26</td>
<td>0.00096</td>
<td>0.031</td>
<td>0.018</td>
</tr>
<tr>
<td>1723</td>
<td>HexNAc2Hex4Fuc2</td>
<td>0.085</td>
<td>0.080</td>
<td>0.049</td>
<td>0.071</td>
<td>0.00025</td>
<td>0.016</td>
<td>0.0092</td>
</tr>
<tr>
<td>1783</td>
<td>HexNAc2Hex6</td>
<td>0.0099</td>
<td>0.0093</td>
<td>0.017</td>
<td>0.012</td>
<td>1.2E-05</td>
<td>0.0035</td>
<td>0.0020</td>
</tr>
<tr>
<td>1897</td>
<td>HexNAc2Hex4Fuc3</td>
<td>0.0090</td>
<td>0.0068</td>
<td>0.0073</td>
<td>0.0077</td>
<td>8.9E-07</td>
<td>0.00094</td>
<td>0.00054</td>
</tr>
</tbody>
</table>

**Table 7.** The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum ([Figure 42, top](#)) of permethylated (35% acetonitrile fraction), PNGase A released N-glycans of adult *H. contortus*. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD- standard deviation, SEM- standard error of the mean, N/D- not detected, N/A- non applicable, s- sample. For formulas and calculations of the mean, variance, SD and SEM see section ‘2.2.14’.
<table>
<thead>
<tr>
<th>m/z value</th>
<th>composition</th>
<th>RA s1</th>
<th>RA s2</th>
<th>RA s3</th>
<th>Mean</th>
<th>Variance</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1315.3</td>
<td>HexNAc₂Hex₂Fuc₂</td>
<td>0.044</td>
<td>0.29</td>
<td>0.15</td>
<td>0.16</td>
<td>0.010</td>
<td>0.10</td>
<td>0.058</td>
</tr>
<tr>
<td>1345.3</td>
<td>HexNAc₂Hex₃Fuc₁</td>
<td>0.31</td>
<td>0.32</td>
<td>0.17</td>
<td>0.27</td>
<td>0.0047</td>
<td>0.068</td>
<td>0.040</td>
</tr>
<tr>
<td>1375.3</td>
<td>HexNAc₂Hex₄</td>
<td>0.16</td>
<td>0.21</td>
<td>0.16</td>
<td>0.18</td>
<td>0.00056</td>
<td>0.024</td>
<td>0.014</td>
</tr>
<tr>
<td>1489.3</td>
<td>HexNAc₂Hex₅Fuc₂</td>
<td>0.20</td>
<td>1</td>
<td>1</td>
<td>0.73</td>
<td>0.14</td>
<td>0.38</td>
<td>0.22</td>
</tr>
<tr>
<td>1519.4</td>
<td>HexNAc₂Hex₆Fuc₂</td>
<td>0.72</td>
<td>1</td>
<td>0.78</td>
<td>0.83</td>
<td>0.014</td>
<td>0.12</td>
<td>0.069</td>
</tr>
<tr>
<td>1579.4</td>
<td>HexNAc₂Hex₇</td>
<td>0.67</td>
<td>0.64</td>
<td>0.72</td>
<td>0.68</td>
<td>0.0011</td>
<td>0.033</td>
<td>0.019</td>
</tr>
<tr>
<td>1693.4</td>
<td>HexNAc₂Hex₃Fuc₃</td>
<td>1</td>
<td>0.82</td>
<td>0.89</td>
<td>0.90</td>
<td>0.0055</td>
<td>0.074</td>
<td>0.043</td>
</tr>
<tr>
<td>1723.4</td>
<td>HexNAc₂Hex₄Fuc₂</td>
<td>0.049</td>
<td>0.095</td>
<td>0.082</td>
<td>0.075</td>
<td>0.00037</td>
<td>0.019</td>
<td>0.011</td>
</tr>
<tr>
<td>1783.4</td>
<td>HexNAc₂Hex₅</td>
<td>0.23</td>
<td>0.085</td>
<td>0.14</td>
<td>0.15</td>
<td>0.0036</td>
<td>0.060</td>
<td>0.035</td>
</tr>
<tr>
<td>1897.4</td>
<td>HexNAc₂Hex₆Fuc₃</td>
<td>0.039</td>
<td>0.059</td>
<td>0.070</td>
<td>0.056</td>
<td>0.00016</td>
<td>0.013</td>
<td>0.0074</td>
</tr>
<tr>
<td>1987.4</td>
<td>HexNAc₂Hex₇</td>
<td>0.22</td>
<td>0.058</td>
<td>0.11</td>
<td>0.13</td>
<td>0.0046</td>
<td>0.068</td>
<td>0.039</td>
</tr>
<tr>
<td>2191.4</td>
<td>HexNAc₂Hex₈</td>
<td>0.20</td>
<td>0.044</td>
<td>0.097</td>
<td>0.11</td>
<td>0.0042</td>
<td>0.065</td>
<td>0.037</td>
</tr>
<tr>
<td>2395.5</td>
<td>HexNAc₂Hex₉</td>
<td>0.023</td>
<td>0.011</td>
<td>0.028</td>
<td>0.021</td>
<td>0.000051</td>
<td>0.0071</td>
<td>0.0041</td>
</tr>
</tbody>
</table>

**Table 8.** The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (Figure 42, bottom) of permethylated (50% acetonitrile fraction), PNGase A released N-glycans of adult *H. contortus*. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD- standard deviation, SEM-standard error of the mean, N/D- not detected, N/A- non applicable, s- sample. For formulas and calculations of the mean, variance, SD and SEM see section ‘2.2.14’.
3.2.12 GC-MS linkage analysis of oligosaccharides released by PNGase A

The GC-MS linkage data analyses from PNGase A released N-glycans are shown in the Table 9. The elution peaks at 18.69 and 18.20 demonstrate the presence of terminal galactose and 2-linked fucose in *H. contortus* which is consistent with suggested Galα1-2Fuc structure 3-linked to the distal GlcNAc (see sections 3.2.13 and 3.2.15). The terminal fucose is consistent with the earlier reported (Haslam et al., 1996) chitobiose core modifications at the 3-position of the distal GlcNAc and 3, 6-positions of the proximal GlcNAc. The latter is also supported by the presence of the 3, 4-linked GlcNAc. The 4-linked GlcNAc is consistent with the non-fucosylated distal GlcNAc. The terminal, 2-linked, 3-linked, 3, 6-linked mannoses agree with some high mannose structures observed in the MALDI-TOF MS spectrum.

Although seen in the MALDI spectra, the absence of 4-linked fucose is explained by lower abundance of Hex-Fuc structure attached to the Asn-linked proximal GlcNAc compared to distal GlcNAc and limitations of GC-MS instrument sensitivity.

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.93</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.47</td>
</tr>
<tr>
<td>18.20</td>
<td>87,89,115,130,131,175,190</td>
<td>2-linked fucose</td>
<td>0.15</td>
</tr>
<tr>
<td>18.42</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal mannose</td>
<td>1.00</td>
</tr>
<tr>
<td>18.69</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal galactose</td>
<td>0.15</td>
</tr>
<tr>
<td>19.58</td>
<td>129, 130, 161, 190</td>
<td>2-linked mannose</td>
<td>0.14</td>
</tr>
<tr>
<td>19.78</td>
<td>101, 118, 129, 161, 234</td>
<td>3-linked mannose</td>
<td>0.16</td>
</tr>
<tr>
<td>21.32</td>
<td>118, 129, 189, 234</td>
<td>3,6-linked mannose</td>
<td>0.06</td>
</tr>
<tr>
<td>23.12</td>
<td>117, 159, 233</td>
<td>4-linked GlcNAc</td>
<td>0.01</td>
</tr>
<tr>
<td>23.97</td>
<td>117, 159</td>
<td>3,4-linked GlcNAc</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 9. GC-MS analysis of partially methylated alditol acetates obtained from *H. contortus* N-glycans released by PNGase A. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.
3.2.13 MALDI-MS/MS analysis of permethylated PNGase A released N-glycans

The molecular ion at $m/z$ 1693 from the MALDI-TOF MS spectrum of 50% acetonitrile PNGase A released N-glycans seen in the Figure 42 was further analysed by the MALDI-TOF-TOF instrumentation. As seen in Figure 43, the presence of a fragment ion at $m/z$ 433 suggests the presence of a Hex-Fuc structure and its corresponding Z, not Y-type ion (Domon, Costello et al. 1988) at $m/z$ 1283 suggests it is 3-linked. The ions at $m/z$ 648 and 1068 correspond to the Asn-linked GlcNAc modified with two fucoses at the 3- and 6-positions as published previously (Haslam et al., 1996). In addition the lack of fragment ions at $m/z$ 678, 852, 1056 corresponding to Asn-linked GlcNAc residues modified with Hex-Fuc, Hex-Fuc(Fuc), (Hex-Fuc)$_2$ structures, rules out the attachment of the observed Hex-Fuc structure to the Asn-linked GlcNAc residue. The ion at $m/z$ 463 represents the fragment consisting of two non-reducing end hexoses and the ion at $m/z$ 676 is a double cleavage Hex$_2$HexNAc fragment which has lost the Hex-Fuc and Asn-linked GlcNAc modified with two fucoses. Although the Hex-Fuc structure attached to the 3-position the internal chitobiose...
core GlcNAc has been introduced earlier (see Figures 33, 34, 35, 36), the data presented in this section for the first time provides a strong evidence for this type of chitobiose core modification in *H. contortus*.

### 3.2.14 MALDI-QIT-TOF MS analyses of permethylated PNGase A released m/z 1693 N-glycans

**Figure 44** MALDI-QIT-TOF MS$^2$ spectrum of 35% acetonitrile permethylated PNGase A released m/z 1693 adult *H. contortus* N-glycan. The blue horizontal arrowed lines represent the loss of indicated structural fragments from the parental molecular ion.

In contrast to the molecular ion at m/z 1693 from the 50% acetonitrile PNGase A released N-glycans spectrum (see earlier MS and MS/MS Figures 42, bottom and 43), the same molecular ion found in the 35% acetonitrile fraction displays an additional major structural isomer as seen in the Figure 44 above. The fragment ions at m/z 852 and 864 correspond to the HexNAcFuc2Hex fragment. Its loss from the parental molecular ion clearly suggests a new structural isomer where proximal GlcNAc is di-fucosylated, as previously reported (Haslam et al., 1996), and one of the fucoses is capped with the hexose. At this point we already know that the capping hexose can form the Galβ1-4Fuc attached to the 6-position of
the proximal GlcNAc (see sections 3.2.3, 3.2.4 and 3.2.7), and the fragment ions at \( m/z \) 628 and 646 suggest its presence, however the hexose can also be attached to the fucose 3-linked to the proximal GlcNAc.

In order to discriminate the 3- and/or 6-linked attachment site for the Hex-Fuc structure on the proximal GlcNAc the fragment ion at \( m/z \) 852 was further analysed using the MALDI-QIT-TOF MS\(^3\) instrument as seen in the Figure 45. The fragment ion at \( m/z \) 646 corresponds to Z, not Y ion-type loss (Domon, Costello et al. 1988) of fucose which favours the 3-position of HexNAc and the most abundant fragment ion at \( m/z \) 573 results from Z-ion type loss (Domon, Costello et al. 1988) of fucose with concomitant elimination of AcNH(CH\(_3\)). The cross-ring \( m/z \) 475 fragment ion indicates the Hex-Fuc epitope is 6-linked to the Asn-linked GlcNAc. Together with the absence of the \( m/z \) 472 peak, which would correspond to Z-ion type loss (Domon, Costello et al. 1988) of Hex-Fuc fragment from the 3-position of GlcNAc, all the data suggest the Hex-Fuc epitope is 6-linked, rather than 3-linked, to the Asn-linked GlcNAc in the PNGase-A released N-glycan pool. The MALDI-TOF-TOF analysis of permethylated PNGase A released N-glycan \( m/z \) 1897 N-glycan (data not shown).
also revealed the presence of a fragment ion at *m/z* 852. In this case the fragmentation pattern also confirmed the Hex-Fuc structure being 6-linked and the fucose 3-linked to Asn-linked GlcNAc.

### 3.2.15 MALDI-QIT-TOF MS^3 analysis of *m/z* 433 fragment ion from permethylated PNGase A released *m/z* 1519 N-glycan

![Figure 46](image)

**Figure 46** MALDI-QIT-TOF MS^2 spectrum of permethylated PNGase A released *m/z* 1519 N-glycan. The blue horizontal arrowed lines represent the loss of indicated structural fragments from the parental molecular ion.

The MALDI-QIT-TOF MS^2 analysis of permethylated PNGase A released *m/z* 1519 N-glycan (see Figure 46) shows the presence of two structural isomers with only one of them bearing the Hex-Fuc epitope which is 3-linked to the distal GlcNAc. The presence of the Hex-Fuc structure is seen by the fragment ion at *m/z* 433 and its elimination, not cleavage, from the parental molecular ion, as seen by the ‘Z-type’ fragment ion at *m/z* 1109, indicates it is linked to the 3-position of the GlcNAc. This structural isomer (see Figure 46 lower structure) carries the monofucosylated proximal GlcNAc as seen by the fragment ions at *m/z* 474 and 1068. The second structural isomer (see Figure 46 upper structure) has the
difucosylated proximal GlcNAc, as indicated by the fragment ions at \( m/z \) 648 and 894, and has no Hex-Fuc modification.

**Figure 47** MALDI-QIT-TOF MS\(^3\) spectrum of \( m/z \) 433 fragment ion from permethylated PNGase A released \( m/z \) 1519 N-glycan. Dashed arrowed lines on the proposed structures represent glycosidic bond cleavages and cross-ring fragments.

As it has now been defined that the Hex-Fuc fragment is certainly coming from the 3-position of the distal GlcNAc, to further characterize it the fragment ion at \( m/z \) 433 was subjected to MALDI-QIT-TOF MS\(^3\) analysis (see **Figure 47**). The fragment ions at \( m/z \) 259, 241, 215 indicate the presence of terminal hexose and internal fucose. The presence of an abundant cross-ring fragment at \( m/z \) 301 and other cross-ring fragments at \( m/z \) 155, 315 and 343 suggest that the hexose is linked to the 2-position of the internal fucose. The data described, together with the presence of the terminal galactose and 2-linked fucose in the GC-MS data, (see section 3.2.12) and the knowledge of biosynthetic pathways in closely related *C. elegans* organisms, allows us to propose the Galα1-2Fuc modification of the distal GlcNAc at the 3-position. The same modification was proposed to be present on the 3-position of the GlcNAc in the non-reducing end LDN structure described earlier in the **Figure 41**.
3.2.16 MALDI-MS and MALDI-MS/MS analyses of permethylated PNGase F and PNGase A released, broad specificity β-galactosidase and broad specificity α-galactosidase digested N-glycans

The adult *H. contortus* N-glycans released by the PNGase F and PNGase A were also subjected to bovine testes broad specificity β-galactosidase and broad specificity green coffee bean α-galactosidase digestions with subsequent MALDI-TOF MS and MALDI-TOF-TOF MS/MS analyses. No evidence of digestion could be observed (data not shown).

3.2.17 MALDI-MS and MALDI-MS/MS analyses of permethylated PNGase A released, reducing end ring-opened N-glycans digested with the range of galactosidase enzymes

The Galα1-2Fuc structure attached to the distal GlcNAc of the chitobiose core is resistant to the broad specificity green coffee bean α-galactosidase enzyme in *C. elegans* (personal communication with Ian B. Wilson), nevertheless the authors managed to show the partial removal of the galactose on the HPLC-fractionated N-glycan using high concentration of enzyme and long incubation time (Paschinger et al., 2011). To eliminate potential steric hindrance of enzyme access, the reducing-end GlcNAc was ring-opened using sodium borodeuteride and then the sample was subjected to green coffee bean α-galactosidase, bovine testes broad specificity β-galactosidase and β1-4 specific galactosidase digests. No differences where observed when comparing the resultant MALDI-TOF MS and MALDI-TOF-TOF MS/MS spectra prior and after the digests.

3.2.18 The summary of N-glycan structures released by PNGase A from adult *H. contortus*

Putting all the data together acquired so far in the thesis about PNGase A released N-glycans from adult *H. contortus*, the proposed structures with the corresponding *m/z* values can be seen below in the *Table 10*. The linkages for the proposed Gal-Fuc structures which have been the major focus of the section 3.2 ‘Structural characterization of N-glycome of adult *H. contortus*’ have been denoted in the structures as well.
<table>
<thead>
<tr>
<th>m/z value</th>
<th>Proposed structure(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1315</td>
<td>![Structure 1315]</td>
</tr>
<tr>
<td>1345</td>
<td>![Structure 1345]</td>
</tr>
<tr>
<td>1375*</td>
<td>![Structure 1375*]</td>
</tr>
<tr>
<td>1489</td>
<td>![Structure 1489]</td>
</tr>
<tr>
<td>1519</td>
<td>![Structure 1519]</td>
</tr>
<tr>
<td>1579*</td>
<td>![Structure 1579*]</td>
</tr>
<tr>
<td>1693</td>
<td>![Structure 1693]</td>
</tr>
<tr>
<td>1723</td>
<td>![Structure 1723]</td>
</tr>
<tr>
<td>1783*</td>
<td>![Structure 1783*]</td>
</tr>
<tr>
<td>1897</td>
<td>![Structure 1897]</td>
</tr>
<tr>
<td>1987*</td>
<td>![Structure 1987*]</td>
</tr>
<tr>
<td>2191*</td>
<td>![Structure 2191*]</td>
</tr>
<tr>
<td>2395*</td>
<td>![Structure 2395*]</td>
</tr>
</tbody>
</table>
Table 10 showing the proposed N-glycan structures released by PNGase A from adult *H. contortus*.

3.2.19 MALDI-TOF MS analysis of α-mannosidase digested, hydrofluoric acid treated N-glycans released by PNGase F enzyme

The presence of the phosphorylcholine (PC) moiety on glycans makes them difficult to detect by mass-spectrometry due to the PC zwitterionic nature. However, an indirect way to detect PC-containing glycans (Haslam et al., 1997) by removing the PC moiety with hydrofluoric acid has previously been developed. The appearance or up-regulation of structures with terminal HexNAc (Harnett et al., 2010) would usually indicate these N-glycans previously being modified by PC.

In order to investigate the potential sub-population of N-glycans in adult *H. contortus* containing PC moieties, the PNGase F released and α-mannosidase digested N-glycans were treated with hydrofluoric acid. The resultant MALDI-TOF MS spectra prior and after the treatment are shown in the Figures 48 and 49. Before the HF treatment the PNGase F released, α-mannosidase digested N-glycan MS spectrum revealed molecular ion peaks up to m/z 3959 with glycans having up to five fucoses and eight HexNAcs. The MALDI-TOF-TOF analyses (see proposed structures in the Figure 49) of molecular ions at m/z 1969, 2010, 2040, 2214, 2388, 2459, 2083, 3328 displayed the fragment ion at m/z 905 in all of the MALDI-TOF-TOF MS/MS spectra. These results suggested the presence of the proposed Galα1-2Fuc1-3(GalNAcβ1-4)GlcNAc structure (see section 3.2.9) in all of the aforementioned molecular ions although for some peaks due to their low abundance more detailed structural analysis could not be performed.

After the hydrofluoric acid treatment the MALDI-TOF MS analysis revealed the presence of molecular ions with m/z values up to 3868 and structures having up to thirteen HexNAc and only one fucose. The MALDI-TOF-TOF MS/MS analyses of molecular ions at m/z 1907, 2040, 2530, 3020, 3061, 3551 revealed structures with poly-HexNAc antennae with up to five residues and there was evidence for two antennae. Fucose, if present, is attached to the 6-position of the proximal GlcNAc.
Figure 48 MALDI-TOF MS m/z 500-5015 range spectra of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (50% acetonitrile fraction) prior (top) and after (bottom) the hydrofluoric acid treatment. The presence of oligomannosyl core is based on the knowledge of biosynthetic pathways in other organisms, previous publication (Haslam et al., 1996) and GC-MS data (see section 3.2.2).
MALDI-TOF MS m/z 1900-4000 range spectra of permethylated, α-mannosidase digested N-glycans released from the *H. contortus* adult parasite by PNGase F enzyme (50% acetonitrile fraction) prior (top) and after (bottom) the hydrofluoric acid treatment. The structures proposed are based on the further MALDI-TOF-TOF MS/MS fragmentations (peaks marked with an asterisk*), the knowledge of biosynthetic pathways in other organisms and data described in the thesis.
3.2.20 MALDI-TOF-TOF MS/MS analysis of α-mannosidase digested, hydrofluoric acid treated N-glycans released by PNGase F enzyme

![MALDI-TOF-TOF MS/MS spectrum](image)

**Figure 50** MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 3551 N-glycan digested with α-mannosidase and treated with hydrofluoric acid. The blue arrowed horizontal lines represent the loss of indicated structural fragments from the parental molecular ion.

The molecular ions at m/z 1907, 2040, 2530, 3020, 3061, 3551 from the MALDI-TOF MS spectrum (see Figure 49) were further analysed by the MALDI-TOF-TOF MS/MS instrumentation and the exemplary data is shown for the molecular ion at m/z 3551 in the Figure 50. The fragment ion at m/z 3100 indicates the fucosylated proximal GlcNAc. The ‘B-type’ (Domon, Costello 1988) fragment ions at m/z 772 and 1017 indicate the presence of the HexNAc₃ and HexNAc₄ fragments. The sequential loss of up to five HexNACs as seen by the fragment ions at m/z 3292, 3047, 2802, 2557, 2312 suggest the presence of the linear poly-HexNAc antenna with up to 5 HexNACs. The fragment ions at m/z 2053, 1808 and 1563 indicate the presence of two separate antennae with up to 4 HexNAc residues on each. Together all this data suggest the most probable structure with two poly-HexNAc antennae with up to five units on each although other structural isomers are not excluded.
3.3 Structural characterization of the O-glycome of adult *H. contortus*

**3.3.1 MALDI-MS analysis of permethylated O-glycans released by reductive elimination from adult *H. contortus***

To date there are no literature reports on O-glycomic data of *H. contortus*. To investigate whether any O-glycans are present the glycopeptides remaining after the PNGase A digestion were subject to the reductive elimination followed by the MALDI-TOF MS analysis (see Figure 51, Table 11).

![MALDI-MS spectrum of permethylated O-glycans released from the *H. contortus* adult stage whole worm extracts by reductive elimination (35% acetonitrile fraction). The structures marked with asterisk* were later subject to MALDI-TOF-TOF analyses. For compositional assignment of each peak see Table 11. The 50% acetonitrile fraction has not been shown as no new structures have been detected.](image-url)
<table>
<thead>
<tr>
<th>m/z value</th>
<th>composition</th>
<th>RA s1</th>
<th>RA s2</th>
<th>RA s3</th>
<th>Mean</th>
<th>Variance</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>534</td>
<td>HexHexNAc</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>575</td>
<td>HexNAc₂</td>
<td>0.50</td>
<td>0.22</td>
<td>0.18</td>
<td>0.30</td>
<td>0.020</td>
<td>0.14</td>
<td>0.082</td>
</tr>
<tr>
<td>738</td>
<td>Hex₂HexNAc</td>
<td>0.29</td>
<td>0.15</td>
<td>0.16</td>
<td>0.20</td>
<td>0.0041</td>
<td>0.064</td>
<td>0.037</td>
</tr>
<tr>
<td>779</td>
<td>Hex HexNAc₂</td>
<td>0.47</td>
<td>0.25</td>
<td>0.23</td>
<td>0.32</td>
<td>0.012</td>
<td>0.11</td>
<td>0.063</td>
</tr>
<tr>
<td>912</td>
<td>FucHex₂HexNAc</td>
<td>0.060</td>
<td>0.024</td>
<td>0.017</td>
<td>0.034</td>
<td>0.00035</td>
<td>0.019</td>
<td>0.011</td>
</tr>
<tr>
<td>942</td>
<td>Hex₃HexNAc</td>
<td>0.080</td>
<td>0.012</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>953</td>
<td>FucHexHexNAc₂</td>
<td>0.030</td>
<td>0.010</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>983</td>
<td>Hex₂HexNAc₂</td>
<td>0.020</td>
<td>0.0074</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>1157</td>
<td>FucHex₂HexNAc₂</td>
<td>0.060</td>
<td>0.043</td>
<td>0.025</td>
<td>0.043</td>
<td>0.00020</td>
<td>0.014</td>
<td>0.0083</td>
</tr>
<tr>
<td>1198</td>
<td>FucHexHexNAc₃</td>
<td>0.24</td>
<td>0.098</td>
<td>0.056</td>
<td>0.13</td>
<td>0.0062</td>
<td>0.079</td>
<td>0.045</td>
</tr>
<tr>
<td>1361</td>
<td>FucHex₁HexNAc₂</td>
<td>0.020</td>
<td>0.075</td>
<td>0.050</td>
<td>0.048</td>
<td>0.00051</td>
<td>0.022</td>
<td>0.013</td>
</tr>
<tr>
<td>1402</td>
<td>FucHex₂HexNAc₃</td>
<td>0.020</td>
<td>0.49</td>
<td>0.37</td>
<td>0.29</td>
<td>0.040</td>
<td>0.20</td>
<td>0.12</td>
</tr>
<tr>
<td>1606</td>
<td>FucHex₃HexNAc₃</td>
<td>0.090</td>
<td>0.073</td>
<td>0.053</td>
<td>0.072</td>
<td>0.00023</td>
<td>0.015</td>
<td>0.0087</td>
</tr>
</tbody>
</table>

**Table 11.** The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 51) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of adult *H. contortus*. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD- standard deviation, SEM- standard error of the mean, N/D- not detected, N/A- non applicable, s- sample. For formulas and calculations of the mean, variance, SD and SEM see section ‘2.2.14’.

As seen in the Figure 51 and Table 11 the MALDI-TOF MS spectrum reveals the previously unreported presence of O-glycans in *H. contortus* together with some of their potential structural features. At this point based on the compositional analysis alone we can propose certain core structures likely present in the *H. contortus* O-glycome. The molecular ion at m/z 534 can potentially represent a core 1 structure Galβ1-3GalNAc. The molecular ion at m/z
575 could be a Core 6 structure GlcNAcβ1-6GalNAc. Based on the previously described schistosome-specific Galβ1–3(Galβ1–6)GalNAc ‘Shisto’ core (Huang et al., 2001) the molecular ions at \( m/z \) 738, 912, 942 could potentially have this type of Core structure. Abundant Core 2 Galβ1-3(GlcNAcβ1-6)GalNAc structure are also possible for the molecular ions at \( m/z \) 779, 953, 983, 1157, 1198, 1361, 1402 and 1606.

At this point it is important to emphasize that MALDI-TOF MS experiment provides us only with compositional information about the glycans but not their arrangement, linkage or conformational isomerism such as glucose, galactose or mannose. For this reason caution must be taken when proposing specific Core structures especially for structures with higher molecular weight.

### 3.3.2 GC-MS linkage analysis of partially methylated alditol acetates of O-glycans released by reductive elimination from adult *H. contortus*

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Assignment</th>
<th>Characteristic fragment ions</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.79</td>
<td>Terminal fucose</td>
<td>115,118,131,162,175</td>
<td>0.18</td>
</tr>
<tr>
<td>14.01</td>
<td>2-linked fucose</td>
<td>87,89,115,130,131,175,190</td>
<td>0.06</td>
</tr>
<tr>
<td>14.52</td>
<td>Terminal galactose</td>
<td>102,118,129,162,161,205</td>
<td>1.00</td>
</tr>
<tr>
<td>15.31</td>
<td>GalNAc-itol</td>
<td>88,89,130,174,218,262</td>
<td>0.64</td>
</tr>
<tr>
<td>16.43</td>
<td>3-linked GalNAc-itol</td>
<td>88,130,133,246,290</td>
<td>0.16</td>
</tr>
<tr>
<td>17.10</td>
<td>6-linked GalNAc-itol</td>
<td>88,101,117,130,161,174,218</td>
<td>0.01</td>
</tr>
<tr>
<td>17.73</td>
<td>Terminal GlcNAc</td>
<td>117,143,145,159,203,205</td>
<td>0.06</td>
</tr>
<tr>
<td>18.18</td>
<td>3,6-linked GalNAc-itol</td>
<td>88,130,161,246,318</td>
<td>0.11</td>
</tr>
<tr>
<td>18.23</td>
<td>Terminal GalNAc</td>
<td>117,143,145,159,203,205</td>
<td>0.08</td>
</tr>
<tr>
<td>18.60</td>
<td>4-linked GlcNAc</td>
<td>117,159,233</td>
<td>0.07</td>
</tr>
<tr>
<td>19.21</td>
<td>3-linked GalNAc</td>
<td>99,101,117,129,159,161</td>
<td>0.05</td>
</tr>
<tr>
<td>19.44</td>
<td>3,4-linked GlcNAc</td>
<td>117,159</td>
<td>0.10</td>
</tr>
</tbody>
</table>

**Table 12.** GC-MS analysis of partially methylated alditol acetates obtained from reductively eliminated *H. contortus* O-glycans. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.
The GC-MS linkage data analyses of partially methylated alditol acetates from the permethylated reductively eliminated O-glycans are summarized in the Table 12. The terminal galactose is consistent with the earlier proposed core 1, core 2 and Schisto core structures. The core 1 structure is also supported by the presence of the 3-linked GalNAc-itol Schisto core as well as core 2 structures are supported by the presence of the 3, 6-linked GalNAcitol. Terminal GlcNAc provides extra evidence for the core 2 structure as well as potential core 6 structure. The latter is also supported by the 6-linked GalNAc-itol.

3.3.3 MALDI-TOF-TOF analyses of permethylated, reductively eliminated lower molecular weight O-glycans

Figure 52 MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 534 adult H. contortus O-glycan. The blue arrowed horizontal line represents the loss of indicated structural fragment from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section ‘2.3.2) and knowledge of biosynthetic pathways in other organisms.
To collect more evidence for the presence of Galβ1-3GalNAc core 1 structure, the molecular ion at m/z 534 observed in the MALDI-MS spectrum (see Figure 51) was fragmented by MALDI-TOF-TOF as seen in the Figure 52. The most abundant fragment observed is at m/z 298 corresponding to the reducing end-HexNAc. The ‘Z’, not ‘Y’ type of this ion (Domon, Costello 1988) suggest it is 3-linked. The minor ‘Y’-type ion of the reducing-end HexNAc is observed at m/z 316. The fragment ion at m/z 259 corresponds to the non-reducing end hexose. The characteristic core-1 O-glycan fragmentation, together with the terminal galactose and 3-linked GalNAc-itol observed in the GC-MS experiment (see section 3.3.2) strongly suggest the m/z 534 molecular ion is a core 1 O-glycan.

Figure 53 MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 575 adult H. contortus O-glycan. The blue arrowed horizontal line represents the loss of the indicated structural fragment from the parental molecular ion.
To obtain more information about the HexNAc$_2$ structure, the molecular ion at $m/z$ 575 seen in the Figure 51, was fragmented by the MALDI-TOF-TOF instrumentation. The resultant spectrum can be seen in Figure 53. The most abundant fragment ion at $m/z$ 298 corresponds to the ‘Z’-type ion (Domon, Costello 1988) of the reducing-end HexNAc suggesting it is 3-linked. On the other hand the abundant ‘Y’-type ion (Domon, Costello 1988) at $m/z$ 316 may suggest the presence of the 6-linked reducing-end HexNAc. The fragment ions at $m/z$ 282 and 300 correspond to ‘B’- and ‘C’-type ions (Domon, Costello 1988) of the non-reducing HexNAc. Together with the terminal GlcNAc, terminal GalNAc, 3-linked GalNAc-itol, 6-linked GalNAc-itol observed in GC-MS experiment (see section 3.3.2) it may suggest the presence of GlcNAcβ1-3GalNAc core 3, GlcNAcβ1-6GalNAc core 6, GalNAcα1-3GalNAc core 5, GalNAcα1-6GalNAc core 7 structures.

Figure 54 MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated $m/z$ 738 adult *H. contortus* O-glycan. The proposed monosaccharide configurations are based on GC-MS data (see section 2.3.2) and knowledge of biosynthetic pathways in other organisms.
To confirm the presence of the Galβ1-3(Gal β1-6)-GalNAc Schisto core (Huang et al., 2001) the molecular ion at m/z 738 observed in MALDI-MS spectrum (see Figure 51) was fragmented by MALDI-TOF-TOF. The resultant spectrum as seen in the Figure 54 is characteristic of the proposed Schisto core structure (Huang et al., 2001). The most abundant fragment ion at m/z 502 corresponds to the ‘Z’-type ion (Domon, Costello 1988) of reducing-end Hex-HexNAc suggesting HexNAc is 3-linked. The minor ‘Y’-type ion (Domon, Costello 1988) of reducing-end Hex-HexNAc is observed at m/z 520. The fragment ion at m/z 259 corresponds to the non-reducing end hexose. Together with the terminal galactose and 3, 6-linked GalNAc-itol observed in GC-MS experiment (see section 3.3.2) these results confirm the earlier proposed Galβ1-3(Gal β1-6)-GalNAc Schisto core structure.

3.3.4 MALDI-QIT-TOF MS^n analyses of permethylated, reductively eliminated m/z 1606 O-glycan

![Figure 55](image)

**Figure 55** MALDI-QIT-TOF MS^2 spectrum of reductively eliminated m/z 1606 O-glycan. The blue horizontal arrowed lines represent the loss of indicated structural fragments from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section ‘2.3.2’), previously in the thesis described N-glycans data and knowledge of biosynthetic pathways in other organisms.
In order to get more structural information the highest molecular weight O-glycan at \( m/z \) 1606 was further fragmented by the MALDI-QIT-TOF MS\(^n\) approach. The MS\(^2\) analysis as seen in the Figure S5 interestingly revealed the presence of the Hex-Fuc structure seen earlier in the N-glycome of the parasite. This structure is supported by the fragment ions at \( m/z \) 433 and 1196. The absence of the fragment ion corresponding to the loss or elimination of the fucose residue alone suggests the fucose is internal rather than a terminal residue. The ‘Z-type’ ion (Domon, Costello 1988) at \( m/z \) 1196 indicates the elimination, not cleavage of the Hex-Fuc structure strongly suggesting linked to the 3-position of the GlcNAc. The most abundant fragment ion at \( m/z \) 1143 with fragment ions at \( m/z \) 486 and 504 indicate the presence of the terminal Hex-HexNAc structure. Other fragment ions assigned in the spectrum confirm the proposed modified core 2 where Hex-HexNAc and Hex-Fuc are attached to the core 2 GlcNAc residue.

The Gal-Fuc modifications of the chitobiose core and non-reducing LDNF antennae observed in the N-glycome of adult \( H. \) contortus together with the terminal galactose observed in the GC-MS linkage data of the adult O-glycans (see section 3.3.2) makes the hexose attached to the fucose likely to be a galactose residue. The proposed terminal Hex-GalNAc structure is based on the presence of the modified LDNF structure in the N-glycome of the parasite (although without the terminal hexose) and the 3-linked GalNAc residues in the GC-MS linkage data (see section 3.3.2).
To get more linkage information about the putative Hex-GalNAc structure the fragment ion at $m/z$ 486 was further fragmented by MALDI-QIT-TOF MS$^3$ as seen in the Figure 56. The fragment ions at $m/z$ 259, 241 and 268 confirm the non-reducing end terminal hexose. The cross-ring fragment ions at $m/z$ 384 and 412 generated by the reverse Diels-Alder mechanism strongly suggest the 3-linkage to HexNAc. Together with the 3-linked GalNAc observed in the GC-MS linkage data (see section 3.3.2) these results strongly support a Hex1-3GalNAc structure.
Figure 57 MALDI-QIT-TOF MS$^3$ spectrum of m/z 433 fragment ion from permethylated reductively eliminated m/z 1606 O-glycan. Dashed arrowed lines on the proposed structures represent glycosidic bond cleavages and cross-ring fragments.

To determine the linkage of the Hex-Fuc structure, the fragment ion at m/z 433 was further fragmented by the MALDI-QIT-TOF MS$^3$ instrumentation as seen in the Figure 57. The fragment ions at m/z 259, 241 and 215 confirm the non-reducing end hexose and the linked fucose. The most abundant cross-ring fragment ions at m/z 301 together with the cross-ring fragment ions at m/z 343 support the presence of the 2-linked fucose. In addition the spectrum looks remarkably similar to the MALDI-QIT-TOF MS$^3$ spectrum of the proposed Galα1-2Fuc structure found on the distal GlcNAc of the chitobiose core N-glycans and modified LDNF antennae (see Figures 47 and 41, respectively). The GC-MS data (see section 3.3.2) showed the presence of 2-linked fucose and the results seen in the Figure 57 are consistent with the Galα1-2Fuc structure. Finally, the proposed Galα1-2Fuc structures found in the N- and O-glycomes of adult H. contortus are attached to the 3-position of the GlcNAc in contrast to the Galβ1-4Fuc structure attached to the 6-position of the proximal GlcNAc of the N-glycan chitobiose core. These results suggest that the fucose 3-linked to the GlcNAc’s of the distal chitobiose core, LDNF antenna and core-2 O-glycan structures may act as substrates for the same putative α1-2 galactosyltransferase in adult H. contortus.
3.3.5 MALDI-TOF-TOF MS/MS analyses of permethylated reductively eliminated other higher molecular weight O-glycans

Figure 58 showing the MALDI-TOF-TOF MS/MS spectrum of reductively eliminated m/z 1402 O-glycan. The blue horizontal arrowed lines represent the loss of indicated structural fragments from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section ‘3.3.2’), previously in the thesis described N-glycans data and knowledge of biosynthetic pathways in other organisms.

In order to gain more insight into the structural features of the O-glycans, the abundant peak at m/z 1402 was fragmented by MALDI-TOF-TOF MS/MS as seen in the Figure 58. As well as for the previously described molecular ion at m/z 1606 (see Figure 55), the Hex-Fuc structure was present as indicated by the fragment ions at m/z 433, 992, 733. The MALDI-QIT-TOF MS3 analysis of the fragment ion at m/z 433 resulted in almost identical fragmentation (data not shown) to that obtained earlier for the fragment ion at m/z 433 originated from the molecular ion at m/z 1606 (see Figure 57) strongly suggesting the proposed Galα1-2Fuc structure is present in both structures. The fragment ions at m/z 282, 300, 1143 indicate the non-reducing end terminal HexNAc. The ‘Z-type’ (Domon, Costello 1988) nature of the fragment ion at m/z 992 indicates the elimination, not cleavage of the
Hex-Fuc structure strongly suggesting it is linked to the 3-position of the HexNAc. The presence of the terminal HexNAc and the Hex-Fuc structure linked to the 3-position of the HexNAc, together with the fragment ions at \( m/z \) 905, 733, 520 provide the evidence for the non-reducing end HexNAc-(Hex-Fuc1-3)-HexNAc structure. Combining this information with the terminal GalNAc and the 3, 4-linked GlcNAc observed in the GC-MS instrument (see section 3.3.2) we can propose a modified LDNF structure where 2-linked fucose is capped with galactose. Excitingly this Galα1-2Fuc- modified LDNF structure has been proposed for the N-glycome as well (see section 3.2.9) suggesting the potential common biosynthetic pathways decorating antennae of both N- and O-glycans *H. contortus*.

The molecular ion at \( m/z \) 1361 was also fragmented by the MALDI-TOF-TOF MS/MS approach as seen in the Figure 59. The most abundant fragment ion at \( m/z \) 951 together with fragment ions at \( m/z \) 433, 715 indicate the presence of the earlier described Hex-Fuc structure which is proposed to be Galα1-2Fuc (see Section 3.3.4, Figure 57). As seen in the

![Figure 59 MALDI-TOF-TOF MS/MS spectrum of reductively eliminated \( m/z \) 1361 O-glycan. The blue horizontal arrowed lines represent the loss of indicated structural fragments from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section '2.3.2'), previously in the thesis described N-glycans data and knowledge of biosynthetic pathways in other organisms.](image-url)
fragmentation of the molecular ions at m/z 1606 and 1402 (see Figures 55 and 58, respectively) the ‘Z-type’ (Domon, Costello 1988) nature of the fragment ion at m/z 951 indicates that the Hex-Fuc structure is 3-linked to the GlcNAc. The fragment ions at m/z 472, 520, 715, 864 together with the absence of the non-reducing end terminal HexNAc provide the evidence for the Hex-(Hex-Fuc1-3) HexNAc structure. The presence of the 3, 4-linked GlcNAc and terminal galactose seen in the GC-MS linkage data (see section 2.3.2) make this structure likely to be a modified LacNAc structure where GlcNAc is 3-linked to the proposed Galβ-2Fuc structure.

3.3.6 MALDI-TOF MS analysis of permethylated reductively eliminated, broad specificity α1-3, 4, 6 galactosidase digested O-glycans

To provide more evidence for the conformation and linkage of the hexoses, the reductively eliminated O-glycans were subject to Green Coffee Bean broad specificity α1-3, 4, 6 galactosidase digestion coupled with MALDI-TOF MS and MALDI-TOF-TOF MS/MS analyses. The resultant MALDI-TOF MS spectrum can be seen in Figure 60, bottom and the most obvious difference in comparison to the MS spectrum prior to the digest (see Figure 60, top) is the disappearance of the molecular ion at m/z 1606. Referring to the proposed structure of this molecular ion (see section 3.3.4) the two potential hexoses susceptible to the α1-3, 4, 6 galactosidase are those capping the fucose and the GalNAc of the proposed modified LDNF structure. The hexose attached to the reducing-end GalNAc is assumed to be the Galβ1-3 as consistent with the proposed Core 2 structure. The MALDI-TOF-TOF analyses of all other molecular ions after the digest (data not shown) did not reveal any structural changes in the molecular ions suggesting the resistance of the proposed Galα1-2Fuc structure to the α1-3, 4, 6 galactosidase. This was also observed in the N-glycome of H. contortus (see sections 3.2.16 and 3.2.17) and is most likely explained by the unusual presentation of the alpha galactose to the enzyme. On the other hand the absence of the new structural isomers in the molecular ion at m/z 1402 and virtually complete absence of the molecular ion at m/z 1606 provide the evidence for the non-reducing end GalNAc-capping hexose being alpha-linked to galactose. Together with the previous data from GC-MS (see section 3.3.2.) and fragment ion at m/z 486 MS³ analysis (see Figure 56), these results strongly suggest the modified LDNF Galα1-3GalNAc-(Galα1-2Fuc)GlcNAc- structure, which, to our knowledge, is not seen in any other organism.
Figure 60 MALDI-TOF MS spectra of permethylated reductively eliminated O-glycans before (top) and after (bottom) the broad specificity α-galactosidase digestion. The proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), GC-MS data (see section ‘2.3.2’), previously in the thesis described N- and O-glycans data and knowledge of biosynthetic pathways in other organisms.
3.3.7 MALDI-MS and MALDI-MS/MS analyses of permethylated reductively eliminated, broad specificity β-galactosidase and β1-4-galactosidase digested O-glycans

Adult *H. contortus* O-glycans were also subjected to the broad specificity β-galactosidase and β1-4-galactosidase digestions with the subsequent MALDI-TOF MS and MALDI-TOF-TOF MS/MS analyses however no structural changes were observed (data not shown).

3.3.8 The summary of reductively eliminated O-glycan structures from adult *H. contortus*

Putting all the data together acquired so far in the thesis about reductively eliminated O-glycans from adult *H. contortus*, the proposed structures with the corresponding m/z values can be seen below in the Table 13. Considering that O-glycans from *H. contortus* have not been reported previously, the proposed linkages for most of the structures have been included as well.

<table>
<thead>
<tr>
<th>m/z value</th>
<th>Proposed structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>534</td>
<td>β1-3</td>
</tr>
<tr>
<td>575</td>
<td>β1-3 β1-6</td>
</tr>
<tr>
<td>738</td>
<td>β1-3 β1-6</td>
</tr>
<tr>
<td>779</td>
<td>β1-3 β1-6</td>
</tr>
<tr>
<td>912</td>
<td>β1-3 β1-6</td>
</tr>
</tbody>
</table>
Table 13 showing the proposed O-glycan structures released by reductive elimination from adult *H. contortus*. 
3.4 Structural characterization of the N-glycome of L3 stage *H. contortus*

3.4.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from L3 stage *H. contortus*

In order to investigate whether the additional N-glycan structures observed in the adult stage are also present in the L3 stage, PNGase F released permethylated N-glycans were analysed by MALDI-TOF-MS. As illustrated in Figure 61, top the most abundant structures observed in MALDI-MS spectrum of 35% acetonitrile fraction are paucimannose structures with or without fucose and having the general composition of Hex$_2$Fuc$_0$HexNAc$_2$ (m/z 967-1375). A second group of structures corresponds to complex N-glycans having the general composition of Hex$_3$Fuc$_0$HexNAc$_3$-5. The majority of the N-glycans observed in 50% acetonitrile fraction (Figure 61, bottom) corresponds to Hex$_4$Fuc$_0$HexNAc$_2$. As with the adult these high mannose N-glycans have not been subject to MS/MS analysis due to the established characterisation of these structures. The molecular ions at m/z 1141, 1345, 1416, 1590, 1835 and 2081 were further characterised by the MALDI-TOF-TOF MS/MS approach.
Figure 61 MALDI-TOF MS spectra of permethylated 35% acetonitrile (top) and 50% acetonitrile (bottom) N-glycans released from L3 *H. contortus* by PNGase F enzyme. The structures proposed are based on the further MALDI-TOF-TOF MS/MS fragmentation (peaks marked with an asterisk*), GS-MS data (see section 3.4.2) and already obtained knowledge of N-glycan structures in adult parasite. The peak at $m/z$ 1655 seen in 50% acetonitrile fraction corresponds to contaminant.
3.4.2 GC-MS linkage analysis of L3 stage H. contortus N-glycans released by PNGase F (35% acetonitrile fraction)

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.93</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.35</td>
</tr>
<tr>
<td>18.45</td>
<td>102,118,129,162,161,205</td>
<td>Terminal mannose</td>
<td>1.00</td>
</tr>
<tr>
<td>19.60</td>
<td>129,130,161,190</td>
<td>2-linked mannose</td>
<td>0.46</td>
</tr>
<tr>
<td>21.16</td>
<td>129,130,189,190</td>
<td>2,6-linked mannose</td>
<td>0.11</td>
</tr>
<tr>
<td>21.33</td>
<td>118,129,189,234</td>
<td>3,6-linked mannose</td>
<td>0.65</td>
</tr>
<tr>
<td>22.25</td>
<td>117,159,205</td>
<td>Terminal GlcNAc</td>
<td>0.14</td>
</tr>
<tr>
<td>22.69</td>
<td>117,159,205</td>
<td>Terminal GalNAc</td>
<td>0.09</td>
</tr>
<tr>
<td>23.14</td>
<td>117,159,233</td>
<td>4-linked GlcNAc</td>
<td>0.50</td>
</tr>
<tr>
<td>23.97</td>
<td>117,159</td>
<td>3,4-linked GlcNAc</td>
<td>0.07</td>
</tr>
<tr>
<td>24.43</td>
<td>117,159,261</td>
<td>4,6-linked GlcNAc</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 14. GC-MS analysis of partially methylated alditol acetates obtained from L3 stage H. contortus N-glycans released by PNGase F. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.

The GC-MS linkage data of partially methylated alditol acetates from the PNGase F released N-glycans can be seen in the Table 14. The terminal, 3, 6-linked, 2-linked mannoses are consistent with high and truncated mannose structures present. The 2, 6-linked as well as 2-linked mannoses can be explained by the addition of the initiating GlcNAc antennae on the trimannosyl core. 4-linked GlcNAc is consistent with the proximal chitobiose core GlcNAc as well as the non-reducing end LDN antennae. The latter is also supported by the presence of the terminal GalNAc. The 3, 4-linked GlcNAc is consistent with the minor amounts of the distal GlcNAc fucosylation observed in the L3 stage parasite.
3.4.3 MALDI-MS/MS analyses of permethylated PNGase F released N-glycans from L3 stage *H. contortus*

Since the MALDI-TOF MS compositional analysis alone cannot reveal structural isomers present, the molecular ions have been further fragmented by the MALDI-TOF-TOF MS/MS analyses. The exemplary data is shown for the molecular ion at \( m/z \) 1345 in the Figure 62, **top**. The fragment ions at \( m/z \) 474 and 894 confirm the core fucosylation of the proximal GlcNAc at 6-position. The fragment ions at \( m/z \) 431 and 667 together with the absence of the fragment ion at \( m/z \) 463 corresponding to the linear Hex-Hex structure, all these data point to the three mannoses on the non-reducing end of the N-glycan arranged in the branched, not linear way. These results are supported by the presence of the 3, 6-linked mannose in the GC-MS linkage data (see Table 14). The \( '1X_5\)-type' (Domon, Costello 1988) cross-ring fragment ion shows the presence of the terminal fucose. Comparing this data to the MALDI-TOF-TOF MS/MS analysis of the same molecular ion from the adult parasite (see Figure 62, **bottom**), apart from the structural isomer described, the additional isomer with the proposed Gal-Fuc structure linked to the 3-position of the distal GlcNAc is also present as supported by the fragment ions at \( m/z \) 433, 953, 1068, 463. The fragmentation of other molecular ions (see the proposed structures in the MALDI-MS spectrum, Figure 61) showed the absence of the Gal-Fuc structures in L3 stage parasite.
Figure 62 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 1345 L3 infective stage *H. contortus* N-glycan (top). The MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 1345 adult *H. contortus* N-glycan (bottom) is given for comparison. The blue arrowed horizontal lines represent loss of indicated structural fragments from the parental molecular ion.
The MALDI-TOF-TOF spectrum of the molecular ion at $m/z$ 1141 can be seen in the Figure 63, revealing two structural isomers are present. The major structural isomer is fucosylated on the proximal GlcNAc as seen by the fragment ions at $m/z$ 474 and 690. The minor isomer is fucosylated on the distal GlcNAc and is represented by the fragment ions at $m/z$ 864 and 676 which correspond to the loss of reducing-end GlcNAc and an additional fucose, respectively. The presence of the fucose 3-linked to the distal GlcNAc is also supported by the ‘Z’-type ion (Domon, Costello 1988) at $m/z$ 935. Other fragment ions at $m/z$ 463, 719, 923, 953 can be ascribed to any of two isomers. It is important to note that previously reported data (Haslam et al., 1998) on *H. contortus* L3 stage N-glycans generated by FAB-MS instrumentation did not reveal the presence of distal GlcNAc fucosylation.

**Figure 63** MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released $m/z$ 1141 L3 infective stage *H. contortus* N-glycan. The blue arrowed horizontal lines represent loss of indicated structural fragments from the parental molecular ion. ‘M’ denotes major isomer and ‘m’ minor.
3.4.4 MALDI-MS analysis of permethylated PNGaseA released N-glycans from L3 stage H. contortus

![MALDI-TOF MS spectrum of permethylated 35% acetonitrile N-glycans released from L3 infective stage H. contortus parasite by PNGase A enzyme. The structures proposed are based on the further MALDI-TOF-TOF MS/MS fragmentation (peaks marked with an asterisk*), GC-MS data (see section 3.4.5) and already obtained knowledge of N-glycan structures in adult parasite. The 50% acetonitrile fraction has not been shown as no new structures apart from high-mannose have been detected.]

The N-glycans remaining on the glycopeptides after the PNGase F digestion were subject to the PNGase A digestion with the resultant MALDI-TOF MS spectrum seen in the Figure 64. The major N-glycan family observed corresponds to the truncated core-fucosylated structures with the general composition Hex$_{2,3}$Fuc$_{1,3}$HexNAc$_2$. The oligomannose structures with the general composition Hex$_3$HexNAc$_2$ as well as one complex truncated Hex$_3$HexNAc$_3$ structure suggest the incomplete removal of glycans by the PNGase F enzyme.
3.4.5 GC-MS linkage analysis of L3 *H. contortus* parasite N-glycans released by PNGase A (35% acetonitrile fraction)

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.93</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.14</td>
</tr>
<tr>
<td>18.41</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal mannose</td>
<td>0.82</td>
</tr>
<tr>
<td>19.56</td>
<td>129, 130, 161, 190</td>
<td>2-linked mannose</td>
<td>0.03</td>
</tr>
<tr>
<td>21.32</td>
<td>118, 129, 189, 234</td>
<td>3,6-linked mannose</td>
<td>0.07</td>
</tr>
<tr>
<td>22.23</td>
<td>117, 159, 205</td>
<td>Terminal GlcNAc</td>
<td>0.01</td>
</tr>
<tr>
<td>23.12</td>
<td>117, 159, 233</td>
<td>4-linked GlcNAc</td>
<td>0.07</td>
</tr>
<tr>
<td>23.98</td>
<td>117, 159</td>
<td>3,4-linked GlcNAc</td>
<td>0.02</td>
</tr>
<tr>
<td>24.43</td>
<td>117, 159 (no 261)</td>
<td>4,6-linked GlcNAc</td>
<td>traces</td>
</tr>
</tbody>
</table>

Table 15. GC-MS analysis of partially methylated alditol acetates obtained from L3 *H. contortus* N-glycans released by PNGase A. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.

The GC-MS linkage data of partially methylated alditol acetates from the PNGase A released N-glycans can be seen in the Table 15. The terminal fucose is consistent with the core-fucosylated structures observed in the MALDI-TOF MS spectrum (see Figure 61). The 3, 6-linked mannose is consistent with the branched trimannosyl part of the all N-glycans except those where only two mannoses are present such as in the molecular ions at *m/z* 1141, 1315, 1489. The 2-linked mannose suggests the attachment site of the non-reducing end terminal GlcNAc seen in the molecular ion at *m/z* 1416. 4-linked GlcNAc is consistent with the non-fucosylated distal GlcNAc observed in all of the molecular ions except at *m/z* 1489. The distal GlcNAc in the latter molecular ion is represented by the 3, 4-linked GlcNAc.
3.4.6 MALDI-MS/MS analyses of permethylated PNGase A released N-glycans from L3 stage H. contortus

![MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase A released m/z 1519 from L3 H. contortus N-glycan. The blue arrowed horizontal lines represent loss of indicated structural fragments from the parental molecular ion.](image)

In order to get more structural information the PNGase A released fucosylated molecular ions from the MALDI-TOF MS spectrum have been further fragmented by the MALDI-TOF-TOF MS/MS approach. The exemplary data are shown for the molecular ions at m/z 1519 in the Figure 65. The most abundant fragment ions at m/z 894 together with the fragment ions at m/z 648 indicate the di-fucosylation of the proximal GlcNAc. The fragment ions at m/z 667 correspond to the trimannosyl part of the N-glycan and the fragment ions at m/z 431 together with the absence of the fragment ions at m/z 463 provide the evidence for the branched, not linear, arrangement of three mannoses. In contrast to the MALDI-QIT-TOF spectrum of the PNGase A-released N-glycan molecular ions at m/z 1519 from adult H. contortus parasite as seen earlier in the Figure 46, in the L3 stage parasite the structural isomer with Gal-Fuc modification on the distal GlcNAc is absent. Together with the MALDI-TOF-TOF MS/MS fragmentations of the other PNGase A-released fucosylated structures (see the proposed
structures in the MALDI-TOF MS spectrum in the Figure 64) these results suggest the stage specificity of the Gal-Fuc chitobiose core modifications observed in adult *H. contortus*.

Figure 66 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase A released m/z 1489 from L3 *H. contortus* N-glycan. The blue arrowed horizontal lines represent loss of indicated structural fragments from the parental molecular ion.

No distal GlcNAc fucosylation has been observed in previous FAB-MS data (Haslam et al., 1998). However minor amounts have been detected by MALDI instrumentation in PNGase F released N-glycans (see Figure 63). To further confirm distal GlcNAc fucosylation the molecular ion at m/z 1489 observed in L3 stage *H. contortus* PNGase-A released N-glycans MALDI-MS spectrum (see Figure 64) has been fragmented as seen in the Figure 66. Fragment ions at m/z 648 and 864 correspond to di-fucosylated proximal GlcNAc and its loss from the molecular ion, respectively. The fragment ion at m/z 463 confirms the presence of two mannoses on the non-reducing end. The ‘Z’-type (Domon, Costello 1988) fragment ion at m/z 1283 corresponds to elimination of fucose from the 3-position of the GlcNAc. The fragment ion at m/z 676 corresponds to the loss of fucose and di-fucosylated proximal
GlcNAc. Overall the data confirms the tri-fucosylated N-glycan with two fucoses being attached to the proximal GlcNAc and the third fucose to the distal GlcNAc.
3.5 Structural characterization of the O-glycome of L3 stage *H. contortus*

3.5.1 MALDI-MS analysis of reductively eliminated, permethylated O-glycans from infective L3 stage *H. contortus* parasite

![MALDI-TOF MS spectrum of permethylated reductively eliminated O-glycans from L3 *H. contortus* (35% acetonitrile fraction). All of the proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), the knowledge of biosynthetic pathways in adult *H. contortus* and other organisms. ‘amu’ stands for atomic mass units. 50% acetonitrile fraction has not been shown as no new structures have been detected.](image)

**Figure 67** MALDI-TOF MS spectrum of permethylated reductively eliminated O-glycans from L3 *H. contortus* (35% acetonitrile fraction). All of the proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), the knowledge of biosynthetic pathways in adult *H. contortus* and other organisms. ‘amu’ stands for atomic mass units. 50% acetonitrile fraction has not been shown as no new structures have been detected.

To our knowledge there are no literature reports available for O-glycans of L3 *H. contortus*. For this reason the O-glycans from the glycopeptides remaining after the PNGase A digestion were reductively eliminated and the MALDI-TOF MS spectrum is seen in **Figure 67**. A simpler spectrum was obtained compared to the adult stage O-glycans (see **Figure 60, top**) with a lack of fucosylated structures. The major molecular ions at *m/z* 534, 575, and 738 suggest the presence of Core 1, HexNAc₂, and Schisto Core structures. In addition a family of putative O-glycans at *m/z* 628, 832, 926 corresponding to the potential Core 1 structure with an extra 94 atomic mass unit modification and the Schisto Core structure with one and two
The aforementioned modifications, respectively. A mass increment of 94 is consistent with a fully methylated phosphate group.

### 3.5.2 GC-MS linkage analysis of partially methylated alditol acetates obtained from L3 *H. contortus* O-glycans

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.93</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.03</td>
</tr>
<tr>
<td>18.42</td>
<td>102,118,129,162,161,205</td>
<td>Terminal glucose/mannose</td>
<td>0.13</td>
</tr>
<tr>
<td>18.71</td>
<td>102,118,129,162,161,205</td>
<td>Terminal galactose</td>
<td>0.46</td>
</tr>
<tr>
<td>19.70</td>
<td>88,89,130,174,218,262</td>
<td>GalNAc-itol</td>
<td>1.00</td>
</tr>
<tr>
<td>20.40</td>
<td>102, 118, 129, 189, 233</td>
<td>6-linked galactose</td>
<td>0.85</td>
</tr>
<tr>
<td>20.89</td>
<td>88,130,133,246,290</td>
<td>3-linked GalNAc-itol</td>
<td>0.13</td>
</tr>
<tr>
<td>21.69</td>
<td>88,117,130161,174,218</td>
<td>6-linked GalNAc-itol</td>
<td>0.17</td>
</tr>
<tr>
<td>22.70</td>
<td>117, 159, 205</td>
<td>Terminal GalNAc</td>
<td>0.41</td>
</tr>
<tr>
<td>23.23</td>
<td>117, 159, 233</td>
<td>4-linked GlcNAc</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 16. GC-MS analysis of partially methylated alditol acetates obtained from L3 infective stage *H. contortus* O-glycans released by reductive elimination. The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.

The GC-MS linkage analysis of partially methylated alditol acetates obtained from L3 O-glycans is seen in Table 16. The terminal galactose is consistent with the proposed Galβ1-3GalNAc Core 1 and the Galβ1-3(Galβ1-6)GalNAc Schisto Core structures. The first is also supported by the presence of the 3-linked GalNAc-itol. The presence of the terminal GalNAc together with the 3- and 6-linked GalNAc-itols may represent potential GalNAcα1-3GalNAc Core 5 and GalNAcβ1-6GalNAc Core 7 structures, respectively. The presence of the 6-linked galactose may suggest the position of the potential phosphorylation. The presence of the 4-linked GlcNAc together with the co-eluting glucose/mannose peak can be explained by reductive elimination of some residual mannose N-glycan structures and/or often observed glucose contamination. The presence of the terminal fucose is also most likely explained by reductive elimination of some residual N-glycans. The 3, 6-linked GalNAc-itol has not been
observed and it is most likely explained by the decreased detection sensitivity of later time-eluting monosaccharides usually observed in GC-MS linkage analysis as well as signal suppression effects of co-eluting contaminants. On the other hand the presence of 3, 6-linked GalNAc-itol is strongly suggested by the MALDI-TOF-TOF fragmentation of the molecular ion at \( m/z \) 738 as seen later in the Figure 70. It is interesting to note the presence of the most abundant peak corresponding to GalNAc-itol which acts as a substrate for other O-glycans. However, the presence of a single HexNAc was not observed in the MALDI-MS spectra (Figure 67) due to the low cut-off mass value of 500 \( m/z \) as below this value the spectrum is usually dominated by low MW contaminant ions, for example coming from the matrix.

3.5.3 MALDI-MS/MS analyses of reductively eliminated, permethylated O-glycans from L3 \textit{H. contortus}

![Figure 68 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated \( m/z \) 534 O-glycan from L3 stage \textit{H. contortus}. The blue arrowed horizontal; line represents the loss of depicted fragment from the parental molecular ion.](image-url)
To confirm the presence of Galβ1-3GalNAc core 1 structure the molecular ion at m/z 534 (see Figure 67) was fragmented by MALDI-TOF-TOF MS/MS instrumentation. The resultant spectrum seen in the Figure 68 demonstrates fragments typical for core 1 O-glycan. The most abundant fragment ion at m/z 298 corresponds to the ‘Z’-type (Domon, Costello 1988) ion of the reducing-end HexNAc suggesting it being 3-linked. The minor ‘Y-type’ ion (Domon, Costello 1988) for the reducing-end HexNAc is observed at m/z 316. The fragment ion at m/z 259 corresponds to the non-reducing end hexose. Together with the terminal galactose and 3-linked GalNAc-itol seen in GC-MS experiment (see section 3.5.2) these results strongly suggest the presence of Galβ1-3GalNAc core 1 O-glycan structure.

To obtain more structural information about the HexNAc2 core, the molecular ion at m/z 575 has been fragmented by MALDI-TOF-TOF as seen in the Figure 69. One of the most abundant ions at m/z 298 is a ‘Z-type’ ion (Domon, Costello 1988) which is favoured when the non-reducing end HexNAc is 3-linked to the reducing-end GalNAc. The corresponding
‘C-type’ ion (Domon, Costello 1988) is at m/z 300. Together with the 3-linked GalNAc-itol observed in the GC-MS data (see section 3.5.2), it provides strong evidence for the HexNAc1-3GalNAc-itol structure. On the other hand ‘the Y-type’ ion (Domon, Costello 1988) at m/z 316, together with the corresponding ‘B-type’ ion (Domon, Costello 1988) at m/z 282, suggests that terminal HexNAc may be 4- or 6-linked to the reducing-end GalNAc. The GC-MS data (see section 3.5.2) indicates the presence of the 6-linked GalNAc-itol and, together with the terminal GalNAc, suggests the possible GalNAcα1-3GalNAc core 5 and/or GalNAcβ1-6GalNAc core 7 structures.

![Figure 70](image-url) MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 738 O-glycan from L3 H. contortus

To confirm the presence of the proposed Galβ1-3(Galβ1-6)GalNAc Schisto Core structure, the molecular ion at m/z 738 has been further fragmented by the MALDI-TOF-TOF MS/MS approach (see Figure 70). The molecular ions at m/z 259 and 502 together with the absence of the fragment ion at m/z 463, indicate that both hexoses are linked to the reducing end GalNAc. The ‘Z-type’ ion for the HexHexNAc fragment at m/z 502 being more abundant that
the ‘Y-type’ ion (Domon, Costello 1988) at m/z 520, strongly suggests one of the hexoses being attached to the 3-position of the GalNAc. All these results are consistent with the proposed Galβ1-3(Galβ1-6)GalNAc Schisto Core structure.

![Figure 71](image_url)  
**Figure 71** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 832 O-glycan from L3 *H. contortus*. The blue arrowed horizontal line represents the loss of depicted fragment from the parental molecular ion.

To test whether the molecular ions at m/z 628, 832, 926 actually correspond to the sugars modified with proposed modification(s) of the 94 amu, the molecular ion at m/z 832 has been further fragmented by the MALDI-TOF-TOF approach as seen in the **Figure 71**. Indeed fragment ions at m/z 502 and 706 suggest the presence of the modified O-glycan molecular ion. The fragment ion at m/z 149 corresponds to the 94 amu modification ion alone and the fragment ion at m/z 353 is explained by the 94 amu-modified hexose fragment ion. The abundant fragment ion at m/z 502 suggests the modified hexose linked to the 3-position of the HexNAc and the fragment ion at m/z 596 to the putative 6-position. The most abundant fragment ion at m/z 706 corresponds to the elimination of the proposed modification. The most likely explanation for the 94 amu modification is phosphorylation of the sugar where
both hydroxyl groups of the phosphate have been methylated during the standard permethylation procedure whilst preparing the O-glycans for the MALDI analysis.

![Diagram](MS/MS of m/z 926 HC L3 O-glycan)

**Figure 72** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 926 O-glycan from L3 stage *H. contortus*. The blue arrowed horizontal lines represent the loss of depicted fragments from the parental molecular ion.

To investigate whether two phosphate groups seen in the molecular ion at m/z 926 are attached to one or two separate hexoses, the MALDI-TOF-TOF analysis has been performed as seen in the Figure 72. The presence of the fragment ions at m/z 353 and 596 indicate one phosphate being attached to one hexose residue. Together with the absence of the fragment ions at m/z 447, 502, 520 these results strongly suggest each hexose being phosphorylated by the one phosphate group.
3.6 Structural conclusions of *H. contortus* parasite glycomes

3.6.1. Structural conclusions of adult N-glycans

![Diagram](attachment:figure73.png)

*Figure 73:* proposed structural conclusions of novel N-glycan structures discovered in the *Haemonchus contortus* adult parasite.

In this thesis I report the discovery of novel Gal-Fuc chitobiose core modifications in adult *H. contortus* (see *Figure 73*) which were not observed previously using older FAB-MS instrumentation (Haslam et al., 1996). When this moiety is 6-linked to the proximal GlcNAc it has been unequivocally established as Galβ1-4Fuc. The Galβ1-4Fuc structure 6-linked to the proximal GlcNAc has been previously reported in *C. elegans* (Hanneman et al., 2006). The potential galactosyltransferase GALT-1 carrying out the galactosylation of 6-linked fucose has been found (Titz et al., 2009) and galectins LEC-6, LEC-10 from *C. elegans* as well as CGL2 from ink cap mushroom were shown to be able to bind this type of chitobiose core modification (Butschi et al., 2010; Maduzia et al., 2011; Takeuchi et al., 2008). The structure has also been found in octopus rhodopsin (Zhang et al., 1997), squid rhodopsin (Takahashi et al., 2003), *Ascaris suum* (Yan et al., 2012), *Oesophagostomum dentatum* (Yan et al., 2012), keyhole limpet haemocyanin (Wuhrer et al., 2004) and planaria *Dugesia*...
japonica (Paschinger et al., 2011). Interestingly in keyhole limpet haemocyanin one galactose can be further modified at the 4-position with another Galβ1 residue (Wuhrer et al., 2004) and in D. japonica it can be capped with methylated hexose (Paschinger et al., 2011). Also the opportunistic human pathogen Acanthamoeba also has a Hex-Fuc modification 6-linked to the proximal GlcNAc. However, the identity of hexose was proven to be mannose (Schiller et al., 2012). Very recently the Hex-Fuc modification on the 6-position of the proximal GlcNAc has been reported in H. contortus (Paschinger and Wilson, 2015). Based on the knowledge of C. elegans biosynthetic pathways the authors proposed this modification to be Gal-Fuc (Paschinger and Wilson, 2015) which is in full agreement with the data presented in my thesis.

The Hex-Fuc structure 3-linked to the distal GlcNAc was first observed in C. elegans by Hanneman and colleagues (Hanneman et al., 2006). Later the hex-2 and hex-3 genes coding for hexosaminidases, the enzymes removing non-reducing end GlcNAc were knocked out in C. elegans (Yan et al., 2012). This strain showed up-regulation of Hex-Fuc structure 3-linked to the distal GlcNAc (Yan et al., 2012). MS and MS/MS analysis coupled with α-galactosidase treatment showed it to be Galα1-2Fuc 3-linked to the distal GlcNAc (Yan et al., 2012). In our case Gal-Fuc 3-linked to the distal GlcNAc proved to be resistant not only to β1-4 specific galactosidase and broad specificity β-galactosidase but also to broad specificity α-galactosidase. On the other hand the presence of 2-linked fucose in GC-MS analysis as well as indicative cross-ring fragments of this structure suggest that Gal-Fuc 3-linked to the distal GlcNAc is most likely a Galα1-2Fuc structure. The inability of broad specificity α-galactosidase to remove the hexose can be explained by the unusual three-dimensional galactose presentation for the enzyme. In 2015 Wilson and colleagues reported the Galα1-2Fuc modification on the 3-position of the distal GlcNAc in H. contortus (Paschinger and Wilson, 2015). Interestingly they managed to show the partial digestion of galactose from this structure by green coffee bean α-galactosidase on the HPLC-purified N-glycan fraction (Paschinger and Wilson, 2015). Nevertheless the proportion of the digested N-glycan was very small (Paschinger and Wilson, 2015) confirming the earlier proposed difficulties for available exoglycosidases to act unusual nematode glycans.

Unexpectedly the Galα1-2Fuc structure was also observed on non-reducing terminus HexNAc₂ antenna where it is 3-linked to the innermost HexNAc thus forming to our knowledge previously unseen in any organisms modified LDNF structure where fucose is 2-linked to the capping hexose (likely galactose). The proposed LDNF structures are also
supported by the presence of the terminal GalNAc in the GC-MS data. The presence of LDNF structure in *H. contortus* has been suggested earlier based on vaccination studies in sheep (Geldhof et al., 2005; van Stijn et al., 2010b; Vervelde et al., 2003). Upon hydrofluoric acid treatment the poly-HexNAc antennae comprising up to 5 and potentially more HexNAc units were observed. Due to sample complexity and low abundance it is hard to make definitive conclusions, however the presence of up to five fucoses and Hex-Fuc modified LDN antennae prior to HF treatment and only fucose being present on the proximal GlcNAc after the treatment indicates the α1-3 hydrolysis of Hex-Fuc structures from the modified LDN antennae. The presence of poly-Hex-Fuc-(HexNAc)HexNAc antennae prior to HF treatment has not been confirmed, but not ruled out due to the low abundance of higher molecular weight peaks potentially explaining the nature of poly-HexNAc structures. On the other hand it could also be explained by the new family of N-glycans bearing PC moieties which have been reported in *H. contortus* using mass spectrometry (Paschinger and Wilson, 2015) and Western blot (Kooyman et al., 2009) approaches. Poly-GlcNAc chitobiose oligomers with capping GalNAc or GlcNAc structures have been observed in human *O. volvulus*, bovine *O. gibsoni* and rodent *A. vitae* parasites. However, the origin of *H. contortus* poly-HexNAc antennae as PC-bearing N-glycans and/or as poly-modified LDNF antennae remains an open question.

### 3.6.2. Structural conclusions of adult O-glycans

In this thesis data on *H. contortus* O-glycans has been presented for the first time (see Figure 74). The majority of adult O-glycans are Galβ1-3(GlcNAcβ1-6)GalNAc Core 2 structures with modifications on the GlcNAcβ1-6 branch. The GlcNAc can be fucosylated at the 3-position and then the galactose can be added to the likely 4-position of the GlcNAc forming Lewisx terminal structure. Gal α1-2 can also cap the fucose forming the previously unseen modified Lewisx structure. Instead of galactose, GalNAc can be added to the 4-position of the GlcNAc forming previously reported LDN structures (Vervelde et al., 2003) which can be further modified by galactose forming Gal α1-2Fuc structures. In addition galactose can cap the terminal GalNAc forming the Galα1-3GalNAc structures previously suggested by (van Stijn et al., 2010b) using antibody arrays technology. As previously described Galα1-2Fuc modifications have also been observed in the adult N-glycome where it was found on the 3-position of the distal chitobiose core GlcNAc and 3-position of the non-reducing end LDN.
antenna GlcNAc monosaccharide. Together these findings may suggest the presence of α1-2 galactosyltransferase acting on the fucoses attached to the 3-positions of GlcNAc’s. However, more than one α1-2 galactosyltransferase is not ruled out. Interestingly the closely related *C. elegans* has been reported to have over 20 putative α1-2 galactosyltransferases (Zheng et al., 2002) making it likely that more than one α1-2 galactosyltransferase is active in *H. contortus*. In addition the potential Galβ1-3GalNAc Core 1, GalNAcβ1-6GalNAc Core 6 and Schisto Core O-glycans (Khoo et al., 2001) have been observed as well in the *H. contortus* adult O-glycome although these type of O-glycans did not form extended families as seen for Core-2 structures.

![Diagram](image_url)

**Figure 74** proposed structural conclusions of O-glycan structures discovered in the *Haemonchus contortus* adult parasite
3.6.3. Structural conclusions of L3 stage N- and O-glycans

When comparing the L3 stage N- and O-glycomes with adult *H. contortus*, the diminished complexity of structures detected is consistent with transcriptomics studies of this parasite where down-regulation of many genes in L3 stage has been reported (Laing et al., 2013; Schwarz et al., 2013). High mannose structures, complex structures with terminal GlcNAc, proximal GlcNAc fucosylation at 3- and 6-positions and to a smaller extent fucosylation of distal GlcNAc at 3-position, these N-glycan structures were observed in L3 stage as well as adult parasites however no chitiobiose core or LDNF galactosylation has been observed in L3 based on MALDI experiments. The O-glycome of L3 revealed the presence of Galβ1-3GalNAc Core 1, potential GalNAcα1-3GalNAc Core 5 and GalNAcβ1-6GalNAc Core 7 as well as Schisto Core (Khoo et al., 2001) -containing O-glycans. On the other hand no adult Core 2 structures with Galα1-2Fuc modifications on 3-positions of GlcNAc’s in LacNAc, LDN, Galα1-3GalNAcβ1-4GlcNAc structures were observed. Interestingly the phosphorylation of the Schisto Core galactose(s) has been observed however whether it is phosphorylation of galactoses only or it forms a larger post-translational modification with part of it being lost during preparation procedures remains unknown. The phosphorylation of glycans in the form of phosphorylcholine residue is relatively well characterized post-translational modification happening in nematodes (Harnett and Harnett, 2001; Harnett et al., 2010) and recently suggested to be present in *H. contortus* (Kooymann et al., 2009; Paschinger and Wilson, 2015). The phosphorylated mannose has been previously observed in lysosome-targeted enzymes (Coutinho et al., 2012) and in α-dystroglycan protein (Yoshida-Moriguchi et al., 2010). The mannoses modified at 2- and 6-positions with phosphoethanolamine residues are known to be present in GPI-anchors (Paulick and Bertozzi, 2008). The phosphorylation at 3-position of GlcNAc residue in LacdiNAc-modified extracellular matrix glycoproteins has been also reported (Breloy et al., 2012). To our knowledge no phosphorylated galactoses have been reported to date.
Chapter 4

Structural characterization of

 Dictyocaulus viviparus N- and O-glycomes
4.1 Introduction

An introduction to *Dictyocaulus viviparus* is provided in Chapter 1. This section summarises significant information that places my thesis work in context. *D. viviparus* is an economically important parasite which infects cattle in the temperate regions across the world. It resides in the lungs causing bronchitis, loss of weight and occasional death of the animal bringing financial damage to the farmers. To control *D. viviparus* infection a nematode vaccine called ‘Huskvac’ or ‘Dictol’ is available based on attenuated L3 larvae. However it has many drawbacks such as the need for calf donors, instability, short shelf-life and batch-to-batch variations (Bain, 1999). Another way to control infection is by using anthelminth drugs. However, they also have major limitations such as an increase in parasite resistance (Kaplan, 2004; Taman and Azab, 2014). Early vaccination attempts with purified *D. viviparus* acetylcholinesterase showed promising results (McKeand et al., 1994a, b). However immunization attempts with recombinant acetoeholinesterase failed (Matthews et al., 2001). With the failure of the recombinant approach the need for cheap, reproducible and effective vaccines focused scientific attention onto the development of carbohydrate-based vaccines (Astronomo and Burton, 2010). Nevertheless one of the major bottlenecks in their development is the lack of well-defined carbohydrate structures.

The only structural studies on adult *D. viviparus* glycome were performed by Haslam and colleagues in 2000 using FAB-MS technologies (Haslam et al., 2000). The study revealed an abundance of Lewis\(^{x}\) epitopes in N-glycans as well as the presence of LacN\(\text{Ac}\), proximal Glc\(\text{Nac}\) fucosylation at the 6-position and high mannose structures (Haslam et al., 2000). No data have been reported for adult O-glycans. Neither has structural characterization been performed on L3 stage N- or O-glycans. As discussed in sections 1.3.4 and 1.3.5, with the advent of new mass-spectrometry technologies FAB has been replaced by MALDI ionisation method and new MALDI-TOF instruments have proven to be more sensitive, detecting glycans with higher \(m/z\) ratios and have lower signal to noise ratios (Babu et al., 2009b; North et al., 2010a) compared to older FAB-MS instruments (Fukuda et al., 1984). For this reason the aim of the studies described in this chapter was to revisit structural characterization of adult *D. viviparus* N-glycans using MALDI-MS instrumentation as well as for the first time to provide data for N-glycans of L3 stage and O-glycans of both L3 and adult stage parasites.
4.2 Structural characterization of the N-glycome of adult *D. viviparus*

4.2.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from adult *D. viviparus*

The PNGase-F released N-glycans from adult *D. viviparus* were analysed by MALDI-TOF MS to assess if any structures previously not detected by FAB-MS analysis (Haslam et al., 2000) could be found. The MALDI-TOF MS spectrum can be seen in the Figure 75. The major glycan families seen are high mannose structures with the general composition of man₅₋₉ GlcNAc₂, paucimannosidic Man₁₋₄ GlcNAc₂, truncated fucosylated Man₂₋₃FucGlcNAc₂ and complex bi-, tri- and tetra-antennary structures with the general...
composition of Hex$_3$-7Fuc$_0$-3 HexNAc$_3$-6. All these glycan families seen in the MALDI-TOF MS spectrum have been observed earlier in the FAB-MS analysis (Haslam et al., 2000).

4.2.2 MALDI-MS/MS analysis of permethylated PNGase F released N-glycans from adult *D. viviparus*

![Figure 76 MALDI-TOF-TOF MS/MS spectrum of permethylated PNGase F released m/z 3215 N-glycan. The blue horizontal arrowed lines represent the loss of depicted fragments from the parental molecular ion. The proposed monosaccharide configurations is fully consistent with the previous publication of (Haslam et al., 2000) and knowledge of biosynthetic pathways in other organisms.](image)

In order to investigate the presence of new possible structural isomers not observed previously by the FAB-MS instrumentation (Haslam et al., 2000) the molecular ions at m/z 1171, 1345, 1416, 1375, 1579, 1590, 1661, 1794, 1835, 1906, 2396, 3041, 3215 from the MALDI-TOF MS spectrum (see Figure 75) were further fragmented by MALDI-TOF-TOF. Exemplar data is shown for the molecular ion at m/z 3215 in Figure 76. The abundant fragment ions at m/z 2578 and 660 indicate the presence of the earlier observed (Haslam et al., 2000) Lewis$^x$ (Galβ1-4(Fucα1-3)GlcNAc) structure. The fragment ion at m/z 1941 originates from the loss of two Lewis$^x$ structures. The fragment ions at m/z 474 and 2764
indicate the fucosylation of the reducing-end GlcNAc. Overall the fragment ions suggest a core-fucosylated complex N-glycan with three Lewis^x antennae which was reported previously using FAB-MS instrumentation (Haslam et al., 2000). MALDI-TOF-TOF MS/MS analyses of the other molecular ion peaks (data not shown) are also consistent with the earlier reported FAB-MS derived structures (Haslam et al., 2000) and indicate no new N-glycan structures detected by ultrasensitive MALDI instrumentation.

4.2.3 MALDI-MS analysis of permethylated PNGase A released N-glycans from adult *D. viviparus*

Adult *D. viviparus* glycopeptided remaining after the PNGase F digestion were subjected to subsequent PNGase A digestion to test for the presence of N-glycans with 3-linked fucose on the reducing-end GlcNAc. The MALDI-TOF MS spectrum (data not shown) did not reveal any glycan peaks indicating the absence of this N-glycan population in adult *D. viviparus.*
4.3 Structural characterization of the O-glycome of adult *D. viviparus*

4.3.1 MALDI-TOF MS analysis of reductively eliminated, permethylated O-glycans from adult *D. viviparus*

To our knowledge no literature reports have been previously published for adult *D. viviparus* O-glycans. Therefore glycopeptides remaining after enzymatic removal of N-glycans were subject to reductive elimination with the resultant MALDI-TOF MS spectrum of the permethylated O-glycans shown in the [Figure 77](#) and Table 17. The most abundant peak at *m/z* 534 compositionally corresponds to the Galβ1-3GalNAc Core 1 structure. From the MALDI-MS compositional analysis we can propose core 1, Schisto core and core 2 structures for the molecular ions at *m/z* 534, 738 and 779. However to provide more evidence and collect more structural information about the molecular ions at *m/z* 575, 983, 1157 all of the aforementioned peaks have been fragmented by MALDI-TOF-TOF with the exemplar data shown for the peaks at *m/z* 534, 575, 738, 779, 1157 as seen in the section 4.3.3.

[Figure 77](#) MALDI-TOF MS spectrum of permethylated, 35% acetonitrile fraction O-glycans released from the adult *D. viviparus* adult stage whole worm extracts by reductive elimination (50% acetonitrile fraction is not shown as no new structures have been observed). The structures marked with asterisk* were later subject to MALDI-TOF-TOF analyses. It is important to note that MALDI-TOF-MS experiment provides compositional data only shown in Table 17 and the proposed structures are based on other experiments described further in the chapter.
Table 17 The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 77) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of adult *D. viviparus*. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD- standard deviation, SEM- standard error of the mean, N/D- not detected, N/A- non applicable, s- sample. For formulas and calculations of the mean, variance, SD and SEM see section ‘2.2.14’.

<table>
<thead>
<tr>
<th>m/z value</th>
<th>composition</th>
<th>RA s1</th>
<th>RA s2</th>
</tr>
</thead>
<tbody>
<tr>
<td>534</td>
<td>HexHexNAc</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>575</td>
<td>HexNAc$_2$</td>
<td>0.12</td>
<td>0.056</td>
</tr>
<tr>
<td>738</td>
<td>Hex$_2$HexNAc</td>
<td>0.20</td>
<td>0.029</td>
</tr>
<tr>
<td>779</td>
<td>Hex HexNAc$_2$</td>
<td>0.29</td>
<td>0.054</td>
</tr>
<tr>
<td>983</td>
<td>Hex$_2$HexNAc$_2$</td>
<td>0.20</td>
<td>0.025</td>
</tr>
<tr>
<td>1157</td>
<td>FucHex$_2$HexNAc$_2$</td>
<td>0.10</td>
<td>0.014</td>
</tr>
</tbody>
</table>

4.3.2 GC-MS linkage analysis of adult *D. viviparus* O-glycans released by β-reductive elimination

The GC-MS linkage data analyses for the reductively eliminated O-glycans are summarized in the Table 18. The presence of the terminal fucose, 3, 4-linked GlcNAc and terminal galactose are consistent with the Lewis$^\text{x}$ structure proposed for the molecular ion at m/z 1157. The 4-linked GlcNAc and earlier mentioned terminal galactose support the proposed LacNAc (Galβ1-4GlcNAc) structure seen in the molecular ion at m/z 983. The presence of the terminal GlcNAc may indicate that the molecular ion peak at m/z 779 is Core 2 O-glycan although the presence of the terminal GalNAc may indicate other configurational isomers. In the same way the molecular ion peak at m/z 575 may contain both terminal GlcNAc and terminal GalNAc. The linkage between two HexNAcs in the molecular ion at m/z 575 remains unknown since the only mono-substituted GalNAc-itol is 3-linked although the biosynthetically likely 6-linked GalNAc-itol was possibly not detected due to the low abundance of the molecular ion at m/z 575. The presence of both 3-and 6-linked GalNAc-itol is supported by the MALDI-TOF-TOF fragmentation of the molecular ion at m/z 575
discussed earlier. The 3-linked GalNAc-itol is consistent with the proposed Core 1 structure and the 3, 6-linked GalNAc-itol complies with the Core 2 and Schisto Core structures. Interestingly the most abundant peak corresponds to GalNAc-itol which is a precursor for the biosynthesis of other O-glycans observed. This monosaccharide was not detected by MALDI-MS due to the 500 m/z lower mass limit setting of the instrument, chosen because of high levels of contaminant matrix peaks in the lower mass region.

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.93</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.19</td>
</tr>
<tr>
<td>18.42</td>
<td>102,118,129,162,161,205</td>
<td>Terminal glucose/ mannose (likely contaminant)</td>
<td>0.37</td>
</tr>
<tr>
<td>18.71</td>
<td>102,118,129,162,161,205</td>
<td>Terminal galactose</td>
<td>0.64</td>
</tr>
<tr>
<td>19.68</td>
<td>88,130,174,218,262</td>
<td>GalNAc-itol</td>
<td>1.00</td>
</tr>
<tr>
<td>20.86</td>
<td>88,130,246,290</td>
<td>3-linked GalNAc-itol</td>
<td>0.59</td>
</tr>
<tr>
<td>22.24</td>
<td>117,159,205</td>
<td>Terminal GlcNAc</td>
<td>0.17</td>
</tr>
<tr>
<td>22.61</td>
<td>88,130,161,246,318</td>
<td>3,6-linked GalNAc-itol</td>
<td>0.08</td>
</tr>
<tr>
<td>22.69</td>
<td>117,159,205</td>
<td>Terminal GalNAc</td>
<td>0.12</td>
</tr>
<tr>
<td>23.13</td>
<td>117,159,233</td>
<td>4-linked GlcNAc</td>
<td>0.47</td>
</tr>
<tr>
<td>23.98</td>
<td>117,159,244,301</td>
<td>3,4-GlcNAc</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Table 18 GC-MS analysis of partially methylated alditol acetates obtained from adult *D. viviparus* O-glycans released by reductive elimination (35% acetonitrile fraction). The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.
4.3.3 MALDI-TOF-TOF MS/MS analysis of reductively eliminated, permethylated O-glycans from adult *D. viviparus*

![MALDI-TOF-TOF MS/MS spectrum](image)

**Figure 78** MALDI-TOF-TOF MS/MS spectrum of 35% acetonitrile permethylated reductively eliminated m/z 534 adult *D. viviparus* O-glycan. The blue arrowed horizontal line represents the loss of indicated structural fragment from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section 4.3.2) and knowledge of biosynthetic pathways in other organisms.

To confirm the presence of the Galβ1-3GalNAc core 1 structure, the molecular ion at m/z 534 observed in the MALDI-MS spectrum (see Figure 77) was fragmented by MALDI-TOF-TOF (see Figure 78). The most abundant fragment ion observed at m/z 298 corresponds to the reducing end ‘Z-type’ (Domon, Costello 1988) HexNAc ion suggesting it is 3-linked. The minor ‘Y’-type ion of the reducing-end HexNAc is observed at m/z 316. The fragment ion at m/z 259 corresponds to the non-reducing end hexose. Overall the spectrum is typical for core-1 O-glycan fragmentation pathways and, together with the terminal galactose and 3-linked GalNAc-itol observed in GC-MS experiment (see section 4.3.2), confirms the presence of the core 1 O-glycan.
Figure 79. MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 575 adult *D. viviparus* O-glycan. The proposed monosaccharide configurations are based on GC-MS data (see section 3.3.2) and knowledge of biosynthetic pathways in other organisms.

The MALDI-TOF-TOF fragmentation of the molecular ion at m/z 575 was performed (see Figure 79) to obtain some information on the linkage between the HexNAc residues. The presence of the Y-fragment ion (Domon, Costello 1988) at m/z 316 is consistent with the HexNAc attachment to the 6-position of GalNAc. On the other hand the presence of the most abundant Z-fragment ion (Domon, Costello 1988) at m/z 298 suggest the HexNAc is attached to the 3-position of the GalNAc. The fragment ions at m/z 282 and 300 correspond to B-type and C-type ions (Domon, Costello 1988) of the non-reducing end terminal HexNAc, respectively. This indicates the molecular ion at m/z 575 contains two isomers with the reducing end GalNAc being modified both at the 3- and 6- positions.
To confirm the Galβ1-3(Gal β1-6)-GalNAc Schisto core (Khoo et al., 2001) the molecular ion at m/z 738 was fragmented by MALDI-TOF-TOF instrumentation. The resultant spectrum seen in the Figure 80 is characteristic of the proposed Shisto core structure (Khoo et al., 2001). The most abundant fragment ion at m/z 502 is the ‘Z’-type ion (Domon, Costello 1988) of the reducing-end Hex-HexNAc suggesting HexNAc being 3-linked. The minor fragment ion at m/z 520 corresponds to the ‘Y’-type ion (Domon, Costello 1988) of reducing-end Hex-HexNAc. The fragment ion at m/z 259 represents the non-reducing end hexose. Together with the 3,6-linked GalNAc-itol and terminal galactose observed in GC-MS experiment (see section 4.3.2) these results confirm the earlier proposed Galβ1-3(Gal β1-6)-GalNAc Schisto core structure.
To get more structural information the molecular ion at \( m/z \) 779 has been further fragmented by MALDI-TOF-TOF instrumentation (see Figure 81). The most abundant fragment ion at \( m/z \) 543 results by eliminating hexose from the molecular ion and the Z-type fragment nature (Domon, Costello 1988) indicates it is attached to the 3-position of the HexNAc. The fragment ion at \( m/z \) 520 results from the loss of the non-reducing end HexNAc and the Y-type fragment nature (Domon, Costello 1988) suggests HexNAc is attached to the 6-position of the GalNAc. The fragment ions at \( m/z \) 259 and 282 correspond to terminal hexose and HexNAc, respectively. The fragment ion at \( m/z \) 284 corresponds to the internal reducing end GalNAc. Together with the terminal GlcNAc, terminal galactose and 3, 6-linked GalNAc-itol seen in GC-MS data (see section 4.3.2), the MALDI-TOF-TOF data are consistent with a GlcNAcβ1-6(Galβ1-3)GalANc Core 2 structure.
Figure 82 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 1157 adult *D. viviparus* O-glycan. The blue horizontal arrowed lines represent the loss of depicted fragments from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section 4.3.2) and knowledge of biosynthetic pathways in other organisms.

The fragment ions at m/z 660 compositionally correspond to a Lewis^5^ structure and the fragment ions at m/z 520 correspond to the loss of the Lewis^5^ structure. The more abundant ‘Z-type’ (Domon, Costello 1988) ions at m/z 951 in comparison to the ‘Y-type’ Domon, Costello 1988) ions at m/z 969 suggest more favoured elimination, not cleavage of the fucose residue indicating it being 3-linked. These results are consistent with the proposed Lewis^5^ structure. The most abundant ‘Z-type’ (Domon, Costello 1988) fragment ion at m/z 921 suggests the hexose is 3-linked which is consistent with the Core 2 structure having galactose attached to the 3-position of the reducing-end GalNAc. The eliminated galactose is seen as the fragment ion at m/z 259 although it can also originate from the Lewis^5^ part of the molecular ion. Overall the fragment ions support the proposed modified Core 2 O-glycan where the GlcNAc branch forms the Lewis^5^ structure which is supported by the 3,4-GlcNAc seen in GC-MS data (see section 4.3.2) and is also very abundant in the N-glycome of adult *D. viviparus* (see previous Figure 75).
4.3.4 The summary of reductively eliminated O-glycan structures from adult *D. viviparus*

Putting all the data together acquired so far in the thesis about reductively eliminated O-glycans from adult *D. viviparus*, the proposed structures with the corresponding *m/z* values can be seen below in Table 19.

<table>
<thead>
<tr>
<th><em>m/z</em> value</th>
<th>Proposed structurec</th>
</tr>
</thead>
<tbody>
<tr>
<td>534</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>575</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>738</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>779</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>983</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>1157</td>
<td><img src="image" alt="Structure" /></td>
</tr>
</tbody>
</table>

Table 19 showing the proposed O-glycan structures released by reductive elimination from adult *D. viviparus*. 
4.4 Structural characterization of the N-glycome of L3 stage *D. viviparus*

4.4.1 MALDI-MS analysis of permethylated PNGase F released N-glycans from L3 stage *D. viviparus*

![MALDI-MS spectrum](image)

**Figure 83** MALDI-TOF MS spectrum of permethylated 50% acetonitrile N-glycans released from *D. viviparus* L3 stage whole worm extracts by PNGase F (35% acetonitrile fraction was not shown as no new structures were detected). The structures proposed are based on the further MALDI-TOF-TOF MS/MS fragmentation (peaks marked with an asterisk*), previous knowledge about the adult N-glycome (Haslam et al., 2000) and knowledge of biosynthetic pathways in other organisms.

There are no literature reports on structural data of L3 *D. viviparus* N-glycans. The PNGase F released permethylated N-glycans were analysed by the MALDI-TOF MS analysis as seen in **Figure 83**. Abundant paucimannosidic structures with or without fucose and with the general composition of Man$_1$$_2$Fuc$_0$$_1$GlcNAc$_2$ are observed. The other relatively abundant structures are high mannose Man$_3$$_6$GlcNAc$_2$ structures. In addition complex structures having the composition of Man$_3$Fuc$_0$$_4$GlcNAc$_3$$_5$ are also relatively abundant in the L3 stage. In terms of the higher molecular weight complex structures in comparison with the adult parasite, these are more minor components suggesting stage specific differential expression in *D. viviparus*. 
### 4.4.2 GC-MS linkage analysis of L3 *D. viviparus* N-glycans released by PNGase F

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.96</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.12</td>
</tr>
<tr>
<td>18.44</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal mannose</td>
<td>1.00</td>
</tr>
<tr>
<td>18.70</td>
<td>102,118,129,162,161, 205</td>
<td>Terminal galactose</td>
<td>0.09</td>
</tr>
<tr>
<td>19.59</td>
<td>129, 130, 161, 190</td>
<td>2-linked mannose</td>
<td>0.84</td>
</tr>
<tr>
<td>20.78</td>
<td>113, 129, 130, 190, 233</td>
<td>2,4-linked mannose</td>
<td>0.11</td>
</tr>
<tr>
<td>21.17</td>
<td>129, 130, 189, 190</td>
<td>2,6-linked mannose</td>
<td>0.12</td>
</tr>
<tr>
<td>21.33</td>
<td>118, 129, 189, 234</td>
<td>3,6-linked mannose</td>
<td>0.50</td>
</tr>
<tr>
<td>22.26</td>
<td>117, 159, 205</td>
<td>Terminal GlcNAc</td>
<td>0.05</td>
</tr>
<tr>
<td>23.12</td>
<td>117, 159, 233</td>
<td>4-linked GlcNAc</td>
<td>0.32</td>
</tr>
<tr>
<td>23.98</td>
<td>117, 159</td>
<td>3,4-linked GlcNAc</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Table 20** GC-MS analysis of partially methylated alditol acetates obtained from L3 stage *D. viviparus* N-glycans released by PNGase F (35% acetonitrile fraction). The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.

The GC-MS linkage data analyses from the PNGase F released permethylated N-glycans are summarized in **Table 20**. The terminal, 2 and 3, 6-linked mannoses are consistent with the most abundant paucimannosidic and high mannose class of N-glycans in the L3 stage. The terminal GlcNAc is consistent with the initiation of antennae later becoming LacNAc and Lewis^x^ structures. Terminal galactose and 4-linked GlcNAc are consistent with the LacNAc structure although the latter is also consistent with the distal GlcNAc of the chitobiose core. Low abundance of 3, 4-linked GlcNAc shows the Lewis^x^ antennae as minor structures in the L3 stage in comparison to adult *D. viviparus* (see **Figure 75**). The terminal fucose is consistent with the proximal GlcNAc fucosylation and minor Lewis^x^ structures in the sample.
4.4.3 MALDI-TOF-TOF MS/MS analysis of permethylated PNGase F released N-glycans from L3 stage *D. viviparus*

**Figure 84** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 2694 L3 stage *D. viviparus* N-glycan. The blue horizontal arrowed lines represent the loss of depicted fragments from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section 4.4.2), previous knowledge about adult N-glycome (Haslam et al., 2000) and knowledge of biosynthetic pathways in other organisms.

To get more structural information the molecular ions at m/z 1345, 1416, 1590, 1661, 1783, 1906, 2080, 2693 from the PNGase F released permethylated N-glycans were subjected to MALDI-TOF-TOF analyses and the exemplar spectrum of the ion at m/z 2694 is shown in the **Figure 84**. The fragment ions at m/z 486, 2231, 1767 suggest the presence of multiple LacNAc units in the N-glycan. The absence of the fragment ions corresponding to one antenna having more than one LacNAc unit strongly suggests the presence of three separate LacNAc units rather than elongated antennas. The fragment ions at m/z 474, 2243 indicate fucosylation of the reducing-end GlcNAc.
4.4.4 MALDI-MS analysis of permethylated PNGase A released N-glycans from L3 stage D. viviparus

The L3 stage D. viviparus glycopeptides remaining after PNGase F digestion were subjected to subsequent PNGase A digestion to test for the presence of any N-glycans with 3-linked fucose on the reducing-end GlcNAc. The MALDI-TOF MS spectrum (data not shown) did not reveal any glycan peaks indicating the absence of this N-glycan population in D. viviparus L3.
4.5 Structural characterization of the O-glycome of L3 stage *D. viviparus*

4.5.1 MALDI-MS analysis of permethylated reductively eliminated O-glycans from L3 stage *D. viviparus*

The O-glycans remaining on *D. viviparus* L3 stage glycopeptides after PNGase F digestion were released by reductive elimination and the MALDI-TOF MS spectrum obtained can be seen in the **Figure 85**. In contrast to adult *D. viviparus* parasite where the O-glycans contain mostly Core 1, Core 2, and Schisto Core structures, the compositional MALDI-TOF MS O-glycan profiling of the L3 stage parasite revealed the presence of a single potential Core 1 structure at \( m/z \) 534 and a large family of hexose rich O-glycans bearing the likely Galβ1-3-(Galβ1-6)-GalNAc Schisto Core structure. The O-glycans represented by the molecular ion peaks at \( m/z \) 912, 1116, 1321, 1525 also contain one fucose residue. The other O-glycans represented by the molecular ion peaks at \( m/z \) 738, 942, 1146, 1351, 1555, 1579, 1963, 2167 have the molecular composition of HexNAc3Hex3-9.

![Figure 85](image-url)  
**Figure 85** MALDI-TOF MS spectrum of permethylated reductively eliminated 35% acetonitrile fraction O-glycans from *D. viviparus* L3 stage parasite (50% acetonitrile fraction is not shown as no new O-glycans were observed). The peaks marked with asterisk ‘*’ have been subjected to MALDI-TOF-TOF analyses at later stages. It is important to note that MALDI-TOF-MS experiment provides compositional data only shown in Table 21 and the proposed structures are based on other experiments described further in the chapter.
Table 21. The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 85) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of L3 stage *D. viviparus*. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD- standard deviation, SEM- standard error of the mean, N/D- not detected, N/A- non applicable, s- sample. For formulas and calculations of the mean, variance, SD and SEM see section ‘2.2.14’.

<table>
<thead>
<tr>
<th>m/z value</th>
<th>Composition</th>
<th>RA s1</th>
<th>RA s2</th>
<th>RA s3</th>
<th>Mean</th>
<th>Variance</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>534</td>
<td>HexHexNAc</td>
<td>0.55</td>
<td>0.40</td>
<td>0.53</td>
<td>0.49</td>
<td>0.0044</td>
<td>0.066</td>
<td>0.038</td>
</tr>
<tr>
<td>738</td>
<td>Hex$_2$HexNAc</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>912</td>
<td>FucHex$_2$HexNAc</td>
<td>0.059</td>
<td>0.026</td>
<td>0.018</td>
<td>0.034</td>
<td>0.00031</td>
<td>0.018</td>
<td>0.010</td>
</tr>
<tr>
<td>942</td>
<td>Hex$_3$HexNAc</td>
<td>0.36</td>
<td>0.14</td>
<td>0.10</td>
<td>0.20</td>
<td>0.013</td>
<td>0.11</td>
<td>0.066</td>
</tr>
<tr>
<td>1116</td>
<td>FucHex$_3$HexNAc</td>
<td>0.16</td>
<td>0.047</td>
<td>0.018</td>
<td>0.075</td>
<td>0.0038</td>
<td>0.061</td>
<td>0.035</td>
</tr>
<tr>
<td>1146</td>
<td>Hex$_4$HexNAc</td>
<td>0.11</td>
<td>0.032</td>
<td>0.019</td>
<td>0.054</td>
<td>0.0016</td>
<td>0.040</td>
<td>0.023</td>
</tr>
<tr>
<td>1321</td>
<td>FucHex$_4$HexNAc</td>
<td>0.0078</td>
<td>0.0035</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>1351</td>
<td>Hex$_5$HexNAc</td>
<td>0.048</td>
<td>0.0092</td>
<td>0.0016</td>
<td>0.020</td>
<td>0.00041</td>
<td>0.020</td>
<td>0.012</td>
</tr>
<tr>
<td>1525</td>
<td>FucHex$_5$HexNAc</td>
<td>0.0038</td>
<td>N/D</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>1555</td>
<td>Hex$_6$HexNAc</td>
<td>0.018</td>
<td>N/D</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>1759</td>
<td>Hex$_7$HexNAc</td>
<td>0.0067</td>
<td>N/D</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>1963</td>
<td>Hex$_8$HexNAc</td>
<td>0.0031</td>
<td>N/D</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2167</td>
<td>Hex$_9$HexNAc</td>
<td>0.00090</td>
<td>N/D</td>
<td>N/D</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 21. The assignment and relative quantification of molecular ions observed in MALDI-TOF-MS spectrum (see Figure 85) of permethylated (35% acetonitrile fraction), reductively eliminated O-glycans of L3 stage *D. viviparus*. The relative abundance (RA) was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak. RA-relative abundance, SD- standard deviation, SEM- standard error of the mean, N/D- not detected, N/A- non applicable, s- sample. For formulas and calculations of the mean, variance, SD and SEM see section ‘2.2.14’.
4.5.2 GC-MS linkage analysis of partially permethylated alditol acetates of O-glycans released by reductive elimination from L3 stage *D. viviparus*

The GC-MS linkage analysis was performed on partially methylated alditol acetates of the O-glycans released by reductive elimination from L3 stage *D. viviparus*. The presence of the 3, 6-linked GalNAc-itol is consistent with the earlier proposed Galβ1-3-(Galβ1-6)-GalNAc Schisto Core O-glycan structure. 3-linked GalNAc-itol is consistent with the Galβ1-3GalNAc Core 1 O-glycan structure proposed on the O-glycans MALDI-MS spectrum. The most abundant peak of terminal galactose together with the 2-linked, 3-linked, 6-linked, 2, 3-linked, 3, 6-linked galactoses suggest the potential configurations of the hexoses.

<table>
<thead>
<tr>
<th>Elution time</th>
<th>Characteristic fragment ions</th>
<th>Assignment</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.78</td>
<td>115,118,131,162,175</td>
<td>Terminal fucose</td>
<td>0.015</td>
</tr>
<tr>
<td>14.53</td>
<td>102,118,129,162,161,205</td>
<td>Terminal galactose</td>
<td>1.000</td>
</tr>
<tr>
<td>15.30</td>
<td>88,89,130,174,218,262</td>
<td>GalNAc-itol</td>
<td>0.077</td>
</tr>
<tr>
<td>15.61</td>
<td>129,130,161,190</td>
<td>2-linked galactose</td>
<td>0.006</td>
</tr>
<tr>
<td>15.66</td>
<td>118,129,161,234</td>
<td>3-linked galactose</td>
<td>0.007</td>
</tr>
<tr>
<td>16.15</td>
<td>102,118,129,162,189,233</td>
<td>6-linked galactose</td>
<td>0.045</td>
</tr>
<tr>
<td>16.42</td>
<td>88,130,133,246,290</td>
<td>3-linked GalNAc-itol</td>
<td>0.027</td>
</tr>
<tr>
<td>16.47</td>
<td>129,161,262</td>
<td>2,3-linked galactose</td>
<td>trace</td>
</tr>
<tr>
<td>17.26</td>
<td>118,129,189,234</td>
<td>3,6-linked galactose</td>
<td>0.004</td>
</tr>
<tr>
<td>18.18</td>
<td>118,129,189,234</td>
<td>3,6-linked GalNAc-itol</td>
<td>0.050</td>
</tr>
</tbody>
</table>

Table 22. GC-MS analysis of partially methylated alditol acetates obtained from L3 stage *D. viviparus* O-glycans released by reductive elimination (35% acetonitrile fraction). The relative abundance was calculated by measuring the total ion count (TIC) of each peak and dividing the value by the TIC of the most abundant peak.
4.5.3 MALDI-MS/MS analyses of permethylated reductively eliminated $m/z$ 534 and 738 O-glycans

Figure 86 MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated $m/z$ 534 L3 stage *D. viviparus* O-glycan. The blue horizontal arrowed line represents the loss of depicted fragment from the parental molecular ion. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2) and knowledge of biosynthetic pathways in other organisms.

To confirm the presence of the Galβ1-3GalNAc core 1 structure, the molecular ion at $m/z$ 534 observed in the MALDI-MS spectrum (see Figure 85) was fragmented by MALDI-TOF-TOF as seen in the Figure 86. The spectrum shows the classical fragmentation of a core 1 O-glycan. The most abundant fragment ion observed at $m/z$ 298 corresponds to the reducing end ‘Z-type’ (Domon, Costello 1988) HexNAc ion suggesting it is 3-linked. The minor ‘Y’-type ion of the reducing-end HexNAc is observed at $m/z$ 316. The fragment ion at $m/z$ 259 corresponds to the non-reducing end hexose. Together with the terminal galactose and 3-linked GalNAc-itol seen in the GC-MS experiment (see section 4.5.2), the data confirm the presence of the core 1 O-glycan.
To demonstrate the presence of the Galβ1-3(Gal β1-6)-GalNAc Schisto Core (Khoo et al., 2001) the molecular ion at m/z 738 observed in the MALDI-MS spectrum (see Figure 85) has been fragmented by MALDI-TOF-TOF as seen in Figure 87. The most abundant fragment ion at m/z 502 corresponds to the ‘Z’-type ion (Domon, Costello 1988) of reducing-end Hex-HexNAc suggesting HexNAc being 3-linked. The minor ‘Y’-type ion (Domon, Costello 1988) of reducing-end Hex-HexNAc is observed at m/z 520. The fragment ion at m/z 259 corresponds to the non-reducing end hexose. Absence of a fragment ion at m/z 463 corresponding to Hex-Hex, or fragment ions at m/z 298 and/or 316, corresponding to ‘Z’- and ‘Y’-type ions (Domon, Costello 1988), excludes the possibility of two hexoses being linked. Together with the terminal galactose and 3, 6-linked GalNAc-itol observed in the GC-MS experiment (see section 4.5.2) these results confirm the earlier proposed Galβ1-3(Gal β1-6)-GalNAc Schisto core structure.
4.5.4 MALDI-MS\textsuperscript{n} analyses of permethylated reductively eliminated m/z 1116 O-glycan

Additional levels of the structural information were obtained by selecting molecular ion peaks for a MALDI-QIT-TOF MS\textsuperscript{n} fragmentation. The MS\textsuperscript{2} fragmentation of the molecular ion peak at m/z 1116 was performed (see Figure 88) to obtain information about the position of the fucose residue within the O-glycan. The major fragmentation peaks at m/z 502 and 637 indicate two hexoses and a fucose moiety being 3-linked to the GalNAc. The 3-position was concluded based on the nature of the m/z 502 ion fragment which is Z-, rather than Y-ion (Domon, Costello 1988).
Figure 89 MALDI-QIT-TOF MS³ spectrum of m/z 637 fragment ions from permethylated reductively eliminated m/z 1116 O-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2) and knowledge of biosynthetic pathways in other organisms.

To provide additional structural confirmation the fragment ion at m/z 637 was subjected to MS³ analysis as seen in the Figure 89. The presence of the fragment ions at m/z 259 and 419 corresponds to the non-reducing end hexose and its loss. Together with the absence of the Fuc-Hex fragment ion at m/z 433, this demonstrates that the fucose is attached to the innermost galactose rather than terminal galactose. The cross-ring fragment ions at m/z 271, 389 and 359 could indicate fucose being 2-linked. The cross-ring fragment ions at m/z 315 and 345 can be explained by the galactose being 3-linked to the other galactose. The cross-ring fragments at m/z 489 and 519 are in accordance with the proposed Gal1-3(Fucα1-2)-Gal-OH fragment ion which is also supported by the earlier presented GC-MS data (see section 4.5.2) providing the evidence for the 2,3-linked galactose. It has to be noted that the GC-MS linkage data (see section 4.5.2) also provided evidence for the 3, 6-linked galactose.

Nevertheless, the large combination of possible cross-ring fragments generated by the MSⁿ analysis of the trisaccharide, in contrast to the disaccharide, makes these linkage assignments still tentative. In addition the cross-ring fragment ions at m/z 301, 359, 329 and 331 may
suggest a second structural isomer with the galactose being attached not to 3-, but 6-position. The absence of the 2, 6-linked galactose in the GC-MS experiment may be explained by the sensitivity limits of the instrument.

4.5.5 MALDI-MS/MS analysis of permethylated reductively eliminated m/z 1146 O-glycan

To get more structural information about the poly-hexose family of O-glycans with the general composition of Hex$_3$HexNAc$_1$ one of its members giving the molecular ion at m/z 1146, with a composition of Hex$_4$HexNAc$_1$, was subjected to MALDI-TOF-TOF MS/MS fragmentation as seen in the Figure 90. The fragment ion at m/z 910 corresponds to the Z-type fragment ion (Domon, Costello 1988) which has lost a hexose from the 3-position of the GalNAc suggesting the structural isomer with three hexoses comprising the 6-linked branch and one galactose at the 3-position of the GalNAc. This structural isomer is also supported by
the fragment ion at m/z 520 resulting from the loss of three hexoses from the 6-linked branch. The fragment ion at m/z 259 can be explained by any non-reducing end terminal hexose. The fragment ions at m/z 706 and 463 can be explained by the structural isomer bearing two hexoses at both the 6-linked and the 3-linked branches of the Schisto Core O-glycan. It has to be noted that the fragment ion at m/z 463 can also originate from other structural isomers. Finally the fragment ions at m/z 502 and 667 correspond to the structural isomer bearing three hexoses on the 3-linked branch of the Schisto Core and one galactose on the 6-linked branch, although the fragment ion at m/z 667 can be generated from the structural isomer bearing three hexoses on the 6-linked branch. It needs to be emphasized that the linear or branched nature of the 3-linked and the 6-linked branches comprised of three hexoses at this stage remains ambiguous. The data therefore indicate the presence of several potential structural isomers.

4.5.6 MALDI-MS/MS analyses of other permethylated reductively eliminated O-glycans

**Figure 91** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated m/z 1524 O-glycan from the *D. viviparum* infective L3 stage parasite. ‘M’ denotes the major isomer and ‘m’ minor. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
The molecular ion at \( m/z \) 1320 from MALDI-MS spectrum (see Figure 85) has been further fragmented by the MALDI-TOF-TOF MS/MS instrumentation as seen in the Figure 91. Major fragment ions at \( m/z \) 637 and 706 are observed corresponding to the Hex-Hex-Fuc fragment and its loss from the 3-position of the GalNAc. The fragment ions at \( m/z \) 259 correspond to the terminal hexose ion. The fragment at \( m/z \) 415 corresponds to the internal Hex-Fuc structure. The fragment at \( m/z \) 463 corresponds to Hex-Hex structure. All these fragments originate from the major structural isomer present depicted in the Figure 91, left (denoted ‘M’). The fragment ions at \( m/z \) 841 and 502 correspond to a Hex1Fuc structure and its loss from the 3-position of the HexNAc and indicate the presence of the minor isomer which has three hexoses and fucose on the 3-branch of the GalNAc as seen in the Figure 91, right (denoted ‘m’). The MALDI-TOF-TOF fragmentation of another fucosylated molecular ion at \( m/z \) 1524 resulted in a spectrum of low quality (data not shown). Nonetheless two major fragment ions at \( m/z \) 637 and 910 suggested an analogous structure to that seen in Figure 91, left (denoted ‘M’), but with a total of three, not two, hexoses on the 6-linked branch.

**Figure 92** MALDI-TOF-TOF MS/MS spectrum of permethylated reductively eliminated \( m/z \) 1759 O-glycan from the L3 stage *D. viviparus*. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
The MALDI-TOF-TOF fragmentation of the molecular ion at \( m/z \) 1759 (see Figure 92) showed a similar fragmentation pattern to that observed for the molecular ion at \( m/z \) 1146 (see Figure 90). Mixtures of isomers are likely to be present with both 3-and 6-linked branches of the Schisto Core containing up to six hexoses. The MS/MS fragmentation of the molecular ion at \( m/z \) 1963 with a composition Hex\textsubscript{6}HexNAc, although of low quality (data not shown), also suggested the similar pattern with seven potential structural isomers bearing from one to seven hexoses on both 3-and 6-linked branches. Due to its extremely low abundance, the MS/MS fragmentation of the highest molecular weight ion at \( m/z \) 2167 was not performed.

### 4.5.7 MALDI-MS and MALDI-MS/MS analyses of permethylated reductively eliminated, broad specificity β-galactosidase digested O-glycans

To try to assign a higher level of structural definition to the hexose rich O-glycans, the reductively eliminated O-glycans of *D. viviparus* L3 stage were subjected to bovine testes broad specificity β\textsubscript{1-3,4,6}-galactosidase digestion as seen in the Figure 93. It has to be noted for the rest of the chapter that enzymatic digestion experiments do not always go to completion and when digesting unusual glycan structures the removal of monosaccharides could be extremely difficult to accomplish (Paschinger et al., 2011). Also the Shisto core appears to be resistant to galactosidase digestion which is consistent with previous observations (Khoo et al., 2001). Despite these issues, changes in relative abundances can be very informative. The first observation when comparing the MALDI-TOF MS spectra before and after the digest is the decrease of the peak corresponding to the molecular ion at \( m/z \) 534 confirming the earlier proposed Gal\textsubscript{β1-3}GalNAc Core 1 structure (see Figure 86). The appearance of the molecular ion peak at \( m/z \) 708, the appearance of the new structural isomer in the molecular ion peak at \( m/z \) 738, the disappearance of the old and appearance of the new structural isomer in the molecular ion peak at \( m/z \) 912, taken together is explained by the removal of the proposed β1-6 galactose from the GalNAc residue thus confirming the presence of the earlier proposed (see Figure 87) Gal\textsubscript{β1-3-(Galβ1-6)}GalNAc Schisto Core. Another important observation is that the peaks above \( m/z \) 1146 corresponding to the higher molecular weight O-glycans with the general composition of Gal\textsubscript{4-9}Fuc\textsubscript{0-1}GalNAc\textsubscript{1} have disappeared suggesting they carry β1-3/4/6-galactose(s). Finally, in contrast to the 6-linked branch hexoses, those attached to the 3-linked branch galactose of the Schisto Core seem to be resistant to the broad specificity β\textsubscript{1-3, 4, 6}-galactosidase enzyme.
Figure 93 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with bovine testes broad specificity β1-3, 4, 6 galactosidase. The MALDI-TOF-MS spectrum prior to the digest is given for reference purposes and has been shown earlier (Figure 85). The proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), previous results, GC-MS data (see section 4.5.2) and the knowledge of biosynthetic pathways. M’ and ‘m’ represent major and minor structures, respectively.
To exemplify the broad specificity β1-3, 4, 6-galactosidase ability to remove the hexoses from the 6-linked branch compared to the hexoses attached to the 3-linked branch galactose the MALDI-TOF/TOF MS/MS spectra of the molecular ion peak at m/z 1146 before and after the enzymatic digest are shown in Figure 94. Prior to the enzymatic treatment the molecular ion at m/z 1146 was considered to potentially include three structural isomers having one and three, two and two, three and one hexoses on the 6-linked and 3-linked GalNAc branches, respectively. These structural isomers are supported by the similarly abundant ‘Z-type’ fragment ions (Domon, Costello 1988) at m/z 502, 706 and 910 resulting from the elimination of the hexose(s) from the 3-position of the GalNAc. Upon sample treatment with the broad specificity β1-3, 4, 6-galactosidase, the fragment ions at m/z 706 and 910 corresponding to the structural isomers with two and three hexoses respectively on the 6-linked GalNAc branch are significantly reduced. The only remaining structural isomer is with three hexoses on the 3-linked GalNAc branch as seen by the fragment ion at m/z 502. These results suggest that the hexoses on the 6-linked GalNAc branch are β1-3/4/6-linked galactoses. In addition, the presence of the fragment ion at m/z 667 containing three hexoses, and the absence of the fragment ion at m/z 463 containing two hexoses, suggest a branched, not linear arrangement of the three hexoses on the 3-position of GalNAc.
Figure 94 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated m/z 1146 L3 D. viviparus O-glycans before and after the bovine testes broad specificity β1-3, 4, 6 galactosidase digestion. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
4.5.8 MALDI-TOF MS analyses of permethylated reductively eliminated O-glycans which were subjected to consecutive broad specificity β-galactosidase and broad specificity α-galactosidase treatment

The reductively eliminated, broad specificity β1-3, 4, 6-galactosidase digested O-glycans from the *D. viviparus* L3 stage were subjected to green coffee bean broad specificity α-galactosidase digestion. The MALDI-TOF MS spectrum is presented in the Figure 95, bottom. The disappearance of the peaks at *m/z* 1146 and 1116, the reduction in the relative abundance of the peaks at *m/z* 942 and 912, and the increase in the relative abundance of the peak at *m/z* 708 suggest the presence of α-linked galactoses in the sample. The reduction in the relative abundance of the peak at *m/z* 912 and the complete disappearance of the structural isomer having a second hexose attached to the β1-3-linked branch galactose, together with the increase in the relative abundance of the structural isomer in the molecular ion peak at *m/z* 708 provides strong evidence for this hexose being α-linked galactose. The appearance of the minor structural isomer at *m/z* 912 having galactose 6-linked to the GalNAc can be explained by the removal of the α-linked galactose from the molecular structure at *m/z* 1116. The disappearance of the structural isomer bearing a hexose attached to the 3-linked branch galactose at *m/z* 738 provides the proof for this hexose being α-linked galactose. Additional MALDI-TOF-TOF MS/MS structural evidence is discussed later (see section 4.5.9). Prior to the broad specificity α-galactosidase digestion, the O-glycans at *m/z* 1146 and one of the structural isomers at *m/z* 942 have two hexoses attached to the 3-linked branch galactose and the question arises whether both of them are α-linked galactoses. If only one of them was α-linked galactose, after the broad specificity α-galactosidase digestion the molecular ion at *m/z* 738 would have had a structural isomer bearing two hexoses on the 3-linked GalNAc branch. The fact that it has not been observed (see next section 4.5.9) suggests both hexoses attached to the 3-linked branch galactose are α-linked galactoses.
Figure 95 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with bovine testes broad specificity β1-3, 4, 6 galactosidase and broad specificity α-galactosidase enzymes. The MALDI-TOF-MS spectrum of O-glycans digested by broad specificity β1-3, 4, 6 galactosidase is given for the reference purposes and has been shown earlier (Figure 93, bottom). The proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), GC-MS data (see section 4.5.2), previous results and the knowledge of biosynthetic pathways. M’ and ‘m’ represent major and minor structures, respectively.
4.5.9 MALDI-TOF-TOF MS/MS analyses of permethylated reductively eliminated O-glycans which were subject to consecutive broad specificity β-galactosidase and broad specificity α-galactosidase treatment

The MALDI-TOF-TOF MS/MS spectra of the molecular ion at \( m/z \) 738 both before and after the broad specificity β1-3, 4, 6-galactosidase digestion are shown in the Figure 96. After the broad specificity β1-3, 4, 6-galactosidase treatment the molecular ion at \( m/z \) 738 appears to contain two structural isomers. One is a Schisto Core O-glycan with the galactoses being attached to the 3- and 6-positions of the GalNAc and is represented by fragment ions at \( m/z \) 502 and 259. The other structural isomer has a hexose attached to the 3-branch galactose and is represented by the fragment ions at \( m/z \) 463 and 298. Upon the subsequent sample treatment with the broad specificity α-galactosidase the structural isomer with two hexoses on the 3-linked branch disappears as is seen by the virtual disappearance of the fragment ions at \( m/z \) 463 and 298. The presence of the Schisto core structure in both samples is also supported by the presence of a low abundance internal reducing-end GalNAc fragment ion at \( m/z \) 284. Together these results provide evidence for the hexose attached to the 3-linked branch galactose being an α-linked galactose.
Figure 96 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated, β1-3, 4, 6 galactosidase digested m/z 738 L3 D. viviparus O-glycans before and after the green coffee beans broad specificity α1-3, 4, 6 galactosidase digestion. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
The molecular ion peaks at \( m/z \) 912 from the broad specificity \( \beta \)-galactosidase digested O-glycans, and from broad specificity \( \beta \)-galactosidase and subsequently \( \alpha \)-galactosidase digest O-glycans, were subjected to MALDI-TOF-TOF structural analyses. The resultant MS/MS spectra are shown in the Figure 97. The main structural isomer found in the molecular ion at \( m/z \) 912 from the broad specificity \( \beta \)-galactosidase digested O-glycans has a Hex-(Fuc)-Gal structure attached to the 3-position of the reducing-end GalNAc as seen by the most abundant fragment ions at \( m/z \) 637, the corresponding fragment ion at \( m/z \) 298 and the secondary fragment ions at \( m/z \) 415 and 449. The traces of another structural isomer having fucose attached to the 3-linked branch galactose and the 6-linked branch with one galactose are also present as seen by the low-abundance fragment ions at \( m/z \) 502 and 433. After the broad specificity \( \alpha \)-galactosidase treatment, the previously most abundant structural isomer with the Hex-(Fuc)-Gal structure attached to the 3-position of the GalNAc almost completely disappears. This is consistent with the complete disappearance of the fragment ions at \( m/z \) 298 and by the trace presence of the fragment ion at \( m/z \) 637. Instead, a previously low-abundance Shisto Core fucosylated structural isomer becomes the most abundant as is seen by the most abundant fragment ions at \( m/z \) 502 and the corresponding fragment ions at \( m/z \) 433. These results show the susceptibility of the Hex-(Fuc)-Gal containing structural isomer to the broad specificity \( \alpha \)-galactosidase enzyme digestion and prove the non-reducing end terminal hexose of the ‘lower’ branch being an \( \alpha \)-linked galactose.
Figure 97 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated, β1-3, 4, 6 galactosidase digested m/z 912 L3 D. viviparus O-glycans before and after green coffee bean broad specificity α1-3, 4, 6 galactosidase digestion. The blue arrowed horizontal lines represent the loss of depicted structures. ‘M’ represents the major structure. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
The Figure 98 shows the MALDI-TOF-TOF MS/MS spectra of the molecular ions at m/z 942 of the broad specificity β-galactosidase digested O-glycans prior and after the broad specificity α-galactosidase digestion. In the broad specificity β-galactosidase digested O-glycans the molecular ion at m/z 942 is made up of at least three structural isomers. O-glycan isomers having two and three hexoses on the 3-linked branch are represented by the fragment ions at m/z 502, 463 and 298, 667 respectively. There is also a trace of the third structural isomer having only one 3-linked galactose on the 3-linked branch and two on the 6-linked branch. This structural isomer is best represented by the fragment ions at m/z 706 since the corresponding fragment ions at m/z 259 can also result from fragmentation of the other structural isomers. After the broad specificity α-galactosidase digestion the molecular ion peak at m/z 942 as seen in MALDI-MS spectrum (see Figure 95) is greatly reduced in terms of relative abundance and upon MALDI-TOF-TOF MS/MS fragmentation only traces of the structural isomers with one and two hexoses on the 3-linked branch of the Schisto Core are observed as indicated by fragment ions at m/z 502 and 706 respectively. The structural isomer with three hexoses on the 3-linked branch virtually disappears as seen by the absence of the fragment ions at m/z 667 and 298. Together with the presence of only one Schisto Core structural isomer in the molecular ion at m/z 738 after broad specificity α-galactosidase digestion (see previous Figure 96) these results show the susceptibility to digestion of both hexoses on the 3-linked branch galactose establishing them as α-linked galactoses.
Figure 98 MALDI-TOF-TOF MS/MS spectra of permethylated reductively eliminated, β1-3, 4, 6 galactosidase digested m/z 942 L3 D. viviparus O-glycans before and after the green coffee bean broad specificity α1-3, 4, 6 galactosidase digestion. Blue arrowed horizontal lines represent the loss of depicted structure. The peak at m/z 538 most likely represents the contaminant. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
4.5.10 MALDI-MS<sup>n</sup> analysis of permethylated reductively eliminated m/z 708 O-glycan after consecutive broad specificity β-galactosidase and broad specificity α-galactosidase treatment

Although the potential α1-2 linkage between the fucose and galactose residues has been suggested earlier (see Figure 89), to provide additional evidence the molecular ion at m/z 708 from the reductively eliminated, broad specificity β-galactosidase and broad specificity α-galactosidase digested and permethylated O-glycans (see Figure 95), were subjected to MS<sup>2</sup> analysis on the MALDI-QIT-TOF instrument as seen in the Figure 99. The most abundant fragment ions at m/z 433 was subjected to MS<sup>3</sup> analysis as seen in the Figure 100, top. The most abundant fragment ions at m/z 245 together with the fragment ions at m/z 229, 227 and 211 confirm a linked galactose and a terminal fucose. Cross-ring fragment ions at m/z 271, 315 and 329 suggest the fucose is 2-linked to the galactose. To get even stronger evidence a glycan standard containing the Fucα1-2Gal structure was also subjected to MS<sup>n</sup> analysis. The MS<sup>3</sup> spectrum of the Fuc-Gal fragment ions at m/z 433 is shown in the Figure 100, bottom. Together with the 2-linked galactose seen in the GC-MS data (see section 4.5.2) the almost
identical spectra provide an additional evidence for the fucose being 2-linked to the galactose in the *D. viviparus* L3 stage O-glycan.

**Figure 100** MALDI-QIT-TOF MS^3^ spectrum of m/z 433 fragment ion from permethylated reductively eliminated, bovine testes β1-3,4,6 galactosidase and green coffee bean α1-3,4,6 galactosidase digested m/z 708 O-glycan from *D. viviparus* L3 stage. The MALDI-QIT-TOF MS^3^ spectrum of m/z 433 Fucα1-2Gal fragment is shown for the reference. Dashed arrowed lines on the proposed/known structures represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations for the top figure are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.
4.5.11 MALDI-MS analysis of permethylated reductively eliminated, *X. manihotis* β-galactosidase digested O-glycans

The reductively eliminated *D. viviparus* L3 stage O-glycans were also subjected to β-galactosidase from *X. manihotis*. Although this enzyme has been marketed as β1-3,6 galactosidase, in fact the enzyme 100 times prefers β1-3 over β1-6 linked galactose and 500 times β1-3 over β1-4 linked galactose (Wong-Madden and Landry, 1995) (unpublished observations obtained from New England Biolabs (NEB) website). The resultant MALDI-TOF MS spectrum can be seen in Figure 101 where all the molecular ion peaks have also been subjected to the MALDI-TOF-TOF MS/MS analyses. As with the bovine testes broad specificity β1-3, 4, 6 galactosidase digestion, the molecular ion peaks corresponding to O-glycans above *m/z* 1146 disappear suggesting the presence of the β-linked galactoses in the sample. On the other hand, due to the not strict specificity of the enzyme used, the specific linkages of the galactoses cannot be determined from the MALDI-TOF MS and MALDI-TOF-TOF MS/MS experiments alone.
Figure 101 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with X. manihotis β-galactosidase. The MALDI-TOF-MS spectrum prior to the digest is given for the reference purposes and has been shown earlier (Figure 85). The proposed structures are based on the MALDI-TOF-MS/MS analyses (peaks marked with an asterisk*), GC-MS data (see section 4.5.2), previous results and the knowledge of biosynthetic pathways.
4.5.12 MALDI-MS² analyses of permethylated reductively eliminated, *X. manihotis* β-galactosidase digested O-glycan at m/z 942

![MALDI-QIT-TOF MS² spectrum of permethylated reductively eliminated, X. manihotis β-galactosidase digested m/z 942 O-glycan from the D. viviparus L3 stage. The blue arrowed horizontal line represents the loss of depicted fragment. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.](image)

The molecular ion at m/z 942 from the *X. manihotis* β-galactosidase digested, permethylated O-glycans (see previous Figure 101, bottom) was further fragmented by the MALDI-QIT-TOF MS² instrumentation. In contrast to the bovine testes broad specificity β-galactosidase digestion experiment (see Figure 98, top) where three structural isomers were present, the *X. manihotis* β-galactosidase digestion resulted in only one structural isomer as seen in Figure 102. The two most abundant fragment ions at m/z 502 and 463 suggest the presence of the structural isomer where two galactoses comprise the 3-linked branch of the Schisto Core and one galactose comprises the 6-linked branch. No fragment ion at m/z 706 indicates the absence of the structural isomer with two galactoses on the 6-linked branch and no fragment ions at m/z 298 and 667 indicate the absence of the structural isomer with three galactoses on the 3-linked Schisto Core branch.
Figure 103 MALDI-QIT-TOF MS$^3$ spectrum of m/z 463 fragment ions from permethylated reductively eliminated, X. manihotis β-galactosidase digested m/z 942 O-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

In order to obtain linkage information, the presence of only one structural isomer makes the fragment ion at m/z 463 a good target for MALDI-QIT-TOF MS$^3$ structural analysis. In this way we know that the two galactoses are originating only from the 3-linked branch of the Schisto Core. The MALDI-QIT-TOF MS$^3$ spectrum of the fragment ion at m/z 463 is shown in Figure 103. When determining the linkage the most informative cross-ring fragments in the spectrum are at m/z 345, 141, 315, 171 and 389 suggesting presence of Gal1-3Gal structure attached to the 3-position of the reducing-end GalNAc.

To support this suggestion we can consider the following. The GC-MS data of the D. viviparus L3 Stage O-glycans (see section 4.5.2) showed the presence of 2-linked, 3-linked and 6-linked galactoses. The Gal-Gal structure attached to the 3-position of the GalNAc has been proved to be prone to the green coffee bean broad specificity α1-3,4,6 galactosidase. Putting these facts together we can assume Galα1-3Gal or Galα1-6Gal structures on the 3-linked branch of the Schisto Core.
In order to get more information about the fragmentation of the 6-linked and 3-linked hexoses, the fragment ions at m/z 463 corresponding to Glcβ1-3Glc and Manα1-6Man structures from the laminaritriose and the oligomannose standards were subjected to MALDI-QIT-TOF MS^3 analyses (see Figure 104). Although the relative abundance of cross-ring fragment ions upon the fragmentation of various hexoses can be different (Ashline et al., 2005), the actual m/z values of the cross-ring fragments are expected to be the same (Ashline et al., 2005). When comparing the MS^3 spectrum of the fragment ions at m/z 463 from the molecular ion at m/z 942 (see Figure 103) to the MS^3 spectrum of the Glcβ1-3Glc fragment ion (see Figure 104, top), there is the presence of the common cross-ring fragment ions at m/z 171, 315, 345, 359 and 389. The MS^3 spectrum of the fragment ions derived from the *D. viviparus* L3 O-glycan lacks characteristic fragments for the 6-linked hexose such as ions at m/z 301 and 329 seen in the MS^3 spectrum of the Manα1-6Man structure (see Figure 104, bottom). All these data put together strongly suggest the Galα1-3Gal structure forming the 3-linked branch of the Shisto Core.
Figure 104 MALDI-QIT-TOF MS$^3$ spectra are showing $m/z$ 463 Glcα1-3Glc and Manα1-6Man fragment ions from permethylated known standard compounds. Dashed arrowed lines on the known structures represent glycosidic bond cleavages and cross-ring fragments. The spectrum above has the zoomed in the $m/z$ 300-460 region of the spectrum inserted.
4.5.13 MALDI-MS\(^n\) analyses of permethylated reductively eliminated, \textit{X. manihotis} β-galactosidase digested O-glycan at \(m/z\) 1146

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure105.png}
\caption{MALDI-QIT-TOF MS\(^2\) spectrum of permethylated reductively eliminated, \textit{X. manihotis} β-galactosidase digested \(m/z\) 1146 O-glycan from the \textit{D. viviparus} L3 stage. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.}
\end{figure}

The molecular ions at \(m/z\) 1146 from the \textit{X. manihotis} β-galactosidase digested, permethylated O-glycans (see previous Figure 101, bottom) were further fragmented by the MALDI-QIT-TOF MS\(^2\) instrument as seen in the Figure 105. The most abundant fragment ions at \(m/z\) 667 and 502 indicate the structural isomer with the three galactoses on the 3-linked branch and one galactose on the 6-linked branch of the Schisto Core. In addition very small amounts of another structural isomer having two galactoses on both Schisto Core branches are also present as indicated by the fragment ions at \(m/z\) 706 and 463.
To obtain linkage information, the fragment ion at m/z 667 was subjected to MALDI-QIT-TOF MS³ fragmentation as seen in the Figure 106. The absence of the fragment ions at m/z 463 corresponding to the Gal-Gal structure confirms the branched nature of these three galactoses as proposed earlier (see Figure 94, bottom). Instead the fragment ions at m/z 449 and 259 corresponding to the cleavage of one hexose are observed. The presence of the characteristic 6-linked hexose cross-ring fragment ions at m/z 301 and 329 (see Figure 104) strongly suggests the presence of the 6-linked galactose. The presence of the cross-ring fragment ions at m/z 315, 345, 375 (see Figure 104) also suggests the presence of the 3-linked galactose. The cross ring fragments at m/z 593 and 563 suggest that the reducing end galactose is not 2-linked.

Out of both 2, 3- and 3, 6-linked galactoses seen in the GC-MS data (see section 4.5.2) the linkage between the three galactoses on the -3 branch of the Schisto Core is consistent with the proposed structure.
the latter. These results are also consistent with the fact that both galactoses attached to the galactose of the 3-linked branch of the Schisto Core are sensitive to green coffee bean broad specificity α1-3, 4, 6 galactosidase (see Figures 96 and 98) which does not cleave 2-linked galactose. Put together the results from the MS3, the GC-MS and the α1-3,4,6 galactosidase digestion experiments provide strong evidence for the Galα1-3-(Galα1-6)Gal structure attached to the 3-position of the reducing-end GalNAc.

4.5.14 MALDI-MS analysis of permethylated reductively eliminated, α-galactosidase digested O-glycans

From the green coffee beans broad specificity α1-3,4,6 galactosidase digestion of reductively eliminated, bovine testes broad specificity β-galactosidase digested O-glycans (see previous Figure 95) we already know that the broad specificity α1-3,4,6 galactosidase removes hexoses from the 3-linked branch of the Schisto Core. Nevertheless the same broad specificity α1-3,4,6 galactosidase digestion was performed on reductively eliminated O-glycans in order to obtain good targets for the MALDI-QIT-TOF MSn analyses. The resultant MALDI-TOF MS spectrum can be seen in the Figure 107. The increase in the relative abundance of the molecular ion peak at m/z 738, the decrease in the relative abundance of the molecular ion peak at m/z 1116 in comparison to molecular ion peak at m/z 912, and the disappearance of the glycan molecular ions above the m/z 1350, are all consistent with the already described enzymatic removal of α-linked galactose(s) on the 3-linked branch of the Schisto Core (see sections 4.5.8 and 4.5.9).
Figure 107 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with green coffee bean broad specificity α1-3, 4, 6 galactosidase. The MALDI-TOF-MS spectrum prior to the digest is given for the reference purposes and has been shown earlier (see Figure 85). The proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), previous results and the knowledge of biosynthetic pathways. ‘M’ and ‘m’ corresponds to major and minor structures, respectively.
4.5.15 MALDI-MS<sup>n</sup> analyses of permethylated reductively eliminated, α-galactosidase digested O-glycan at m/z 942

**Figure 108** MALDI-QIT-TOF MS<sup>2</sup> spectrum of permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 942 O-glycan from the *D. viviparus* L3 stage parasite. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

In order to obtain the linkage information between the galactoses on the 6-linked branch of the Schisto Core O-glycans the molecular ion at m/z 942 after the broad specificity α-galactosidase digestion was subjected to the MALDI-QIT-TOF MS<sup>n</sup> analyses. Initially it was subjected to the MS<sup>2</sup> structural analysis (see Figure 108) where fragment ions at m/z 502, 463 indicate the structural isomer with two galactoses on the 3-linked and one galactose on the 6-linked branch of the Schisto Core. In addition the abundant fragment ion at m/z 706 indicates another structural isomer with two galactoses on the 6-linked, and one galactose on the 3-linked, branch of the Schisto Core. The expected fragment ion at m/z 259 corresponding to one hexose is not seen due to the low sensitivity in the lower molecular weight region of
the MALDI-QIT-TOF instrument. It has to be noted that although the majority of the fragment ion at \( m/z \) 463 is expected to originate from the more favourable Gal-Gal structure elimination from the GalNAc 3-position, it could also potentially be generated from the cleavage of the Gal-Gal structure from the GalNAc 6-position.

![Figure 109](image)

**Figure 109** MALDI-QIT-TOF MS\(^3\) spectrum of \( m/z \) 706 fragment ion from permethylated reductively eliminated, green coffee bean \( \alpha \)-galactosidase digested \( m/z \) 1942 O-glycan. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

To obtain the fragment ion at \( m/z \) 463 generated from the 6-linked branch of the Schisto Core, the MALDI-QIT-TOF MS\(^3\) structural analysis of the fragment ion at \( m/z \) 706 was performed as seen in the **Figure 109**. Although with MS\(^n\) analyses the sensitivity decreases with each consecutive ‘n’ experiment, in this case the most abundant generated fragment ion was the desired \( m/z \) 463 thus enabling us to perform the MS\(^4\) analysis on it. Other fragment ions at \( m/z \) 227, 245, 259, 284, 488 confirmed the proposed structure for the fragment ion at \( m/z \) 706.
Figure 110 MALDI-QIT-TOF MS4 spectrum of m/z 463 fragment ion from m/z 706 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1146 O-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

The MALDI-QIT-TOF MS4 spectrum of the fragment ion at m/z 463 is shown in the Figure 110. The fragment ions at m/z 227, 245, 241 and 259 originate from the glycosidic bond cleavage and, together with fragment ion at m/z 431 resulting from the loss of methanol, do not yield any information about the linkage. However, the cross-ring fragment ions at m/z 301, 329 are characteristic for 6-linked hexose (see Figure 104, bottom). In addition the cross-ring fragment ions at m/z 345 and 389 are also consistent with 6-linked hexose. The most abundant fragment ion at m/z 359 could originate from the analogous to m/z 389 fragmentation but with an extra loss of methanal coming from the methyl group. Overall the MS4 spectrum of the fragment ion at m/z 463 from the O-glycan at m/z 942 (see Figure 110) looks remarkably similar to the MS3 spectrum of the Manα1-6Man fragment as seen in Figure 104, bottom providing strong evidence for the Galβ1-6Gal structure comprising the 6-linked branch of the Schisto Core. The proposed Galβ1-6Gal structure on the 6-linked
branch is also supported by the presence of the 6-linked galactose in the GC-MS data shown earlier (see section 4.5.2).

4.5.16 MALDI-MS² analyses of permethylated reductively eliminated, α-galactosidase digested m/z 1146 O-glycan

![Figure 111](image)

**Figure 111** MALDI-QIT-TOF MS² spectrum of permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1146 O-glycan from the *D. viviparus* infective L3 stage parasite. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

In order to get the linkage information between the second and the third galactoses of the 6-linked branch of the Schisto Core MALDI-QIT-TOF MS² analyses was performed on the permethylated reductively eliminated, broad specificity α-galactosidase digested molecular ions at m/z 1146. The MALDI-QIT-TOF MS² analysis revealed three structural isomers present as seen in the **Figure 111**. The fragment ions at m/z 502 and 667 correspond to the structural isomer with three galactoses on the 3-linked branch and one galactose on the 6-linked branch of the Schisto Core. It has to be noted that the fragment ion at m/z 667 can also originate from another structural isomer with three galactoses on the 6-linked branch and one
galactose on the 3-linked branch of the Schisto Core. The best evidence for this second structural isomer is the fragment ion at $m/z$ 910. The expected additional fragment ion at $m/z$ 259 corresponding to one hexose is not seen, possibly due to the low sensitivity of the instrument in the low molecular weight region. The third structural isomer has two galactoses on both branches of the Schisto Core and is represented by the fragment ions at $m/z$ 706 and 463 although the latter fragment ion can be generated from other structural isomers as well.

**Figure 112** MALDI-QIT-TOF MS$^3$ spectrum of $m/z$ 910 fragment ion from permethylated reductively eliminated, green coffee bean $\alpha$-galactosidase digested $m/z$ 1146 O-glycan. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

To be confident that the galactoses we fragment are originating from the 6-linked branch of the Schisto Core, the fragment ion at $m/z$ 910 was selected for further MALDI-QIT-TOF MS$^3$ analysis as seen in the Figure 112. The fragment ions at $m/z$ 463 and 488 indicate that three galactoses on the 6-linked branch are arranged in a linear, not branched form. Other fragment ions at $m/z$ 259, 284, 449, 667 692 support the proposed structure. The abundance of the fragment ions at $m/z$ 667 and 463 corresponding to three and two galactoses encouraged us to perform MALDI-QIT-TOF MS$^4$ analyses on them.
Figure 113 MALDI-QIT-TOF MS$^4$ spectrum of m/z 667 fragment ions from m/z 910 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1146 O-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

The MALDI-QIT-TOF MS$^4$ analysis of the fragment ion at m/z 667 was carried out as seen in the Figure 113. The presence of the fragment ion at m/z 463 confirms the earlier proposed linear, not branched, structure for the three 6-linked branch galactoses. Fragment ions at m/z 259 and 449 correspond to a hexose and two hexoses with two hydroxyl groups, although they are not as informative since these fragment ions can be generated from both linear and branched structures. The cross-ring fragment ions at m/z 533 and 505 confirm the earlier proposed β1-6 glycosidic bond between the first two galactoses of the 6-linked branch. The fragment ions at m/z 329 may suggest a 1-6 glycosidic bond between the second and third galactoses. MALDI-QIT-TOF MS$^4$ of the fragment ion at m/z 463 was carried out.
Figure 114 MALDI-QIT-TOF MS\textsuperscript{4} spectrum of \textit{m/z} 463 fragment ion from \textit{m/z} 910 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean \(\alpha\)-galactosidase digested \textit{m/z} 1146 O-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2) and previous results.

The MALDI-QIT-TOF MS\textsuperscript{4} analysis of the fragment ion at \textit{m/z} 463 corresponding to the second and third galactoses from the 6-linked branch of the Schisto Core was carried out as seen in the Figure 114. The fragment ions at \textit{m/z} 245 and 259 correspond to reducing and non-reducing end hexoses confirming the hexose-hexose structure although, together with the fragment ion at \textit{m/z} 431 generated by the loss of a methanol, they are not linkage-informative. However, characteristic for 6-linked hexose, the cross-ring fragment ions at \textit{m/z} 329 and 301 suggest the Gal\(\beta\)1-6Gal structure for the second and third galactoses in the 6-linked branch of the Schisto Core. The cross-ring fragment ion at \textit{m/z} 345 supports the Gal\(\beta\)1-6Gal structure. Although not assigned, the cross-ring fragments at \textit{m/z} 359 and 375 are also present in the MS\textsuperscript{3} spectrum of the Man\(\alpha\)1-6Man structure (see previous Figure 104, bottom). Overall the spectrum profile, although of lower quality due to the MS\textsuperscript{4} nature of the experiment looks very similar to the MS\textsuperscript{3} spectrum of the Man\(\alpha\)1-6Man fragment ion (previous Figure 104, bottom). Together with the presence of the 6-linked galactose in the GC-MS results (see...
section 4.5.2) this data provides a strong evidence for the Galβ1-6Gal structure for the second and third galactoses on the 6-linked linear branch of the Schisto Core.

4.5.17 MALDI-MS analysis of permethylated reductively eliminated, broad specificity α-galactosidase and subsequently broad specificity β1-3, 4, 6 galactosidase digested O-glycans

The reductively eliminated and broad specificity α-galactosidase digested *D. viviparus* L3 stage O-glycans described in the Figure 107, bottom have also been subjected to consecutive bovine testes broad specificity β1-3, 4, 6 galactosidase digestion. The resultant MALDI-TOF MS spectrum can be seen in the Figure 115 where the structures presented are based on the MALDI-TOF-TOF MS/MS fragmentation and previous results. The appearance of the new structural isomer in the molecular ion at *m/z* 534 having galactose on the 6-position of the GalNAc, the decrease of the molecular ion peak at *m/z* 942, the disappearance of the molecular ion peaks at *m/z* 1116 and 1146, are all consistent with the removal of up to three already-proven β1-6 galactoses from the 6-linked branch of the Schisto Core. The key result of this experiment is the disappearance of the molecular ion peak at *m/z* 1350. As shown in the MALDI-TOF| MS spectrum in Figure 115, top and proven later by the MALDI-QIT-TOF MS² results in the Figures 116, the most abundant structural isomer comprising this molecular ion is the one having four hexoses on the 6-linked branch of the Schisto Core. Although the initial molecular ion peak at *m/z* 1350 in the Figure 115, top has relatively low abundance, considering that it completely disappears after the broad specificity β1-3, 4, 6 galactosidase digestion (see Figure 115, bottom), and that the most abundant structural isomer has four hexoses on the 6-linked branch, all this data points to the fourth hexose on the 6-linked branch of the Schisto Core being β-galactose.
Figure 115 MALDI-TOF MS spectrum (35% acetonitrile fraction) of permethylated O-glycans released by reductive elimination and digested with broad specificity α-galactosidase and bovine testes broad specificity β1-3, 4, 6 galactosidase enzymes. The MALDI-TOF-MS spectrum of O-glycans digested by broad specificity α1-3, 4, 6 galactosidase is given for the reference purposes and has been shown earlier (Figure 107, bottom). The proposed structures are based on the MALDI-TOF-TOF MS/MS analyses (peaks marked with an asterisk*), previous results and the knowledge of biosynthetic pathways. ‘M’ and ‘m’ represent major and minor structures, respectively.
4.5.18 MALDI-MS\textsuperscript{n} analyses of permethylated reductively eliminated, α-galactosidase digested \( m/z \) 1350 O-glycan

Figure 116 MALDI-QIT-TOF MS\textsuperscript{2} spectrum of permethylated reductively eliminated, green coffee bean α-galactosidase digested \( m/z \) 1350 O-glycan from *D. viviparus* L3 stage parasite. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

In a similar way to the molecular ion at \( m/z \) 1146 (see section 4.5.16), the permethylated reductively eliminated, broad specificity α-galactosidase digested molecular ion at \( m/z \) 1350 was further analysed by MALDI-QIT-TOF MS\textsuperscript{n} approach in order to get information about the linkage between the third and fourth galactoses. The MALDI-QIT-TOF MS\textsuperscript{2} spectrum of the molecular ions at \( m/z \) 1350 is shown in the Figure 116 and reveals four structural isomers are present. The fragment ions at \( m/z \) 1114 represent the structural isomer with four galactoses on the 6-linked branch of the Schisto Core. The corresponding fragment ion at \( m/z \) 259 is not observed due to the MALDI-QIT-TOF instrument sensitivity limitations in the lower molecular mass region. The fragment ions at \( m/z \) 910 with corresponding fragment ions at \( m/z \) 463 indicate the presence of the structural isomer with three galactoses on the 6-linked branch and two on the 3-linked branch of the Schisto Core. The fragment ions at \( m/z \) 706, 667 originate from the structural isomer with two galactoses on the 6-linked branch and three on the 3-linked branch branches of the Schisto Core. Finally the fragment ions at \( m/z \) 502 and
871 indicate the presence of the structural isomer with one galactose on the 6-linked branch and four galactoses on the 3-linked branch of the Schisto Core. It has to be noted that the fragment ions at m/z 463, 667, 871 corresponding to two, three and four galactoses are not unique to the structural isomers assigned and can originate from other structural isomers as well.

![Figure 117](image.png)

**Figure 117** MALDI-QIT-TOF MS³ spectrum of m/z 1114 fragment ion from permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1350 O-glycan. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2), previous results and knowledge of biosynthetic pathways in other organisms.

To be confident that the galactoses originate from the 6-linked branch of the Schisto Core, the fragment ion at m/z 1114 were further fragmented in the MALDI-QIT-TOF MS³ experiment and the resultant spectrum is shown in the **Figure 117**. The presence of the fragment ions at m/z 463 and 667 prove the linear, not branched arrangement of the four galactoses on the 6-linked branch of the Schisto Core. This arrangement is supported by the series of fragment ions at m/z 488, 692 and 896. The fragment ions at m/z 871 comprised of four galactoses support the proposed structure of the parental fragment ion.
Figure 118 MALDI-QIT-TOF MS\textsuperscript{4} spectrum of m/z 463 fragment ion from m/z 1114 ion precursor which in turn originated from permethylated reductively eliminated, green coffee bean α-galactosidase digested m/z 1350 O-glycan. Dashed arrowed lines on the proposed structure represent glycosidic bond cleavages and cross-ring fragments. The proposed monosaccharide configurations are based on GC-MS data (see section 4.5.2) and previous results.

In order to determine the linkage between the third and fourth galactoses of the 6-linked branch of the Schisto Core, the fragment ion at m/z 463 was selected for MALDI-QIT-TOF MS\textsuperscript{4} analysis as seen in Figure 118. Due to the initial relatively low abundance of the molecular ions at m/z 1350 in the MS spectrum and the nature of the MS\textsuperscript{n} experiment, the resultant MS\textsuperscript{4} spectrum is not of the highest quality. Nevertheless the cross-ring fragments at m/z 329 and 345 indicate 6- or 4-linked galactose. Although with some missing peaks, the resemblance of this spectrum to the standard MS\textsuperscript{3} spectrum of the Man\textalpha1-6Man fragment (see previous Figure 104, bottom), the proven β1-6 linkage between the first-second and second-third and the presence of 6-linked, not 4-linked galactose in the GC-MS data (see section 4.5.2) makes the β1-6, not β1-4 linkage most probable between the third and the fourth galactoses of the 6-linked branch of the Schisto Core.
4.5.19 The summary of reductively eliminated O-glycan structures from L3 stage *D. viviparus*

Putting all the data together acquired so far in the thesis about reductively eliminated O-glycans from L3 stage *D. viviparus*, the proposed structures with the corresponding m/z values can be seen below in the Table 23. Considering that O-glycans from *D. viviparus* have not been reported previously, the proposed linkages for most of the structures have been included as well.

<table>
<thead>
<tr>
<th>m/z</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>534</td>
<td>β1-3</td>
</tr>
<tr>
<td>738</td>
<td>β1-3, β1-6</td>
</tr>
<tr>
<td>912</td>
<td>α1-2, β1-3, β1-6</td>
</tr>
<tr>
<td>942</td>
<td>β1-3, β1-6</td>
</tr>
<tr>
<td>1116</td>
<td>α1-3/6, β1-3, β1-6</td>
</tr>
<tr>
<td>1146</td>
<td>α1-3, β1-3</td>
</tr>
<tr>
<td>1321</td>
<td>β1-3, β1-6, α1-3/6</td>
</tr>
</tbody>
</table>
Table 23 showing the proposed O-glycan structures released by reductive elimination from L3 stage *D. viviparus*. 

<table>
<thead>
<tr>
<th>1351</th>
<th>( \beta_1-6 )</th>
<th>( \alpha_1-3 )</th>
<th>( \beta_1-6 )</th>
<th>( \alpha_1-6 )</th>
<th>( \beta_1-6 )</th>
<th>( \alpha_1-6 )</th>
<th>( \beta_1-6 )</th>
<th>( \alpha_1-3 )</th>
<th>( \beta_1-3 )</th>
<th>( \beta_1-3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1525</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-6 )</td>
<td>( \alpha_1-6 )</td>
<td>( \beta_1-6 )</td>
<td>( \alpha_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \alpha_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
</tr>
<tr>
<td>1555</td>
<td>( 4X )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
</tr>
<tr>
<td>1759</td>
<td>( 5X )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
</tr>
<tr>
<td>1963</td>
<td>( 6X )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
</tr>
<tr>
<td>2167</td>
<td>( 7X )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-3 )</td>
<td>( \beta_1-6 )</td>
<td>( \beta_1-3 )</td>
</tr>
</tbody>
</table>
4.6 Structural conclusions of *D. viviparus* parasite glycomes

4.6.1. Structural conclusions of adult N and O-glycans

The MALDI analysis of adult *Dictyocaulus viviparus* N-glycans showed the presence of high mannose, paucimannodic, truncated fucosylated and complex bi-, tri- and tetra-antennary structures with the latter carrying terminal GlcNAc(s), LacNAc(s) and Lewis^x^ structures. All these N-glycans have been reported earlier using FAB-MS techniques (Haslam et al., 2000). For the first time the *D. viviparus* O-glycan data have been described and the evidence for Galβ1-3GalNAc Core 1, Galβ1-3(Gal β1-6)GalNAc Schisto Core and Galβ1-3(GlcNAc β1-6)GalNAc Core 2 structures has been presented. Interestingly the GlcNAc of latter O-glycans can be modified into terminal LacNAc and Lewis^x^ structures, the same terminal antennae seen in N-glycans. The abundant Lewis^x^ containing structures have been proposed to interact with host lectins such as DC-SIGN (van Die et al., 2003) and skew immune response towards inflammatory CD4^+^ Th2 immune response (van Die and Cummings, 2010). Newly reported adult O-glycans are summarized in the Figure 119.

![Figure 119](image_url) proposed structural conclusions of O-glycan structures discovered in *D. viviparus* adult and L3 stages.
4.6.2. Structural conclusions of L3 stage N- and O-glycans

For the first time N-glycan data form *D. viviparus* L3 stage has been presented. Overall most of the N-glycan structures are also seen in the adult stage, although complex bi- and tri-antennary structures carrying terminal GlcNAc(s), LacNAc(s) and Lewis\(^x\) structures were less abundant. No tetra-antennary complex structures have been observed suggesting the wider transcriptomics profile in adult *D. viviparus* as also seen for *H. contortus* (Laing et al., 2013; Schwarz et al., 2013).

Surprisingly the analysis of the L3 stage *D. viviparus* O-glycans revealed not a more simple, but rather a substantially different O-glycome when compared to the adult stage as summarized in Figure 119. None of the Core 2 structures were observed and instead extended families of Schisto Core structures were detected. The first has poly-\(\beta1-6\) galactoses on the ‘upper’ 6-linked branch of the Schisto Core arranged linearly. Up to four \(\beta1-6\) galactoses have been demonstrated although longer, up to 8 hexoses, most likely all galactoses comprising ‘upper branch’ antennae are likely to be present. Interestingly Gal\(\beta1-6\)Gal structure has been detected in free bovine milk oligosaccharides which are not attached to any proteins or lipids (Aldredge et al., 2013). The 3-linked ‘lower branch of the Schisto Core has been demonstrated to contain \(\alpha\)-linked galactoses arranged in a branched manner. Initially \(\alpha\)-galactose is attached to the 3-position of the ‘lower branch’ galactose, then another \(\alpha\)-galactose is attached to the 6-position of the same \(\beta1-3\)-linked galactose. More hexoses, most likely galactoses can be added to the ‘lower branch’ of the Schisto Core. The Gal\(\beta1-3\)(Gal \(\beta1-6\))GalNAc Schisto Core was originally discovered in *Shistosoma mansoni* trematode however both ‘upper’ and ‘lower’ branches have Lewis\(^x\) rather than poly-galactose antennae (Khoo et al., 2001). Interestingly Gal\(\alpha1-3\)Gal structure has been detected in bovine milk oligosaccharides (Aldredge et al., 2013).

The second family of L3 stage *D. viviparus* O-glycans is also based on the Schisto Core however the ‘lower branch’ galactose is 2-linked to the fucose forming the ‘blood group H’ structure. In addition the same galactose can be further modified at the 3-position with an \(\alpha\)-galactose forming the blood group B antigen. No further addition of galactoses to the lower branch has been observed. Up to three galactoses have been observed on the ‘upper branch’, most likely arranged in a linear fashion forming \(\beta1-6\) linkages based on another non-fucosylated O-glycan family. The Fuca\(\alpha1-2\)Gal\(\beta1-3\)GalNAc blood group H structure has also
been observed in *P. tenuis* nematode however no further extensions of this structure were reported (Duffy et al., 2006).

The presence of terminal galactoses in *D. viviparus* L3 stage O-glycans suggest possible interactions with host lectins such as galectins thus modulating host immune responses (van Die and Cummings, 2010). The Galα1-3(Fuc α1-2)Gal blood group B antigen found in *E. coli* 086 strain was shown to interact with human galectins 3, 4 and 8 with galectins 4 and 8 being able to eliminate bacteria (Stowell et al., 2010). Later the same blood group B antigen was shown to interact with the subset of human dendritic cells found in mucus and skin, Langerhans cells via the C-type lectin langerin (Feinberg et al., 2011). To our knowledge no parasite-derived Fucα1-2Gal blood group H antigen-host lectin interactions have been reported, however the presence of this blood group antigen in cows (Aldredge et al., 2013) may suggest the molecular mimicry approach of the parasite to evade the host immune system. It is not known whether poly-β1-6-galactoses would interact with galectins or other host lectins.
Future perspectives

In my thesis I have reported the detailed structural characterizations of a range of novel N- and O-glycans from *Haemonchus contortus* and *Dictyocaulus viviparus*. Obtaining the structural information about the glycans, such as monosaccharide conformation, linkage and anomericity, is a pre-requisite before chemical synthesis and an investigation of their effects as potential vaccine candidates. Although synthesis of glycans still remains a daunting task, the progress in the field is obvious. In 2015 the first commercial automated carbohydrate synthesizer Glyconeer® appeared on the market. Although it is unlikely that at this stage it would be able to synthesize the unusual parasite glycans reported in my thesis.

An alternative solution could be the utilization of enzymatic synthesis which has been proven to be successful for the synthesis of some of the glycans. The Galβ1-4(Fucα1-3)GlcNAc (Lewis²), GalNAcβ1-4GlcNAc (LacdiNAc), LDNF (GalNAcβ1-4(Fucα1-3)GlcNAc), LDNDF (GalNAcβ1-4(Fucα1-2Fucα1-3)GlcNAc) structures found in schistosomes were successfully synthesised in micromolar amounts using recombinant β1-4GalT galactosyltransferase, α1-3FucT fucosyltransferase, β1-4GalNAcT N-acetylgalactosyltransferase, α1-2FucT fucosyltransferase enzymes (van Remoortere et al., 2000). In another study the Globo-H glycolipid-derived hexasaccharide Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galα1-4Glc has been synthesised which is one of the most promising glycan-based anticancer vaccine candidates (Su et al., 2008). Three recombinant glycosyltransferases, namely an α1-4 galactosyltransferase LgtC, a bifunctional β1-3 galactosyl/ β1-3 N-acetylgalactosaminyltransferase and an α1-2 fucosyltransferase WbsJ together with two recombinant epimerases GalE and WbgU for conversion of less expensive UDP-Glc, UDP-GlcNAc to more expensive UDP-Gal, UDP-GalNAc have been employed to achieve millimolar (grams) quantities of the hexasaccharide product (Su et al., 2008). Very recently commercially available recombinant β1-4 galactosyltransferase (B4GALT1) and α1-6 fucosyltransferase (FUT8) as well as recombinant α2-3 sialyltransferase (PmST1m), α2-6 sialyltransferase (Pd2,6ST), CMP-sialic acid synthetase (NmCSS) enzymes have been used to create sialylated Lewis² and sialylated LacNAc antennae and fucosylate the proximal GlcNAc of the chitobiose core on various N-glycans in an attempt to create a large library of N-glycans in milligram quantities (Li et al., 2015). Speaking about enzymatic O-glycan synthesis recombinant commercially available β1-4 galactosyltransferase (β4-GalT1), α2-3 sialyltransferase (ST3Gal3), α1-3 fucosyltransferase (α3-FucT3) enzymes together with
recombinantly expressed polypeptide N-acetylgalactosaminyltransferase 3 (GalNAcT3) and core-2 β6-N-acetylgalcosaminyltransferase (C2GnT) have been used to synthesise core-2 sialyl Lewis\(^x\) O-glycans in milligram quantities (Gutierrez Gallego et al., 2003).

Although many of the glycosyltransferases and glycosidases involved in glycan synthesis of the novel species specific nematode structures are not available commercially, a number of them have been cloned and characterized. Considering adult \(H.\ contortus\) N-glycans with Gal-Fuc modifications the range of glycan biosynthetic enzymes characterized in \(C.\ elegans\) organism immediately suggests the idea of using these enzymes for \(H.\ contortus\) N-glycans synthesis. The three fucosyltransferases FUT-1, FUT-6, FUT-8 generating the trifucosylated chitobiose core have been cloned and characterized (Paschinger et al., 2005; Titz et al., 2009; Yan et al., 2013). The galactosyltransferase GALT-1 involved in the generation of Galβ1-4Fuc structure attached to the 6-position of the proximal GlcNAc has also been cloned and characterized (Titz et al., 2009; Yan et al., 2013). No α1-2 galactosyltransferase(s) acting on the fucose attached to the 3-position of the distal GlcNAc has been cloned and characterized yet. However, over twenty potential α-galactosyltransferases have been identified in \(C.\ elegans\) (Yan et al., 2013; Zheng et al., 2002). It is likely that the required α1-2 galactosyltransferase is one of them. In contrast to \(C.\ elegans\) the N-glycans from \(H.\ contortus\) also contain the proposed GalNAcβ1-4(Galα1-2Fucα1-3)GlcNAc non-reducing end antennae. The GLY-13 protein transferring the GlcNAc onto 2-position of mannose (Chen et al., 2002), the β4GalNAc-T3 and β4GalNAc-T4 enzymes forming GalNAcβ1-4GlcNAc (Gotoh et al., 2004; Sato et al., 2003) and the FUT-6 enzyme being able to form GalNAcβ1-4(Fucα1-3)GlcNAc (LDNF) (Yan et al., 2013) have been cloned and characterized. The earlier discussed putative α1-2 galactosyltransferase generating the proposed Galα1-2Fuc on the 3-position of the distal GlcNAc could also add galactose to the LDNF structure. The summary of the strategies for enzymatic biosynthesis of adult \(H.\ contortus\) N-glycans is shown in Figure 120.

Considering the adult \(H.\ contortus\) O-glycans as potential vaccine candidates I will discuss the potential enzymatic synthesis of the highest molecular weight O-glycan observed, see Figure 120. The core 1 O-glycan T-synthase which adds β-galactose to the 3-position of GalNAc has been cloned and characterized (Ju et al., 2006). The β1-6 GlcNAc transferase responsible for the formation of GlcNAcβ1-6(Galβ1-3)GalNAc core 2 structure has also been cloned and characterized (Bierhuizen and Fukuda, 1992). The formation of the galactosylated LDNF structure has been described earlier for N-glycans. Although no α1-3
galactosyltransferase(s) acting on the terminal GalNAc have been cloned or characterized, a
H. contortus worm extract has been shown to contain such an activity (van Stijn et al., 2010b). When discussing O-glycans from L3 stage it would be worth considering the
synthesis of the Galβ1-3(Galβ1-6)GalNAc Schisto Core structure. Unfortunately to my
knowledge no β1-6 galactosyltransferase(s) acting on Galβ1-3GalNAc core 1 structure have
been reported. However the sequencing of the H. contortus genome (Laing et al., 2013;
Schwarz et al., 2013) allows the expectation that soon this enzyme will be discovered.

Both N- and O-glycans from D. viviparus contain Galβ1-4(Fucα1-3)GlcNAc (Lewis^x)
structures or their precursors. Considering the fact that the Lewis^x structure is abundant in
cows (Siciliano et al., 1993) it makes it less likely to consider these glycans as successful
vaccine candidates. Although the analysis of L3 stage D. viviparus glycans revealed
previously unseen O-glycan structures, most of their separate features such as Fucα1-2Gal
Blood group H, Galα1-3Gal, Galβ1-6Gal have been detected in bovine milk oligosaccharides
(Aldredge et al., 2013). On the other hand Galβ1-3(Galβ1-6)GalNAc Schisto Core could be
an attractive vaccine candidate especially because it is found both in L3 stage and adult
parasites. Its potential enzymatic synthesis has been described above as this structure is also
seen in H. contortus parasite. Although the Galβ1-6Gal structure has been observed in cows
(Aldredge et al., 2013) the poly-(Galβ1-6) structure, to my knowledge was not been observed
in any mammals. For this reason this structure can also be considered as a potential vaccine
candidate. Unfortunately no β1-6- galactosyltransferase(s) have been identified. However the
generated transcriptomics data (Cantacessi et al., 2011; Ranganathan et al., 2007; Strube et
al., 2012) and ongoing research on the genome sequencing of D. viviparus (Martin et al.,
2015) allows the expectation that soon these β1-6- galactosyltransferase(s) and other enzymes
involved in biosynthesis of L3 stage O-glycans will be identified, cloned and characterized. In
addition the fact that certain glycan structures are present in the host does not eliminate those
structures as potential vaccine candidates. The similar rational is used for the development of
carbohydrate-based vaccines against cancer (Astronomo and Burton, 2010) where tumour-
associated carbohydrate antigens (TACA) are present on cancer and other parts of the body
not necessarily accessible to antibodies against TACA. Finally, the synthesis of these unusual
H. contortus and D. viviparus N- and O-glycans as well as testing of their immunological
properties may not only reveal the potential vaccine candidates but also immune-modulators
with unexpected properties and a range of different applications in treating immune disorders.
Alternatively cell lines or organisms can be glycoengineered to produce proteins with the parasite glycan structures. The major advantage of this approach is that once engineered the expression of the desired glycoprotein can be achieved in large quantities. If whole cells are used for vaccination as shown for glycoengineered Lec 8 Chinese Hamster Ovary (CHO) cells expressing poly- GalNAcβ1-4GlcNAc (poly-LDN) and poly- GalNAcβ1-4(Fucα1-3)GlcNAc (poly-LDNF) (Prasanphanich et al., 2014) there is no need for protein purification and immunization material can be obtained even easier. On the other hand the major drawback for this approach when compared to enzymatic in vitro synthesis is that cell engineering provides an extra layer of complexity as not only expression of desired glycan-modifying enzymes in the ER or Golgi has to be considered, but also the availability of corresponding sugar donors as well as the deletion of undesired glycosyltransferases/glycosidases. For the creation of the cell line which would produce adult *H. contortus* N-glycans the *C. elegans* organism as an expression system would be the logical candidate. Since both *C. elegans* and *H. contortus* seem to share the same chitobiose core modifications, only non-reducing end antennae have to be changed. The *C. elegans* non-reducing antennae are known to contain single GlcNAc residue as well as mannose and galactose residues which can be modified with fucose (see section 1.2.2.5). To create
GalNAcβ1-4(Galα1-2Fucα1-3)GlcNAc structures which have been suggested to be present in *H. contortus*, GalNAc transferase has to be introduced into the Golgi of *C. elegans*. Since no GalNAc antennae have been observed in *C. elegans* N- or O-glycans, it is likely that the GalNAc-UDP transporter would have to be inserted into the Golgi as well. If needed the levels of the GalNAc-UDP could be enhanced by the insertion of the GlcNAc-UDP epimerase WbgU (Su et al., 2008). Likely α1-3 fucosyltransferase FUT-6 and putative α1-2 fucosyltransferase forming the Galα1-2Fuc structure on the 3-position of the distal chitobiose core GlcNAc would also be able to modify the GalNAcβ1-4GlcNAc antennae. If not, the insertion of additional putative fucosyltransferase and galactosyltransferase has to be performed as well. Since both fucose and galactose are found on the chitobiose core of *C. elegans* there would be no need to insert GDP-Fuc and UDP-Gal transporters. If homogenous N-glycosylation is needed the putative fucosyltransferase(s) and galactosyltransferase enzymes modifying *C. elegans* non-reducing end N-glycans would have to be deleted. These glycoengineered *C. elegans* could be used to make recombinant proteins with *H. contortus* N-glycans as vaccine candidates or whole worms could be tried as well assuming that other molecules present in *C. elegans* would not interfere with immunization.

In attempt to glycoengineer the cell line which would express unusual L3 stage *D. viviparus* O-glycans, already existing CHO mutant cell lines can be considered. For example Lec2 CHO cell line has been demonstrated to contain only core 1 Galβ1-3GalNAc O-glycans without any structures containing sialic acid (North et al., 2010a). Putative β1-6 galactosyltransferase(s) would have to be inserted into Golgi to form the poly- β1-6-galactose branch. Putative α-galactosyltransferases as well as α1-2 fucosyltransferase enzymes would have to be inserted into Golgi as well. Since galactose is present in both N- and O-glycans of this cell line as well as fucose is present in N-glycans (North et al., 2010a) no additional sugar donor transporters would be required. The proteins with O-glycosylation site(s) could be expressed using that system and used as potential vaccine candidates. Alternatively if human N-glycans present on cell surface would not interfere with cow immunization studies, whole cells could be used as vaccination material as shown for immunization studies with glycoengineered CHO cells expressing schisto parasite antigens (Prasanphanich et al., 2014).

The discovered unusual glycans can also be used to raise antibodies against them. These antibodies could be used in a range of scientific applications such as staining the worm to find the exact position of the antigens within the worm as well as antibody-based antigen
purification. Similarly the antibodies raised against parasite-specific glycans can be used in the development of tests to diagnose the worm infection. Currently apart from general symptoms and faecal examination the *H. contortus* parasite infection is detected by two tests. The Dipstick Test was developed in Australia and is based on blood detection in dung (Colditz and Le Jambre, 2008). Second test was developed later by University of Georgia, Athens and is based on staining parasite eggs with fluorescein-labeled peanut agglutinin and examination of faeces under the microscope in the presence of ultraviolet light (Jurasek et al., 2010). Although the second test is less expensive and less labour intensive it still takes two days to perform. Having an antibody against glycans found, for example, on excretory-secretory proteins of adult *H. contortus* would potentially enable us to perform the sandwich-ELISA experiment to detect the excretory-secretory antigens in the blood. Alternatively the synthesized glycan antigens from L3 and/or adult parasite would enable us to detect antibodies elicited against them in the host blood. In combination with new technologies such as microfluidics these tests could be performed within minutes as shown for ELISA-based detection of HIV antigen (Chin et al., 2011).

In terms of *D. viviparus* parasite, two tests based on ELISA (Cornelissen et al., 1997) and dipstick immunoasay (Schnieder, 1993) are available to detect the infection. However both of them are based on adult Major Sperm Protein antigen which binds antibodies circulating in cow’s blood or milk. Two current problems are associated with these tests. The first problem is that it takes a couple of weeks needed for the development of antibodies against adult antigen. The second problem is that these tests do not show the ‘reinfection’ phenomenon where L3 stage parasites accumulate in bronchi, cause strong immune response and eventually are cleared out of bronchi (Sekiya et al., 2013). The discovered unusual O-glycans of L3 stage *D. viviparus* parasite could act as potential antigens to raise antibodies against them for the alternative ELISA-based diagnostics.
References


Jarrett, W.F., Jennings, F.W., Mc, I.W., Mulligan, W., and Urquhart, G.M. (1960a). Immunological studies on Dictyocaulus viviparus infection; active immunization with whole worm vaccine. Immunology 3, 135-144.


