MECHANISTIC STUDIES
ON THE ENZYMES INVOLVED IN THE BIOSYNTHESIS
OF CMP-\(N,N'-\)DIACETYLLEGIONAMINIC ACID AND UDP-\(d\)-APIOSE

by

Pavel Alexander Glaze

B.Sc., The University of British Columbia, 2003

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

May, 2009

© Pavel Alexander Glaze, 2009
Abstract

This thesis focuses on the biosynthesis of two sugar nucleotides. The enzymes responsible for the biosynthesis of \(N,N'-\text{diacetyllegionaminic acid}\) in \textit{Legionella pneumophila} are identified for the first time. All three genes (\textit{neuA,B,C}) demonstrated sequence homology to the genes involved in sialic acid biosynthesis. The first gene \textit{neuC} encodes a hydrolyzing 2-epimerase, which is found to catalyze the conversion of UDP-\(N,N'-\text{diacetylbacillosamine}\) (UDP-Bac2,4diNAc) into 2,4-diacetamido-2,4,6-trideoxymannose (6-deoxyMandiNAc) and uridine diphosphate (UDP). The incubation of UDP-Bac2,4diNAc with NeuC in deuterated buffer generated \(\alpha\)-[2-\(^2\text{H}\)]-6-deoxyMandiNAc. This indicates that the reaction catalyzed by the hydrolyzing 2-epimerase proceeds with a net retention of configuration at C-1, and that C-2 is deprotonated and reprotonated with a solvent-derived deuterium atom. An enzymatic reaction in H\(_2\)\(^{18}\text{O}\) demonstrated that the loss of UDP occurs through a C-O bond cleavage process. These results support a mechanism involving the \textit{anti}-elimination of UDP, forming a 6-deoxy-2,4-diacetamidoglucal intermediate, followed by a \textit{syn}-hydration, to generate 6-deoxyMandiNAc. \(N,N'-\text{diacetyllegionaminic acid synthase}\) (NeuB) is a potential phosphoenolpyruvate-condensing synthase involved in the biosynthesis of \(N,N'-\text{diacetyllegionaminic acid}\) (Leg5Ac7Ac). This enzyme is proposed to catalyze the condensation of phosphoenolpyruvate (PEP) and 6-deoxyMandiNAc to form Leg5Ac7Ac and phosphate. NMR spectroscopic analysis confirmed that NeuB is an acid synthase and that the \(N,N'-\text{diacetyllegionaminic acid}\) product has the D-glycero-D-galacto configuration. Incubation with [2-\(^{18}\text{O}\)]-PEP demonstrated that NeuB operates via a C-O bond cleavage mechanism. Finally, the NeuA homolog was demonstrated to possess...
CMP-$N,N'$-diacetylleagionaminic acid synthetase activity generating CMP-Leg5Ac7Ac, which is activated for use in lipopolysaccharide biosynthesis.

UDP-D-apiose is biosynthesized from UDP-D-glucuronic acid by a bifunctional enzyme UDP-D-apiose/UDP-D-xylose synthase (AXS1). NMR spectroscopic analysis confirmed that AXS1 produces a roughly 1:1 mixture of UDP-D-apiose (UDP-Api) and UDP-D-xylose (UDP-Xyl). Incubation of a potential reaction intermediate, UDP-4-ketoxylose, resulted in the slow formation of either UDP-Xyl and possibly UDP-Api. AXS1 catalyzed the formation of exclusively UDP-2-deoxy-2-fluoroxylose when incubated with UDP-2-deoxy-2-fluoroglucuronic acid, while fluoride, UDP and CO$_2$ was formed when AXS1 was incubated with UDP-3-deoxy-3-fluoroglucuronic acid. The enzymatic incubation with these substrate analogs provided further evidence for the retro-aldol mechanism.
<table>
<thead>
<tr>
<th>Chapter One: Sugar Nucleotide-Modifying Enzymes</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Sugar Nucleotides</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Epimerases and Racemases</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Short-Chain Dehydrogenase/Reductase (SDR) Family</td>
<td>5</td>
</tr>
<tr>
<td>1.4 SDR Sugar Nucleotide Epimerases</td>
<td>7</td>
</tr>
<tr>
<td>1.4.1 UDP-Galactose 4-Epimerase (GalE)</td>
<td>7</td>
</tr>
<tr>
<td>1.4.2 GDP-mannose 3,5-Epimerase (GME)</td>
<td>12</td>
</tr>
<tr>
<td>1.4.3 GDP-fucose Synthase (GFS)</td>
<td>14</td>
</tr>
<tr>
<td>1.5 Cofactor-Independent Sugar Nucleotide Epimerases</td>
<td>17</td>
</tr>
<tr>
<td>1.5.1 Epimerases Operating at Substituents Bearing Activated Stereocenters</td>
<td>17</td>
</tr>
<tr>
<td>1.5.2 UDP-N-acetylglucosamine 2-Epimerases</td>
<td>19</td>
</tr>
<tr>
<td>1.6 Sugar Nucleotide Deoxygenation</td>
<td>23</td>
</tr>
<tr>
<td>1.6.1 SDR Sugar Nucleotide 4,6-Dehydratases</td>
<td>23</td>
</tr>
<tr>
<td>1.6.2 Sugar Nucleotide Aminotransferases</td>
<td>25</td>
</tr>
<tr>
<td>1.7 Sugar Nucleotide Dehydrogenases</td>
<td>27</td>
</tr>
<tr>
<td>1.8 Sugar Nucleotide Decarboxylases</td>
<td>29</td>
</tr>
<tr>
<td>1.9 Biosynthesis of Higher Order Sugars</td>
<td>30</td>
</tr>
<tr>
<td>1.9.1 Sialic Acid Biosynthesis</td>
<td>30</td>
</tr>
<tr>
<td>1.9.2 Pseudaminic Acid Biosynthesis</td>
<td>34</td>
</tr>
<tr>
<td>1.9.3 Legionaminic Acid Biosynthesis</td>
<td>36</td>
</tr>
<tr>
<td>1.10 Branched-Chain Sugar Nucleotides</td>
<td>41</td>
</tr>
<tr>
<td>1.10.1 Methylation Reaction in the Biosynthesis of Mycarose</td>
<td>41</td>
</tr>
<tr>
<td>1.10.2 Biosynthesis of UDP-D-Apiose</td>
<td>42</td>
</tr>
<tr>
<td>1.11 Project Goals</td>
<td>51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter Two: Biosynthesis of CMP-N,N'-Diacetyllegionaminic Acid</th>
<th>53</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>54</td>
</tr>
<tr>
<td>2.2 Preparation of UDP-Bac2,4diNAc</td>
<td>55</td>
</tr>
<tr>
<td>2.2.1 Expression and Purification of PglIF, PglE and PglD enzymes</td>
<td>56</td>
</tr>
<tr>
<td>2.2.2 Chemoenzymatic synthesis of UDP-Bac2,4diNAc</td>
<td>57</td>
</tr>
<tr>
<td>2.3 Identification and Mechanistic Studies on the Hydrolyzing UDP-Bac2,4diNAc 2-Epimerase from <em>Legionella pneumophila</em></td>
<td>60</td>
</tr>
<tr>
<td>2.3.1 Expression and Purification of the Hydrolyzing UDP-Bac2,4diNAc 2-epimerase</td>
<td>60</td>
</tr>
</tbody>
</table>
2.3.2 Testing the Activity of the Hydrolyzing UDP-Bac2,4diNAc 2-epimerase .... 61
2.3.3 Kinetic Characterization of the Hydrolyzing UDP-Bac2,4diNAc 2-epimerase...

2.3.4 Stereochemical Analysis and Solvent Isotope Incorporation .................. 63
2.3.5 Test for C-O vs. P-O Bond Cleavage ............................................. 69

2.4 Identification and Mechanistic Studies on $N,N'$-Diacytelylegionaminic Acid Synthase (NeuB) ................................................................. 72
2.4.1 Expression and Purification of $N,N'$-Diacytelylegionaminic Acid Synthase. 73
2.4.2 Test for $N,N'$-Diacytelylegionaminic Acid Synthase (NeuB) Activity .......... 75
2.4.3 Isolation and Characterization of $N,N'$-Diacytelylegionaminic Acid ......... 80
2.4.4 Potential Synthase Mechanisms and Test for C-O vs P-O Bond Cleavage .... 82

2.5 CMP-$N,N'$-Diacytelylegionaminic Acid Synthetase ................................... 85
2.5.1 Identification, Expression and Purification of CMP-$N,N'$-Diacytelylegionaminic Acid Synthetase ................................................................. 85
2.5.2 Activity of CMP-$N,N'$-Diacytelylegionaminic Acid Synthetase .................. 86

2.6 Conclusions .............................................................................................. 87
2.7 Future Directions ....................................................................................... 90

2.8 Experimental ............................................................................................ 93
2.8.1 Materials and General Methods ............................................................. 93
2.8.2 Cloning of L. pneumophila neuA, neuB and neuC .................................. 94
2.8.3 Over-expression and Purification of L. pneumophila NeuA, NeuB and NeuC. 95
2.8.4 Sub-cloning of L. pneumophila neuB for Fusion Protein, Overexpression and Purification. ................................................................. 96
2.8.5 Over-expression and Purification of His-Tagged PglF and PglE .................. 97
2.8.6 Chemo-enzymatic Synthesis of UDP-Bac2,4diNAc .............................. 98
2.8.7 Characterization of Hydrolyzing 2-Epimerase Activity ............................ 99
2.8.7.1 NeuC Homolog Activity Assay ........................................................ 99
2.8.7.2 NeuC Kinetic Studies ................................................................. 100
2.8.7.3 Stereochemistry and Solvent Deuterium Isotope Incorporation Studies 100
2.8.7.4 Metal Dependency of NeuC ........................................................... 101
2.8.8 Test for C-O vs. P-O Bond Cleavage Mechanism ................................. 101
2.8.9 Characterization of $N,N'$-Diacytelylegionaminic Acid Synthase ......... 102
2.8.10 Isolation and Characterization of the $N,N'$-Diacytelylegionaminic Acid 102
2.8.11 Test for C-O vs. P-O Bond Cleavage Mechanism .............................. 103
2.8.12 Characterization of CMP-$N,N'$-Diacytelylegionaminic Acid Synthetase Activity ................................................................. 104

Chapter Three: Mechanistic Studies on UDP-$d$-Apiose Synthase .................. 105
3.1 Introduction ............................................................................................... 106
3.2 Cloning, Expression and Purification of UDP-D-Apiose/UDP-D-Xylose

3.2.1 Preparation of cDNA from Arabidopsis thaliana ........................................... 107
3.2.2 Expression and Purification of AXS1 ............................................................... 109
3.3 Testing the Activity of AXS1 ................................................................. 113
3.3.1 Monitoring the Activity of AXS1 Using 1H and 31P NMR Spectroscopy 113
3.3.2 Preparation of UDP-[U-13C]-D-Glucuronic Acid ........................................... 119
3.3.3 Testing the Activity of AXS1 Using UDP-[U-13C]-Glucuronic Acid ............ 121
3.4 Catalytic Competence of UDP-4-ketoxylose .................................................. 124
3.4.1 Preparation of UDP-4-ketoxylose ............................................................... 124
3.4.2 Reaction of UDP-4-ketoxylose with AXS1 .................................................... 127
3.5 Testing UDP-Xylose as a Potential Substrate ................................................. 132
3.6 Attempted Synthesis of UDP-D-Apiose ........................................................... 133
3.7 Fluorinated Analogs of UDP-GlcA ................................................................. 139
3.7.1 Introduction ......................................................................................... 139
3.7.2 Synthesis of UDP-2-Deoxy-2-fluoro-D-glucuronic Acid ............................... 145
3.7.3 Testing UDP-2F-GlcA as a Substrate Analog ............................................... 147
3.7.4 Synthesis of UDP-3-Deoxy-3-fluoro-D-glucuronic Acid ......................... 152
3.7.5 Testing UDP-3F-GlcA as a Substrate Analog ............................................ 153
3.8 Conclusions and Future Directions ............................................................... 160
3.9 Experimental ................................................................................................. 162
3.9.1 Materials and General Methods ................................................................. 162
3.9.2 Source and Cloning of AXS1 ................................................................. 162
3.9.3 Over-expression and Purification of AXS1 ............................................... 163
3.9.4 Source and Cloning of arnA ................................................................. 164
3.9.5 Over-expression and Purification of ArnA ................................................. 164
3.9.6 Over-expression and Purification of UDP-Glucose Dehydrogenase ........ 164
3.9.7 Monitoring Enzyme Incubation by NMR Spectroscopy ............................ 165
3.9.8 AXS1 Activity Test with UDP-GlcA ......................................................... 166
3.9.9 Enzymatic synthesis of UDP-[U-13C]-Glucuronic Acid ............................ 166
3.9.10 Studies with UDP-[U-13C]-GlcA ............................................................... 167
3.9.11 Enzymatic Synthesis and Isolation of UDP-4-ketoxylose ....................... 168
3.9.12 Test of UDP-4-ketoxylose as a Potential Intermediate for AXS1 ........ 168
3.9.13 AXS1 Activity Test with UDP-Xyl ......................................................... 169
3.9.14 Attempted Synthesis of UDP-Api 10 ....................................................... 169
3.9.14.1 Methyl 3'-benzyl-2,3-O-isopropylidene-β-D-erythro-apifuranose 5 ... 169
3.9.14.2 3'-Benzy1-1,2,3-O-triacetyl-D-erythro-apifuranose 6 ....................... 170
3.9.14.3 3'-Benzy1-D-erythro-apifuranose 12 ................................................. 171
3.9.14.4 1,2,3,3'-Tetra benzyl-D-erythro-apifuranose 13 ............................. 171
3.9.14.5 2,3,3'-Tribenzyl-D-erythro-apifuranose 14 ...................................... 172
3.9.14.6. Dibenzyl 2,3,3'-Tribenzyl-D-erythro-apifuranose phosphate 15 ....... 173
3.9.15 Chemoenzymatic Synthesis of UDP-2F-GlcA 22........................................... 173
3.9.16 AXS1 Incubation with UDP-2F-GlcA ......................................................... 174
3.9.17 Chemoenzymatic Synthesis of UDP-3F-GlcA 28........................................... 175
3.9.18 Chemoenzymatic Synthesis of UDP-[U-13C]-3F-GlcA ................................. 176
3.9.19 AXS1 Incubation with UDP-3F-GlcA and UDP-[U-13C]-3F-GlcA ............. 176

Bibliography ............................................................................................................. 177
Appendix ..................................................................................................................... 190
List of Figures

Figure 1.1 Generic sugar nucleotide. ................................................................. 2

Figure 1.2 Biosynthesis of a sugar nucleotide from an unactivated sugar .............. 3

Figure 1.3 A) Mechanism employed by epimerase catalyzing a reaction at an activated stereocenter. B) Reaction catalyzed by epimerase at an unactivated center. B and BH represent active site basic and acidic residues, respectively. ............................................. 5

Figure 1.4 Positions and roles of the catalytic residues in SDR enzymes. .............. 6

Figure 1.5 Inversion of configuration catalyzed by UDP-galactose 4-epimerase (GalE). . 7

Figure 1.6 Proposed mechanism of the UDP-galactose 4-epimerase mechanism, showing the UDP-4-ketoglucose. .................................................................... 8

Figure 1.7 Solvent isotope incorporation experiment with GalE showing an absence of incorporation. ......................................................................................... 8

Figure 1.8 Tritium incorporation into the ketone intermediate analog of GalE. .......... 9

Figure 1.9 Accidental release of UDP-4-ketoglucose from the active site. Square brackets indicate species bound to GalE active site. .................................................... 10

Figure 1.10 Mechanism of the reaction catalyzed by GalE. .................................... 11

Figure 1.11 Reactions catalyzed by GDP-mannose 3,5-epimerase (GME). .............. 12

Figure 1.12 Mechanism of the reaction catalyzed by GME. .................................... 13

Figure 1.13 Positioning of Cys 145 and Lys 217 in the active site of the GME. .......... 14

Figure 1.14 Biosynthesis of GDP-L-fucose from GDP-α-D-mannose. .................... 15

Figure 1.15 Proposed mechanism for the reaction catalyzed by GDP-L-fucose synthase. .. 16

Figure 1.16 Deuterium washout experiment with His179Gln GFS. ......................... 17

Figure 1.17 The reaction catalyzed by RmlC. ...................................................... 18

viii
Figure 1.18 The reactions catalyzed byNovW and EvaD. ................................................................. 19
Figure 1.19 Non-hydrolyzing and hydrolyzing UDp-GlcNAc 2-epimerases ........................................... 20
Figure 1.20 Reaction and mechanism catalyzed by the bacterial “non-hydrolyzing” UDp-GlcNAc 2-epimerase .................................................................................................................. 21
Figure 1.21 Positional isotope exchange experiment with RffE ................................................................. 22
Figure 1.22 Mechanism of the hydrolyzing UDp-GlcNAc 2-epimerase ...................................................... 23
Figure 1.23 Proposed mechanism for the dTDP-Glc 4,6-dehydratase reaction ............................................ 24
Figure 1.24 dTDP-glucose 4,6-dehydratase (RmlB) catalyzes a stereospecific hydride transfer. 25
Figure 1.25 Mechanism of the reaction catalyzed by the PLP-dependent TylB enzyme. ......................... 26
Figure 1.26 Formation of TDP-3-keto-6-deoxyglucose ........................................................................... 27
Figure 1.27 Proposed mechanism of the UDP-glucose dehydrogenase reaction ....................................... 28
Figure 1.28 Proposed mechanism for the reaction catalyzed by UDP-glucuronic acid decarboxylase ............................................................................................................................... 30
Figure 1.29 Sialic acid and sialic acid-related sugars ................................................................................. 31
Figure 1.30 Biosynthesis of CMP-NeuAc in mammals and bacteria ......................................................... 33
Figure 1.31 Biosynthesis of CMP-pseudaminic acid .............................................................................. 36
Figure 1.32 Structure of legionaminic acid derivatives and sialic acid ..................................................... 37
Figure 1.33 Structure of L. pneumophila LPS, the major component of the outer surface of the outer membrane of Gram-negative bacteria .................................................................................. 38
Figure 1.34 Proposed biosynthesis of CMP-Leg5Ac7Ac ....................................................................... 39
Figure 1.35 Biosynthesis of UDP-Bac2,4diNAc in Campylobacter jejuni ................................................. 40
Figure 1.36 Structure of L-mycarose and proposed mechanism of the reaction catalyzed by TylC 3 .................................................................................................................................................. 42
Figure 1.37 D-Apiose involvement in formation of borate cross-link between two different homogalacturonan chains ................................................................. 43

Figure 1.38 The UDP-D-apiose/UDP-D-xylose synthase reaction ........................................ 44

Figure 1.39 Retro-aldol mechanism for the UDP-D-apiose/UDP-D-xylose synthase reaction. ... 46

Figure 1.40 Fate of the labeled C-3" of UDP-[3"-14C]-glucuronic acid during the reaction catalyzed by UDP-D-apiose/UDP-D-xylose synthase ................................................................. 47

Figure 1.41 Fate of the labeled 4"-3H of UDP-[U-14C, 4-3H]-GlcA during the reaction catalyzed by UDP-D-apiose/UDP-D-xylose synthase ................................................................. 47

Figure 1.42 Determination of stereochemistry in the conversion of UDP-[4"-3H]-GlcA to UDP-[3"-3H]-apiose................................................................................................................. 48

Figure 1.43 Carbon migration mechanism for the UDP-D-apiose/UDP-D-xylose synthase reaction ......................................................................................................................... 50

Figure 2.1 Chemoenzymatic synthesis of UDP-Bac2,4diNAc ................................................. 55

Figure 2.2 SDS-PAGE gel of enzymes involved in the biosynthesis of UDP-Bac2,4diNAc ... 57

Figure 2.3 The 400 MHz 1H NMR spectrum of UDP-Bac2,4diNAc in D2O ................................ 59

Figure 2.4 SDS-PAGE gel showing NeuC purification ............................................................ 61

Figure 2.5 31P NMR spectra monitoring the reaction of UDP-Bac2,4diNAc with NeuC .......... 62

Figure 2.6 Continuous coupled UDP assay used to determine kinetic parameters of the hydrolyzing epimerase ................................................................................................. 64

Figure 2.7 Kinetic plots of initial velocity (v0)/[E]0 vs. UDP-Bac2,4diNAc ............................... 64

Figure 2.8 1H NMR spectra monitoring the reaction of UDP-Bac2,4diNAc with the NeuC in D2O. A) Enzymatic conversion of UDP-Bac2,4diNAc to [2-2H]-6-deoxy-α-MandiNAc, followed
by mutarotation to [2-²H]-6-deoxy-β-MandiNAc and B) ¹H NMR spectra monitoring the enzymatic reaction

Figure 2.9 HMQC experiment with 6-deoxyMandiNAc

Figure 2.10 C-O vs P-O bond cleavage experiment. A) Products from C-O bond cleavage mechanism. B) Products from P-O bond cleavage mechanism

Figure 2.11 ESI-mass spectra of C-O vs. P-O bond cleavage experiment with NeuC

Figure 2.12 Mechanism of the reaction catalyzed by the hydrolyzing UDP-Bac2,4diNAc 2-epimerase

Figure 2.13 SDS-PAGE gel showing MalE-NeuB purification

Figure 2.14 Reaction catalyzed by N,N'-diacetyllegionaminic acid synthase (NeuB) and Fischer projections of 6-deoxyMandiNAc and N,N'-diacetyllegionaminic acid

Figure 2.15 ³¹P NMR spectra monitoring the reaction of PEP and 6-deoxyMandiNAc with MalE-NeuB. A) Before the addition of MalE-NeuB and B) 18 h after the addition of MalE-NeuB

Figure 2.16 Partial ¹H NMR spectra monitoring the incubation of PEP and 6-deoxyMandiNAc with MalE-NeuB A) Before the addition of MalE-NeuB and B) after the addition of MalE-NeuB

Figure 2.17 ¹H NMR spectrum of Leg5Ac7Ac in D₂O

Figure 2.18 Proposed C-O vs. P-O bond cleavage mechanisms for MalE-NeuB. A) C-O bond cleavage mechanism and B) P-O bond cleavage mechanism

Figure 2.19 ³¹P NMR spectra monitoring the incubation of MalE-NeuB with [2-¹⁸O]PEP and 6-deoxyMandiNAc A)Before the addition of MalE-NeuB and B) after the addition of MalE-NeuB

Figure 2.20 SDS-PAGE gel showing NeuA purification
Figure 2.21 Reaction catalyzed by CMP-Leg5Ac7Ac synthetase.............................................. 87
Figure 2.22 Biosynthesis of CMP-Leg5Ac7Ac in Legionella pneumophila.................................. 89
Figure 2.23 Potential tetrahedral intermediate analogs for NeuB............................................... 92
Figure 3.1 Preparation of cDNA from plant mRNA................................................................. 109
Figure 3.2 SDS-PAGE gel showing AXS1 purification........................................................... 111
Figure 3.3 Partial UV spectrum of 147 µM AXS1 in 20 mM Tris-HCl pH 8.0 before and after the addition of sodium borohydride (to give 1.2 mM sodium borohydride)........................... 112
Figure 3.4 $^1$H NMR spectra monitoring the reaction of UDP-GlcA with the AXS1 at different time intervals.................................................................................................................. 116
Figure 3.5 Non-enzymatic decomposition of the UDP-Api to α-D-apio-D-furanosyl 1,2-cyclic phosphate and UMP...................................................................................................................... 117
Figure 3.6 $^{31}$P NMR spectra monitoring the enzymatic conversion of UDP-GlcA to UDP-Api and UDP-Xyl. UDP-Api decomposes to UMP and 1,2 cyclic phosphate after over-night incubation.. ................................................................. 118
Figure 3.7 Preparation of UDP-[U-$^{13}$C]-D-glucuronic acid...................................................... 120
Figure 3.8 $^{13}$C NMR spectra monitoring the enzymatic conversion of UDP-[U-$^{13}$C]-GlcA to UDP-[U-$^{13}$C]-Api and UDP-[U-$^{13}$C]-Xyl................................................................. 122
Figure 3.9 $^{13}$C NMR spectrum of UDP-[U-$^{13}$C]-Api and UDP-[U-$^{13}$C]-Xyl............................ 123
Figure 3.10 Biosynthesis of UDP-4-amino-4-deoxy-L-arabinose in polymyxin-resistant E. coli and Salmonella typhimurium............................................................... 125
Figure 3.11 SDS-PAGE gel showing ArnA purification............................................................. 126
Figure 3.12 Reaction of UDP-4-ketoxylose with H$_2$O and MeOH........................................... 127
Figure 3.13 A) Proposed reaction between UDP-4-ketoxylose and AXS1[NADH]. B) Partial UV spectra of AXS1-bound NADH. Spectra were collected in 10 mM phosphate buffer, pH 8.0, at 37 °C with [AXS1] = 70 μM and [UDP-4-ketoxylose] = 81 μM.

Figure 3.14 Negative ESI-mass spectra monitoring the incubation of AXS1 with UDP-4-ketoxylose.

Figure 3.15 Structures of UDP-Api and UDP-galactofuranose.

Figure 3.16 Open chain form of D-apiose and possible cyclization products.

Figure 3.17 Proposed synthetic route for the formation of UDP-Api.

Figure 3.18 Alternative synthetic route for preparation of anomerically deprotected apiose.

Figure 3.19 Proposed mechanism for the inactivation of CDP-D-glucose 4,6-dehydratase by CDP-6-deoxy-6,6-difluoro-glucose.

Figure 3.20 A) Reaction catalyzed by 1-deoxy-D-xylulose-5-phosphate (DXP) reductoisomerase. B) α-ketol rearrangement mechanism. C) Retro-aldol rearrangement mechanism. D) Fluorinated substrate analogs.

Figure 3.21 Proposed results of the incubation of fluorinated analogs with AXS1.

Figure 3.22 Synthesis of UDP-2-deoxy-2-fluoro-D-glucuronic acid.

Figure 3.23 Partial $^{19}$F NMR spectra monitoring the reaction of UDP-2F-GlcA with AXS1 over a 24 h period.

Figure 3.24 Partial $^1$H NMR spectrum of isolated product form the reaction between UDP-2F-GlcA and AXS1.

Figure 3.25 Proposed reaction between UDP-2F-GlcA and AXS1.

Figure 3.26 Synthesis of UDP-3-deoxy-3-fluoro-D-glucuronic acid.
Figure 3.27 $^{19}$F NMR spectra monitoring the incubation of UDP-3F-GlcA 28 with AXS1 in 50 mM potassium phosphate pH 8.0 (282.4 MHz, D$_2$O) ................................................................. 154

Figure 3.28 Partial $^{13}$C NMR spectra monitoring the reaction of UDP-[$U-^{13}$C]-3F-GlcA with AXS1 over a 24 h period. ................................................................................................. 157

Figure 3.29 Proposed reaction between UDP-3F-GlcA 28 and AXS1 ................................................................. 159

Figure A.1 $^1$H NMR spectrum of 2,4-diacetamido-2,4,6-trideoxymannose (D$_2$O, 300 MHz) ... 191

Figure A.2. $^1$H NMR spectrum of UDP-4-ketoxylose (hydrate form, D$_2$O, 400 MHz) .......... 192

Figure A.3 $^1$H NMR spectrum of 2,3,3'-tribenzyl-D-apiose 14. (CDCl$_3$, 400 MHz) ............ 193

Figure A.4 $^1$H NMR spectrum of UDP-2-deoxy-2-fluoro-D-glucuronic acid as a triethyl-ammonium salt 22. (D$_2$O, 400 MHz) ..................................................................................... 194

Figure A.5 Partial $^1$H COSY spectrum of UDP-2F-Xyl (D$_2$O, 400 MHz) ................................. 195

Figure A.6 $^1$H NMR spectrum of UDP-3-deoxy-3-fluoro-D-glucuronic acid as a triethyl- ...... 196

Figure A.7 $^{13}$C NMR spectrum of UDP-[$U-^{13}$C]-3-deoxy-3-fluoro-D-glucuronic acid as a triethylammonium salt (D$_2$O, 600 MHz) ............................................................... 197
## List of Symbols and Abbreviations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ</td>
<td>chemical shift (ppm)</td>
</tr>
<tr>
<td>Ac</td>
<td>acetyl</td>
</tr>
<tr>
<td>Ac₂O</td>
<td>acetic anhydride</td>
</tr>
<tr>
<td>AcOH</td>
<td>acetic acid</td>
</tr>
<tr>
<td>Ag₂O</td>
<td>silver(I) oxide</td>
</tr>
<tr>
<td>AlcDH</td>
<td>alcohol dehydrogenase</td>
</tr>
<tr>
<td>AldDH</td>
<td>aldehyde dehydrogenase</td>
</tr>
<tr>
<td>arnA</td>
<td>gene encoding the protein ArnA from <em>E. coli</em></td>
</tr>
<tr>
<td>ArnA</td>
<td>bifunctional enzyme involved in the biosynthesis of UDP-N-formyl-4-amino-4-deoxy-L-arabinose.</td>
</tr>
<tr>
<td>ArnB</td>
<td>pyridoxal phosphate-dependent enzyme from <em>E. coli</em></td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>AXS₁</td>
<td>gene encoding AXS₁ in <em>Arabidopsis thaliana</em></td>
</tr>
<tr>
<td>AXS₁</td>
<td>UDP-D-apiose/UDP-D-xylose synthase</td>
</tr>
<tr>
<td>Bn</td>
<td>benzyl</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>CKABP</td>
<td>2-carboxy-3-keto-D-arabinitol 1,5-bisphosphate</td>
</tr>
<tr>
<td>CMP-Leg5Ac7Ac</td>
<td>cytidine 5'-monophospho-N,N'-acetyllegionaminic acid</td>
</tr>
<tr>
<td>CMP-NeuAc</td>
<td>cytidine 5'-monophospho-N-acetylneuraminic acid</td>
</tr>
<tr>
<td>COSY</td>
<td>correlation spectroscopy</td>
</tr>
<tr>
<td>CTP</td>
<td>cytosine triphosphate</td>
</tr>
<tr>
<td>d</td>
<td>doublet (NMR)</td>
</tr>
<tr>
<td>dt</td>
<td>doublet of triplets (NMR)</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>D</td>
<td>deuterium ($\text{^2H}$)</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>dd</td>
<td>doublet of doublets (NMR)</td>
</tr>
<tr>
<td>DMF</td>
<td>$N,N$-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>1,4-dithio-D,L-threitol</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em> strains BL21 (DE3) or JM 109</td>
</tr>
<tr>
<td>ESI-MS</td>
<td>electrospray ionization mass spectrometry</td>
</tr>
<tr>
<td>EvaD</td>
<td>dTDP-3-amino-2,3,6-trideoxy-3-C-methyl-D-erythro-hexopyranos-ulose 5-epimerase</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GalE</td>
<td>UDP-galactose 4-Epimerase</td>
</tr>
<tr>
<td>GDP-Fuc</td>
<td>guanosine 5'-diphospho-L-fucose</td>
</tr>
<tr>
<td>GDP-Man</td>
<td>guanosine 5'-diphospho-D-mannose</td>
</tr>
<tr>
<td>GDP-Gal</td>
<td>guanosine 5'-diphospho-D-galactose</td>
</tr>
<tr>
<td>GFS</td>
<td>GDP-fucose synthase</td>
</tr>
<tr>
<td>GME</td>
<td>GDP-mannose 3,5-epimerase</td>
</tr>
<tr>
<td>GMER</td>
<td>GDP-4-keto-6-deoxy-mannose 3,5-epimerase/4-reductase</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>$N$-acetylglucosamine</td>
</tr>
<tr>
<td>GlcNAc 6-P</td>
<td>$N$-acetylglucosamine 6-phosphate</td>
</tr>
<tr>
<td>H. pylori</td>
<td><em>Helicobacter pylori</em></td>
</tr>
<tr>
<td>HMQC</td>
<td>heteronuclear multiple-quantum coherence NMR spectroscopy</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz ($s^{-1}$)</td>
</tr>
</tbody>
</table>
IPTG  
isopropyl β-D-thiogalactopyranoside

$J$  
coupling constant (NMR); subscripts indicate coupling partners

$k_{cat}$  
catalytic rate constant

$k_{cat}/K_m$  
specificity constant

KDN  
2-keto-3-deoxy-D-glycero-D-galacto-2-nonulosonic acid

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

$K_m$  
Michaelis constant

LB  
Luria-Bertani medium

Leg  
legionaminic acid

Leg5Am7Ac  
5-N-acetimidoyl-7-N-acetyllegionaminic acid

Leg5Ac7Ac  
$N,N'$-diacetyllegionaminic acid

LPS  
lipopolysaccharide

m  
multiplet (NMR)

Man  
mannose

ManNAc  
$N$-acetylmannosamine

ManNAc 6-P  
$N$-acetylmannosamine 6-phosphate

MeOH  
methanol

mRNA  
messenger ribonucleic acid

MW  
molecular weight

$\text{NAD}^+$  
nicotinamide adenine dinucleotide, oxidized form

NADH  
nicotinamide adenine dinucleotide, reduced form

$\text{NADP}^+$  
nicotinamine adenine dinucleotide phosphate, oxidized form

NADPH  
nicotinamine adenine dinucleotide phosphate, reduced form

NaOMe  
sodium methoxide

$\text{neuA}$  
gene encoding the protein NeuA in $L.\ pneumoniae$
NeuA  CMP-Leg5Ac7Ac synthetase from *L. pneumophila*

*neuB*  gene encoding the protein NeuB in *L. pneumophila*

NeuB  Leg5Ac7Ac synthetase from *L. pneumophila*

*neuC*  gene encoding the protein NeuC in *L. pneumophila*

NeuC  hydrolyzing UDP-Bac2,4diNAc 2-epimerase from *L. pneumophila*

NeuAc  *N*-acetylneuraminic acid or sialic acid

*N. meningitidis*  *Neisseria meningitidis*

NMR  nuclear magnetic resonance

NovW  dTDP-6-deoxy-D-xylo-4-hexulose 3-epimerase

OD<sub>600</sub>  optical dispersion at 600 nm

PCR  polymerase chain reaction

PEP  phosphoenolpyruvate

P<sub>i</sub>  inorganic phosphate

PIX  positional isotope exchange

*pglD*  gene encoding PglD in *C. jejuni*

PglD  UDP-6-deoxy-4-amino-GlcNAc acetyltransferase

*pglE*  gene encoding PglE in *C. jejuni*

PglE  UDP-6-deoxy-4-keto-GlcNAc PLP-dependent aminotransferase

*pglF*  gene encoding PglF in *C. jejuni*

PglF  UDP-N-acetyl-glucosamine 4,6-dehydratase

PMP  pyridoxamine phosphate

PLP  pyridoxal phosphate

ppm  parts per million

Pse  pseudaminic acid

PseB  UDP-N-acetylglucosamine 5-inverting 4,6-dehydratase
PseC  UDP-2-acetyl-2,6-dideoxy-β-L-arabino-4-hexulose aminotransferase
PseF  CMP-pseudaminic acid synthase
PseG  UDP-2,4-diacetamido-2,4,6-trideoxy-L-altrose hydrolase
PseH  UDP-2-acetamido-2,4,6-trideoxy-L-altrose N-acetyltransferase
PseI  pseudaminic acid synthase
pyr  pyridine
RffE  non-hydrolyzing UDP-GlcNAc 2-epimerase from E. coli
RmlB  dTDP-glucose 4,6-dehydratase
RmlC  dTDP-4-dehydrorhamose 3,5-epimerase
RG-II rhamnogalacturonan-II
RNA  ribonucleic acid
rpm  revolutions per minute
RT  room temperature
s  singlet (NMR)
[S]  substrate concentration
SAM  S-adenosyl methionine
SDR  short chain dehydrogenase/reductase
SDS-PAGE  sodium dodecylsulfate polyacrylamide gel electrophoresis
Selectfluor™  1-chloromethyl-4-fluoro-1,4-diaziobicyclo[2.2.2]octane bis(tetrafluoroborate)
t  time
TB  terrific broth
TDP  thymidine diphosphate
THF  tetrahydrofuran
Tris  2-amino-2-(hydroxymethyl)-1,3-propanediol
TylA2  TDP-4,6-glucose dehydratase
Ty1B TDP-6-deoxy-4-keto-glucose isomerase
Ty1C3 TDP-2,6-dideoxy-4-ketoaltrose methyl transferase
U enzyme unit
UDP uridine diphosphate
UDP-Api uridine 5'-diphospho-D-apiose
UDP-Bac2,4diNac uridine 5'-diphospho-N,N'-diacetylbacillosamine
UDP-Gal uridine 5'-diphospho-D-galactose
UDP-Glc uridine 5'-diphospho-D-glucose
UDP-GlcA uridine 5'-diphospho-D-glucuronic acid
UDP-2F-GlcA uridine 5'-diphospho-2-deoxy-2-fluoro-D-glucuronic acid
UDP-3F-GlcA uridine 5'-diphospho-3-deoxy-3-fluoro-D-glucuronic acid
UDPGlcDH UDP-glucose dehydrogenase
UDP-GlcNAc UDP-N-acetylglucosamine
UDP-ManNAc UDP-N-acetylmannosamine
UMP uridine 5'-monophosphate
UTP uridine 5'-triphosphate
UV ultraviolet
$V_{\text{max}}$ maximal reaction velocity (rate)
Acknowledgements

Firstly, I would like to thank my supervisor Martin Tanner for supporting me for the past several years. Martin is an excellent professor who has a great deal of patience. I appreciate his taking the time to help me and share his experiences with me. The ‘Life Lessons’ you have passed on to me will not be forgotten. I would also like to thank Dr. Wayne Chou and Dr. James Morrison for assisting me during my first few years in the lab. Also, without the help and support of the current group members I would not be able to complete my projects. Thank you to Timin Hadi for providing me with an excellent breakdown of both the chemistry and sports worlds. Thank you to Feng Liu for helping me with computer related issues, Louis Luk for giving me lessons in microbiology, and Alaine “PD” Mayer for his help with this thesis. I would also like to thank all other group members: Stephen Lau, Xu Li, Jackie Bassiri, Jennifer Griffith, Niusha Mahmoudi and Yanjie Liu for providing me with an excellent environment to work in. I am grateful to Dr. Elena Polishchuk and Jie Chen for accommodating me in Biological Services and lending a helping hand when I needed it, and to Maria Ezhova of the NMR Facility for her assistance. Thank you also to the Withers group and Perrin group for allowing me to have use of the UV/Vis spectrophotometer. Finally, I would like to thank our collaborators Dr. Martin Young and David Watson of NRC for generously donating the recombinant plasmid encoding neuC, neuB and neuA, and thank you to Dr. Reinhard Jetter for donating the plant leaves of Arabidopsis thaliana.
Dedicated to

My wife Sarb, My parents and My sister Anna
Chapter One
Sugar Nucleotide-Modifying Enzymes
1.1 Sugar Nucleotides

Current progress in glycobiology and carbohydrate biochemistry has caused a re-evaluation of the importance of carbohydrates in biology. Generally, carbohydrates are known to be used as an energy source, as structural elements, and as solubilizing parts of glycoproteins. Recent studies have shown that carbohydrates are also involved in biochemical signaling and cell recognition. A sugar nucleotide is a nucleotide linked via a phosphate ester bridge to the anomeric position of a sugar (Figure 1.1). The base portion of the molecule can be either a purine (adenine or guanine) or a pyrimidine (cytosine, uracil or thymine).

![Figure 1.1 Generic sugar nucleotide.](image)

The nucleotide portion of a sugar nucleotide serves to activate the sugar since it acts as an excellent leaving group in a transferase reaction, where the sugar is transferred to a glycosyl acceptor molecule. Moreover, many transferases and sugar-modifying enzymes will use the nucleotide portion of the sugar to differentiate between various activated sugars.

The biosynthesis of sugar nucleotides commonly occurs via one of two distinct pathways. Firstly, the sugar nucleotide may be generated from an unactivated sugar in a process that involves several enzymatic steps (Figure 1.2). A 6-phosphosugar is initially generated by a reaction between adenosine 5'-triphosphate (ATP) and an unactivated sugar, which is catalyzed by a kinase enzyme. The product is then isomerized to a higher energy 1-phosphosugar by the
action of a phosphomutase enzyme. The sugar nucleotide is finally generated by reaction with a nucleotide triphosphate as catalyzed by a pyrophosphorylase enzyme. Sugar nucleotides such as UDP-glucose (UDP-Glc), UDP- N-acetylglucosamine (UDP-GlcNAc) and GDP-mannose (GDP-Man) are all biosynthesized in this fashion.\textsuperscript{4}

\[
\begin{align*}
\text{kinase} &\quad \text{ATP} &\quad \text{ADP} \\
\text{6-phosphosugar} &\quad \text{1-phosphosugar} &\quad \text{sugar nucleotide}
\end{align*}
\]

Figure 1.2 Biosynthesis of a sugar nucleotide from an unactivated sugar.

An alternative pathway to biosynthesize sugar-nucleotides involves the direct modification of pre-existing sugar nucleotides with the use of sugar nucleotide-modifying enzymes. This is an efficient method as it avoids the requirement of additional enzymes to activate each sugar subsequent to its formation. UDP-galactose (UDP-Gal) and GDP-fucose (GDP-Fuc), as well as a large number of less common sugar nucleotides which are produced by plants and microorganisms are produced in this manner.\textsuperscript{5}

This chapter will discuss the best understood sugar nucleotide-modifying enzymes, with an emphasis on the chemical mechanism and the role of key active site residues. Many of the enzymes that will be discussed belong to the short chain dehydrogenase/reductase (SDR) family. A brief introduction of this family will also be presented in this chapter. Finally, the research goals of this thesis will be introduced.
1.2 Epimerases and Racemases

Epimerases and racemases are enzymes that catalyze an inversion of a carbon stereocenter in biological molecules.⁶ Racemases act on substrates with only one asymmetric center, while epimerases catalyze a stereochemical inversion on molecules with multiple asymmetric centers. The inversion of a carbon stereocenter could conceivably occur by breaking and reforming any of the four bonds at that center, however, most racemases and epimerases utilize C-H bond cleavage during catalysis. Epimerases can be divided into two general types.⁶

The first type of epimerase operates at what is generally called an “activated” center. The activated center is a carbon adjacent to a functional group (e.g. a carbonyl) that can stabilize a carbanion. The pKₐ of the proton at these activated stereocenters is relatively low (generally < 30) allowing for a direct deprotonation/reprotonation mechanism to invert stereocenter (Figure 1.3 A). The second type of epimerase catalyzes the reaction at an “unactivated” stereocenter. The pKₐ of the proton at these unactivated stereocenters is generally above 30. As a result, the enzymes are not able to directly deprotonate the stereocenter and must use an alternative pathway.
1.3 Short-Chain Dehydrogenase/Reductase (SDR) Family

The short-chain dehydrogenase/reductase family (SDR) is a large and evolutionarily old family of NAD(P)(H)-dependent enzymes. Of the approximately 3000 known members of the family, 63 have been identified in the human genome, a number similar to that of the ubiquitous P450 enzymes. The functions of SDR enzymes span several categories including oxidoreductases, lyases, and isomerases, with oxidoreductases making up the majority of the forms. Despite having low sequence identity among different enzymes (typically 10-30%), all available three-dimensional structures display a highly similar α/β folding pattern with a Rossman fold near the N-terminus of the peptide.

Most SDR enzymes have a 250-300 residue core structure, with the binding and the active site regions being most conserved. Binding of the NAD(P)(H) cofactor occurs in the N-terminal part of the molecule, where a conserved Gly-X₃-Gly-X-Gly sequence is located. SDR family enzymes are also identified by their active site catalytic triad, that is responsible for promoting a hydride transfer to and from the NAD(P)(H) cofactor and a carbonyl group of the
substrate. The triad consists of a Ser, Tyr and Lys residue, of which Tyr is the most highly conserved residue throughout the whole family. All previous research indicates that the Tyr functions as a catalytic acid that protonates the carbonyl during the reduction (Figure 1.4). The serine is located in close proximity to the carbonyl of the substrate and the hydroxyl group of the Tyr. Ser is responsible for polarizing the carbonyl group and helping to adjust the pK_a of the Tyr. The Lys interacts with the nicotinamide ribose, and lowers the pK_a of the Tyr-OH to promote the proton transfer. A recent discovery suggests that a fourth residue, Asn, could play a role in the catalytic reaction. It is believed that Asn forms an interaction with a water molecule which in turn is involved in the proton relay system. However, considering the enormous spread of SDR enzymes, it is important to note that this mechanism and the role of the active site residues do not have to be identical for all of the SDR enzymes.

Figure 1.4 Positions and roles of the catalytic residues in SDR enzymes.
1.4 SDR Sugar Nucleotide Epimerases

1.4.1 UDP-Galactose 4-Epimerase (GalE)

One of the most heavily studied SDR sugar-nucleotide modifying enzymes is UDP-galactose 4-epimerase (GalE), which catalyzes the interconversion of UDP-galactose and UDP-glucose (Figure 1.5). GalE is found in the Leloir pathway that converts galactose into glucose 1-phosphate.\textsuperscript{14} Deficiencies in any one of the enzymes that are involved in the Leloir pathway in humans result in the diseased state known as galactosemia.\textsuperscript{15} After its initial discovery, the GalE enzyme was referred to as a “Waldenase” enzyme.\textsuperscript{16} At that time, a reaction at carbon that proceeds with an inversion of stereochemical configuration was referred to as a “Walden inversion”.\textsuperscript{17} Thus, it was thought that nucleophilic attack at C-4” by water would displace the C-4” hydroxyl and result in the inverted configuration. Mechanistic studies later disproved this potential mechanism but the name “Waldenase” remained with the enzyme.

![Figure 1.5 Inversion of configuration catalyzed by UDP-galactose 4-epimerase (GalE).](image)

The epimerase isolated from \textit{Escherichia coli} has been the focus of numerous mechanistic and structural investigations. Initial research on GalE focused on the elucidation of the chemical mechanism. The \textit{E. coli} enzyme functions as a homodimer of subunits containing 338 amino acid residues and one tightly bound NAD\textsuperscript{+} cofactor.\textsuperscript{18,19} The NAD\textsuperscript{+} cofactor cannot be removed from the active site without denaturing the enzyme and is regenerated upon the completion of each
reaction cycle. Based on this information, it was suggested that reaction occurs through an oxidation at C-4'' to generate a ketone intermediate (Figure 1.6).

\[
\begin{align*}
NAD^+ & \quad \text{UDP-galactose} & \quad \text{UDP-glucose} \\
& \quad \Rightarrow & \\
\text{UDP-4-ketoglucose} & \quad \Rightarrow & \quad NAD^+
\end{align*}
\]

Figure 1.6 Proposed mechanism of the UDP-galactose 4-epimerase mechanism, showing the UDP-4-ketoglucose.

In order to confirm the proposed mechanism several experiments were performed. One of the earliest experiments performed on the GalE enzyme was using D$_2$O or H$_2^{18}$O as a solvent for the enzymatic reaction.$^{20}$ Results indicated that neither of the labeled atoms was incorporated into the substrate upon extended incubation (Figure 1.7). This showed that the mechanism of the reaction does not involve a simple deprotonation/reprotonation at C-4'' or a dehydration/rehydration sequence.

\[
\begin{align*}
\text{UDP-galactose} & \quad \text{4-epimerase} & \quad \text{UDP-glucose} \\
& \quad \Rightarrow & \\
\text{UDP-4-ketoglucose} & \quad \Rightarrow & \quad \text{H}_2\text{O or H}_2^{18}\text{O}
\end{align*}
\]

Figure 1.7 Solvent isotope incorporation experiment with GalE showing an absence of incorporation.

The first direct evidence for the formation of the UDP-4-ketoglucose intermediate was reported in 1971. In that experiment Kirkwood initially treated GalE with NaB$_3$H$_4$ which produced a reduced form of the enzyme containing NAD$_3^3$H as the cofactor.$^{21}$ The incubation of
the enzyme with previously prepared UDP-4-keto-6-deoxyglucose produced a mixture of UDP-6-deoxy-[4-3H]-glucose and UDP-6-deoxy-[4-3H]-galactose (Figure 1.8). Further evidence supporting the direct C-4'' oxidation mechanism comes from the observation of a primary kinetic isotope effect.\textsuperscript{22, 23} Using UDP-[4-3H]-hexoses, a primary kinetic isotope effect was observed, indicating that the C-4'' C-H bond is broken during catalysis.

![Chemical Structures]

Figure 1.8 Tritium incorporation into the ketone intermediate analog of GalE.

Lastly, it was reported that the extended incubation of UDP-glucose with GalE leads to the accumulation of UDP-4-ketoglucose and inactive enzyme. The inactivation occurred due to the occasional release of the 4-keto intermediate that leaves the free enzyme bearing a tightly bound NADH cofactor (Figure 1.9). The reduced form of the enzyme will then bind either UDP-glucose or UDP-galactose to form an abortive complex where epimerization cannot occur.\textsuperscript{24, 25} The intermediate was trapped by reduction with NaB\textsubscript{3}H\textsubscript{4} to generate UDP-[4-3H]glucose and UDP-[4-3H]galactose.\textsuperscript{26}
Interesting aspects of the GalE reaction still remained unsolved even after the chemical mechanism of the reaction had been elucidated. Unlike other dehydrogenases in the SDR family, which catalyze stereospecific hydride transfers to and from one particular face of a molecule, GalE reduces the ketone intermediate without a stereochemical preference. The hydride transfer is non-stereospecific with respect to the substrate. This is not the case with the NAD$^+$ cofactor, since the hydride always transfers to the pro-$S$ position on the $B$ side of the nicotinamide ring. In order for enzymatic reaction to occur there are two possible scenarios: i) the ketone intermediate is released into the solution and is rebound with the other face of the carbonyl exposed to the hydride, or ii) the ketone intermediate is retained within the active site and undergoes a dramatic
conformational reorientation. Based on all previous research, it was hard to imagine that the active site of the enzyme could accommodate full reorientation of the substrate.

In order to address the phenomenon, an isotopic crossover experiment was performed. An incubation of GalE with UDP-glucose-\(d_7\) and UDP-glucose-\(d_6\) was monitored by mass spectrometry.\(^{27}\) The results showed that no isotopic scrambling that would form \(d_1\) or \(d_6\) species occurred during the reaction, indicating that the substrate remains in the active site of the enzyme throughout the catalytic cycle. This means that epimerization occurs within a single active site with one NAD\(^+\) cofactor controlling the reaction (Figure 1.10). It was proposed that a 180\(^0\) rotation of the keto intermediate occurred within the active site and enabled both faces of the carbonyl to be exposed to the static cofactor.\(^{28}\)

Figure 1.10 Mechanism of the reaction catalyzed by GalE.

In the late 1990’s, X-ray crystallography was used to help identify key residues responsible for the enzymatic reaction and to verify the proposed conformational reorganization. The crystal structures of \(E. coli\) GalE revealed the location of the three conserved residues of
SDR enzymes: Lys 153, Tyr 149 and Ser 124. Lys 153 was found to be important for enhancing the chemical activity of NAD$^+$ upon the binding of uridine nucleotides in the active site.\textsuperscript{18,29} The Tyr and Ser residues were responsible for deprotonating the C-4’’ hydroxyl group during the hydride transfer. Ultimately, the crystal structures of inactive, mutant forms of GalE in complexes with either UDP-Glc or UDP-Gal confirmed the proposed reorientation mechanism utilized by the enzyme.\textsuperscript{30}

### 1.4.2 GDP-mannose 3,5-Epimerase (GME)

GDP-mannose 3,5-epimerase (GME) catalyzes the interconversion of GDP-\(\alpha\)-D-mannose with GDP-\(\beta\)-L-galactose and GDP-\(\beta\)-L-gulose (Figure 1.11).\textsuperscript{31} The enzyme was first discovered in the green alga \textit{Chlorella pyrenoidosa}.\textsuperscript{32} The enzyme is unique among sugar epimerases because it is able to invert the configuration at two distinct “unactivated” stereocenters using a single active site.\textsuperscript{33} Even though the term “epimerase” is commonly used to describe this enzyme, GME is not an epimerase in the strict sense because GDP-\(\alpha\)-D-mannose and GDP-\(\beta\)-L-galactose are diastereomers and not epimers. The formation of GDP-\(\beta\)-L-galactose is believed to be the first committed step in the biosynthesis of L-ascorbic acid (vitamin C), whereas the fate of GDP-\(\beta\)-L-gulose is still unclear. However, it has been postulated that L-gulose and L-gulo-1,4-lactone could also be involved in the biosynthesis of vitamin C.\textsuperscript{34}

![Figure 1.11 Reactions catalyzed by GDP-mannose 3,5-epimerase (GME).](image-url)
In order to elucidate the catalytic mechanism of GME, an enzymatic incubation was carried out in tritium-labeled water. It was found that solvent-derived tritium was incorporated at the C-3" and C-5" positions of the sugar skeleton. Given that GME is a member of the SDR family and bears a tightly bound NAD\(^+\) cofactor, this finding indicates that epimerization proceeds through ene(di)ol intermediates. The mechanism of the reaction is thought to occur via a transient oxidation at C-4" to give a ketone intermediate and tightly bound NADH (Figure 1.12). The formation of an "activated" center at C-4" lowers the pK\(_a\) of the protons at C-3" and C-5" and, as a result, the reaction can proceed via a deprotonation/reprotonation mechanism at each carbon center. A final reduction at C-4" regenerates the alcohol and the NAD\(^+\) cofactor.

Figure 1.12 Mechanism of the reaction catalyzed by GME.
The structures of complexes between GME from *Arabidopsis thaliana* and different GDP-hexoses bound in the active site were solved. The results indicated that GME has the classical SDR family fold, with all three important catalytic residues present. The mutant GME Tyr174Phe was not able to catalyze the epimerization, indicating that Tyr 174 is involved in the oxidation of C-4". Since the formation of GDP-d-altrose (C-3" epimerization only) was not observed, it was proposed that GME epimerizes C-5" first and C-3" second. The two active site residues located in appropriate positions to act as acid/base catalysts in the epimerization steps are Cys 145 and Lys 217 (Figure 1.13). In the forward direction, Cys 145 is located below the sugar and is thought to act as a base to abstract the proton at C-5", while Lys 217 position above the sugar and delivers the new proton to C-5". The sugar is then believed to undergo a slight movement in the active site to expose the C-3" center for the inversion reaction.

![Figure 1.13 Positioning of Cys 145 and Lys 217 in the active site of the GME.](image)

### 1.4.3 GDP-fucose Synthase (GFS)

GDP-L-fucose is biosynthesized in two enzymatic steps from the precursor GDP-α-D-mannose. The first step is catalyzed by GDP-mannose 4,6-dehydratase (GMD) which converts GDP-α-D-mannose into GDP-6-deoxy-4-keto-mannose (Figure 1.14). The mechanism of
this class of enzymes will be discussed later in this chapter. The second step of the biosynthesis is catalyzed by the SDR enzyme, GDP-fucose synthase (GFS). GFS is able to catalyze three distinct reactions in a single active site. It inverts the configuration at both C-3" and C-5" of the GDP-6-deoxy-4-keto-mannose, and then carries out an NADPH-dependent reduction of a ketone at C-4" to give GDP-L-fucose.\(^2\) L-Fucose is found in many important glycoconjugates in both prokaryotes and eukaryotes.\(^{38}\)

![Biosynthesis of GDP-L-fucose from GDP-\(\alpha\)-D-mannose.](image)

Figure 1.14 Biosynthesis of GDP-L-fucose from GDP-\(\alpha\)-D-mannose.

The mechanism of GFS has recently been elucidated (Figure 1.15).\(^{39}\) Due to the presence of an activated center at C-4", the stereochemical inversions occur via a deprotonation/ reproto- nation mechanism. A subsequent reduction of the C-4" carbonyl by NADPH generates the product. Based on X-ray crystallography data obtained with \textit{E. coli} GFS, Cys 109 and His 179 were identified as the potential acid/base residues required to perform the epimerizations.\(^{36,40}\)
The Cys109Ser mutant produced GDP-6-deoxy-D-altrose as the major product, indicating that C-3\(^\text{"}\) epimerization occurs first. The premature reduction of the GDP-4-keto-6-deoxy-D-altrose intermediate led to the formation of this product. Moreover, the His179Gln mutant was inactive towards normal catalysis, yet catalyzed the wash-out of deuterium label from the C-3\(^\text{"}\) position of \([3^n-^2\text{H}]\)-GDP-6-deoxy-4-keto-mannose (Figure 1.16).\(^{39}\) This result indicates that the mutant is properly folded and able to carry out the first step of catalysis. The unmodified Cys109 can act as a base and removes the deuterium from C-3\(^\text{"}\) to generate the enol(ate) intermediate. The inversion of configuration does not take place because the acidic His179 was mutated,
however, an exchange of the Cys109 thiol deuterium with the bulk solvent can occur and results in the delivery of a proton back to the C-3" position.

![Deuterium washout experiment with His179Gln GFS.](image)

Figure 1.16 Deuterium washout experiment with His179Gln GFS.

As a result, it is believed that GFS first performs an epimerization at C-3", followed by epimerization at C-5". This is unlike GME which carries out the inversions in the reverse order.33 The reduction step catalyzed by GFS uses the same catalytic triad as the other SDR family enzymes, with Tyr 136 acting as acid to deliver the proton to the oxygen at C-4" (Figure 1.15).

1.5 Cofactor-Independent Sugar Nucleotide Epimerases

1.5.1 Epimerases Operating on Substituents Bearing Activated Stereocenters

As discussed earlier, most SDR enzymes carry out epimerization by first creating an activated center next to a carbon which will undergo an inversion. There are few examples of enzymes that do not rely on the presence of an NAD(P)⁺ cofactor. The classic example of this is dTDP-4-dehydrohamnose 3,5-epimerase (RmlC).41,42 This enzyme catalyzes the second step in the biosynthesis of dTDP-rhamnose and represents a potential drug target against bacterial infections. RmlC catalyzes the inversion of configuration at two stereocenters of dTDP-6-deoxy-d-xylo-4-hexulose to produce dTDP-6-deoxy-L-hyxo-4-hexulose (Figure 1.17). Despite the
seeming similarity between GME/GFS and RmlC reactions, the enzymes share no sequence or
structural similarity, as no redox chemistry is required in the latter case.

\[
\begin{align*}
\text{dTDP-6-deoxy-D-xylo-4-hexulose} & \quad \overset{\text{RmlC}}{\xrightarrow{\text{3,5-epimerase}}} \quad \text{dTDP-6-deoxy-L-lyxo-4-hexulose} \\
\end{align*}
\]

Figure 1.17 The reaction catalyzed by RmlC.

The incubation of RmlC with dTDP-6-deoxy-D-xylo-4-hexulose in deuterated solvent
(D\textsubscript{2}O) resulted in the incorporation of deuterium at C-3” and C-5”.\textsuperscript{43} This provides evidence for a
direct deprotonation/reprotonation mechanism. This conclusion is not surprising, considering
that both carbons are activated by their location next to a ketone at C-4”. Investigations of RmlC
from \textit{Streptococcus suis} have identified a histidine residue that acts as a base and a Tyr that acts
as an acid in both epimerization steps.\textsuperscript{44}

There are two other enzymes that are closely related to RmlC (Figure 1.18). One is
dTDP-6-deoxy-D-xylo-4-hexulose 3-epimerase (NovW) and the other is dTDP-3-amino-2,3,6-
trideoxy-3-C-methyl-D-\textit{erythra}-hexo-pyran-4-uloze 5-epimerase (EvaD), which catalyze
cofactor-independent epimerizations at C-3” and C-5”, respectively.\textsuperscript{45, 46} Their chemical
mechanisms are believed to involve deprotonation/ reprotonation steps, similar to RmlC.
1.5.2 UDP-N-Acetylgalactosamine 2-Epimerases

UDP-N-Acetylgalactosamine 2-epimerases are interesting examples of enzymes that act at an “unactivated” stereocenter. There are two known classes of UDP-GlcNAc 2-epimerases. An example of a “non-hydrolyzing” UDP-GlcNAc 2-epimerase (RffE in E. coli), which catalyzes the reversible epimerization of UDP-GlcNAc and UDP-ManNAc, exists only in bacteria (Figure 1.19). A “hydrolyzing” UDP-GlcNAc-2-epimerase is involved in the biosynthesis of sialic acid. This enzyme also hydrolyzes the glycosidic bond such that the release of UDP accompanies the inversion of configuration (Figure 1.19). The latter reaction is not a true epimerization as it is irreversible and the substrate and products are not epimers. Both “non-hydrolyzing” and hydrolyzing epimerases share sequence homology and both catalyze similar reactions on the same substrate, so it is reasonable to expect that they utilize similar catalytic mechanisms.
The unusual mechanism of the “non-hydrolyzing” epimerase is believed to involve an anti-elimination that generates 2-acetamidoglucal and UDP, followed by a syn-addition of the UDP group to generate the final product (Figure 1.20). Initial evidence for this mechanism was provided by performing the enzymatic incubation in D₂O. It was observed that deuterium was incorporated into the sugar at C-2”, forming [2”-²H]UDP-GlcNAc and [2”-²H]UDP-ManNAc. Moreover, a primary kinetic isotope effect of 1.8 was observed when rates of reactions with [2”-²H]UDP-GlcNAc and unlabeled UDP-GlcNAc were compared. This supports the notion that the C-2” C-H bond is broken during the enzymatic reaction.
Figure 1.20 Reaction and mechanism catalyzed by the bacterial “non-hydrolyzing” UDP-GlcNAc 2-epimerase.

A positional isotope exchange (PIX) experiment provided further evidence for this mechanism (Figure 1.21). UDP-GlcNAc was prepared with an $^{18}$O label at the anomeric position. During the enzymatic incubation with RffE it was observed that the $^{18}$O label was scrambled between the bridging and non-bridging positions of the UDP. This scrambling indicates that the anomeric C-O bond is broken during catalysis and the β-phosphate has time to rotate such that any of the three chemically equivalent oxygen atoms could reform the C-O bond. Finally, it was found that, upon an extended incubation of RffE and UDP-GlcNAc, the intermediates 2-acetamidoglucal and UDP are released into solution. It turns out that this process is thermodynamically favorable so the intermediates accumulate during the incubation and can be detected.
The "hydrolyzing" UDP-GlcNAc 2-epimerase is believed to use a very similar mechanism (Figure 1.22). Running the reaction in D$_2$O led to the incorporation of deuterium into the C-2 position of ManNAc. In addition, incubation in H$_2^{18}$O led to the incorporation of $^{18}$O label at the anomeric carbon of ManNAc. Moreover, when $[1-^{18}$O]$\text{UDP-GlcNAc}$ with $^{18}$O in the bridging position between the anomeric carbon and the $\beta$-phosphorus was incubated with the hydrolyzing UDP-GlcNAc 2-epimerase, the products were $\alpha$-ManNAc and $^{18}$O-labeled UDP. The reaction mechanism is thought to be initiated by the anti-elimination of UDP, followed by the syn-hydration of 2-acetamidoglucal, forming $\alpha$-ManNAc.
1.6 Sugar Nucleotide Deoxygenation

1.6.1 SDR Sugar Nucleotide 4,6-Dehydratases

Sugar nucleotide 4,6-dehydratases catalyze the formation of 4-keto-6-deoxy sugar nucleotides via an overall elimination of water (Figure 1.23). These dehydratases use a transient oxidation reaction that requires the use of an NAD$^+$ cofactor.\(^5\) This is reminiscent of the epimerases discussed earlier which create an "activated" center in the molecule through transient oxidation of a hydroxyl group. Sugar nucleotide 4,6-dehydratases are required in the biosynthesis of all 6-deoxyhexoses, which are the most abundant naturally occurring deoxy sugars after 2-deoxy-D-ribose.\(^{51}\) Three of the most intensely studied sugar nucleotide 4,6-dehydratases are dTDP-glucose 4,6-dehydratase (RmlB), GDP-mannose 4,6-dehydratase (GMD), and CDP-glucose 4,6-dehydratase (CGD). GMD was previously mentioned in the discussion of GDP-fucose biosynthesis (Section 1.4.3). All of the enzymes are SDR family members and are believed to employ a similar mechanism that is exemplified by the RmlB reaction.
The first step of the mechanism is believed to be an oxidation at C-4” to create a ketone intermediate and NADH (Figure 1.23).52, 53 The presence of a ketone at C-4” creates an "activated" center at C-5", which can then undergo a dehydration reaction across the C-5”/C-6” bond. This will generate an α,β-unsaturated ketone, which may then undergo a reduction reaction. A hydride from NADH is delivered to C-6”, and a proton from an active site acid is delivered to C-5” to give the product.53

![Proposed mechanism for the dTDP-Glc 4,6-dehydratase reaction.](image)

One of the first pieces of evidence in support of this mechanism came from running the reaction in either ²H₂O or ³H₂O. It was discovered that solvent-derived isotope label was incorporated at the C-5” position of the product.54, 55 More compelling evidence was provided when dTDP-glucose 4,6-dehydratase was incubated with dTDP-([6S]-[4-²H, 6-³H]-glucose.56 The final product was discovered to be dTDP-4-keto-6-deoxy-(6R)-[6-²H, 6-³H]-glucose, indicating
that Rm1B catalyzes an intramolecular hydride transfer from C-4” to C-6” in a stereospecific fashion (Figure 1.24).

![Chemical Structures]

Figure 1.24 dTDP-glucose 4,6-dehydratase (Rm1B) catalyzes a stereospecific hydride transfer.

1.6.2 Sugar Nucleotide Aminotransferases

Many naturally occurring sugars contain nitrogen atoms as a substituent. Amino sugars are found in glycoproteins, glycolipids and a variety of secondary metabolites. I have already discussed UDP-N-acetylglucosamine (Section 1.5.2), which is probably the most common example of a nitrogen-containing sugar nucleotide. The nitrogen functionality is introduced in early steps of the biosynthesis when fructose 6-phosphate is converted to glucosamine 6-phosphate by glucosamine 6-phosphate synthase. The introduction of nitrogen into sugar nucleotides is somewhat rare and generally occurs by a reaction between a corresponding keto sugar and ammonia derived from an amino acid. This reaction is catalyzed by pyridoxal phosphate (PLP)-dependent enzymes called aminotransferases.

One such sugar nucleotide aminotransferase is found in the biosynthetic pathway of 3-amino-3,6-dideoxyglucose which is a part of the antibiotic tylosin. The reaction between TDP-3-keto-6-deoxyglucose and glutamate, as the source of ammonia, is catalyzed by TylB which is a pyridoxal phosphate (PLP)-dependent enzyme (Figure 1.25). TDP-3-keto-6-deoxyglucose is
generated from TDP-glucose by the sequential action of TDP 4,6-glucose dehydratase (Ty1A2) and an isomerase (Ty1M3) that isomerizes the ketone functionality between C-4" and C-3" (Figure 1.26).

![Chemical structures and reaction mechanism]

Figure 1.25 Mechanism of the reaction catalyzed by the PLP-dependent Ty1B enzyme.

It was reported that, in the reaction catalyzed by Ty1B, an equivalent of pyridoxamine phosphate (PMP) could replace PLP and glutamate to give the final product.\textsuperscript{59} The reaction mechanism is believed to involve first a reaction between PLP and glutamate to form PMP and α-ketoglutarate. Afterwards, the second transamination can occur between PMP and the ketone of the sugar (Figure 1.25). Several other aminotransferase enzymes,\textsuperscript{60, 61} such as ArnB\textsuperscript{62} which catalyzes the formation of UDP-4-amino-4-deoxy-β-L-arabinose, have been reported since the discovery of Ty1B.

26
1.7 Sugar Nucleotide Dehydrogenases

Altering the oxidation state of the sugar nucleotide is an important transformation that is carried out by several enzymes.² The majority of NAD⁺-dependent dehydrogenases catalyze reactions involving a single oxidation step. The introduction of a carboxylate into a substrate is commonly achieved via two sequential oxidations that require two different enzymes. An alcohol dehydrogenase (AlcDH) is normally required to oxidize an alcohol to an aldehyde, and an aldehyde dehydrogenase (AldDH) to oxidize an aldehyde to a carboxylic acid. However, there is a small group of enzymes that can carry out the 2-fold oxidation in a single active site.⁶³ The most thoroughly investigated of these is UDP-glucose dehydrogenase. The enzyme catalyzes the conversion UDP-D-glucose to UDP-D-glucuronic acid. The reaction is believed to occur via a two-fold NAD⁺-dependent oxidation with an aldehyde as an intermediate of the reaction (Figure 1.27).⁶⁴
Figure 1.27 Proposed mechanism of the UDP-glucose dehydrogenase reaction.

The first oxidation occurs at C-6 to generate the aldehyde intermediate, which is rapidly attacked by an active site thiol to give a thiohemiacetal intermediate. A second molecule of NAD\(^+\) is then bound to the enzyme and oxidizes the thiohemiacetal to give a covalently bound thioester intermediate. Finally, hydrolysis of the thioester generates the product. A close examination of all known sequences of the enzyme has indicated the presence of a conserved Cys in the active site of homologs of this enzyme. The Cys260Ser mutant in the *Streptococcus pyogenes* enzyme led to an essentially inactive catalyst. Moreover, a prolonged incubation of the serine mutant resulted in the formation of the corresponding ester intermediate, which was detected by mass spectrometry.\(^{65}\) It appears that the mutant can slowly catalyze both oxidation
steps, however the hydrolysis of the unnatural ester linkage is exceptionally slow, so the adduct accumulates.

1.8 Sugar Nucleotide Decarboxylases

There are many reported examples of enzymes that catalyze decarboxylation reactions, however, there are only a few that are known to perform this reaction on sugar nucleotides. One of these enzymes is UDP-glucuronic acid decarboxylase, which catalyzes the conversion of UDP-D-glucuronic acid (UDP-GlcA) to UDP-D-xylose (UDP-Xyl, Figure 1.28). UDP-Xyl is required for the biosynthesis of proteoglycans, molecules that are found in the extracellular matrix and on the cell surface of animal cells. It is also required in the biosynthesis of plant polysaccharides such as xyloglucan and xylan.

This enzyme is also a member of the SDR family and the reaction is initiated by an NAD\(^+\) dependent oxidation at the C-4” carbon to form a β-ketoacid (Figure 1.28). Decarboxylation can readily occur to form an enol(ate) intermediate which is protonated at C-5” to give UDP-4-keto-xylose. UDP-Xyl is finally formed by the delivery of hydride from NADH back to the C-4” ketone. The incubation of UDP-[4-\(^3\)H]-glucuronic acid with UDP-glucuronic acid decarboxylase, isolated from Cryptococcus laurentii, demonstrated that tritium label was retained in the product during the decarboxylation reaction. Later, a primary kinetic isotope effect was observed with UDP-[4”-\(^3\)H]-glucuronic acid, but not the with C-3” or C-5” labeled substrates, indicating that the oxidation at C-4” is a rate limiting step. It was also found that incubation of UDP-[5”-\(^3\)H]-glucuronic acid with the UDP-glucuronic acid decarboxylases derived from C. laurentii and wheat germ led to the inversion of configuration at C-5” after
The other known UDP-glucuronic acid decarboxylase is involved in biosynthesis of UDP-D-apiose (UDP-API) and will be discussed later in this chapter.

![Proposed mechanism for the reaction catalyzed by UDP-glucuronic acid decarboxylase.](image)

**Figure 1.28** Proposed mechanism for the reaction catalyzed by UDP-glucuronic acid decarboxylase.

### 1.9 Biosynthesis of Higher Order Sugars

The term “higher order sugars” is commonly used to refer to carbohydrate structures bearing more than six linear carbon atoms. While many higher order sugars are biosynthesized from “unactivated” precursors, others, such as sialic acids, are generated from already existing sugar nucleotides.

#### 1.9.1 Sialic Acid Biosynthesis

Sialic acids are nine carbon α-keto acids that were first discovered in the 1930s. Since then major efforts have focused on identifying the structures and the functional roles that these carbohydrates play. With over 50 different derivatives of N-acetylneuraminic acid (NeuAc) and 2-keto-3-deoxy-D-glycero-D-galacto-2-nonulosonic acid (KDN, Figure 1.29), sialic acids
represent a large portion of carbohydrates found on the surface of eukaryotic cells. To date the known derivatives of sialic acid include modifications at the C-4, C-7, C-8 and C-9 carbons with lactate, phosphate, sulphate, methoxy ethers and acetate groups. Molecules such as 5,7-diacetamido-3,5,7,9-tetrahydroxy-L-glycero-L-manno-nonulosonic acid (pseudaminic acid, Pse) and 2-keto-3-deoxy-D-manno-2-octulosonic acid (KDO) are examples of molecules that are structurally similar to sialic acids yet differ in stereoconfiguration, substitution patterns or carbon chain length.

![Chemical structures of sialic acid, pseudaminic acid, KDO, and KDN](image)

Figure 1.29 Sialic acid and sialic acid-related sugars.

In mammals, sialic acids are typically found capping the non-reducing ends of many cell-surface glycan chains. These serve important functions in regulating cellular events, such as cellular recognition, adhesion processes and cell development. An interesting use of sialic acid derivatives was reported during the investigation of red blood cells (erythrocytes). During their production, the cell surfaces are initially polysialylated. Over the span of the erythrocyte life time (120 days) the sialic acids are removed by sialidases or spontaneously in solution. Near
the end of the life cycle host macrophages recognize, bind and phagocytose the exposed erythrocytes. This demonstrates that sialylation of mammalian cells is necessary for the host immune system to recognize its own tissues. Many viruses use sialic acids to facilitate an attachment to the host cell. For example, influenza viruses demonstrated almost mandatory dependence on the host cell surface sialic acids for infection.

The presence of sialic acids in bacterial cells also has been reported. The display of sialic acid on the surface of pathogenic bacteria is believed to lead to evasion and protection from the host immune system through molecular mimicry of host sialylated cells. *Neisseria meningitidis* and *E. coli K1* are responsible for causing meningitis in humans. Closer inspection of their cell surfaces have shown the presence of an α-(2→8)-linked *N*-acetylneuraminic acid polymer similar to the one found on the surfaces of mammalian cells.

Some of the enzymes discussed earlier in this chapter are involved in the biosynthesis of NeuAc. Since NeuAc is found in both mammals and bacteria, both biosynthetic pathways will be presented (Figure 1.30).
Figure 1.30 Biosynthesis of CMP-NeuAc in mammals and bacteria.

In mammals, the biosynthesis of NeuAc is initiated by the bifunctional enzyme, UDP-\(N\)-acetylglucosamine 2-epimerase/\(N\)-acetylmannosamine kinase (Figure 1.30, outer pathway).\(^{49}\) This enzyme first acts as a hydrolyzing 2-epimerase that catalyzes an inversion at C-2" and the release of UDP from UDP-\(N\)-acetylglucosamine (UDP-GlcNAc) to form \(N\)-acetylmannosamine (ManNAc). It then phosphorylates ManNAc at C-6, producing \(N\)-acetylmannosamine 6-phosphate (ManNAc 6-P). The following enzyme in the pathway, NeuAc 9-P synthase, introduces the three additional carbons by way of a reaction between phosphoenolpyruvate (PEP) and ManNAc 6-P. This forms \(N\)-acetylneuraminic acid 9-phosphate (NeuAc 9-P), the 9-carbon precursor of NeuAc. The mechanism of this reaction will be presented in Chapter 2. Free NeuAc
is formed by dephosphorylation at C-9 via N-acetylneuraminic acid phosphatase (NeuAc 9-P phosphatase). While many other monosaccharides are activated as uridine or guanine diphosphates (UDP-Glc, UDP-Gal and GDP-Man), sialic acids are activated as cytidine mononucleotides. This is accomplished by CMP-NeuAc synthetase that converts NeuAc into CMP-N-acetylneuraminic acid (CMP-NeuAc) using CTP.

In bacteria, the biosynthetic pathway for sialic acid differs from the mammalian version (Figure 1.30, inner pathway). UDP-GlcNAc is first converted to ManNAc by the action of a hydrolyzing 2-epimerase (NeuC). The formation of ManNAc 6-phosphate is not observed. Instead, the NeuAc synthase enzyme (NeuB) catalyzes the direct reaction between PEP and unphosphorylated ManNAc to form NeuAc. Lastly, the bacterial CMP-NeuAc synthetase (NeuA), catalyzes the coupling of NeuAc and CTP to form CMP-NeuAc. In some ways the biosynthesis of NeuAc in bacteria is more efficient, as it avoids the production of extra intermediates.

1.9.2 Pseudaminic Acid Biosynthesis

Pseudaminic acid (Pse) was initially discovered in Pseudomonas aeruginosa as a modified version of 2-keto-3-deoxy-nonulosonic acid. The actual structure was determined to be 5,7-diacetamido-3,5,7,9-tetrahydroxy-L-glycero-L-manno-nonulosonic acid (Figure 1.29). More recently this sugar was discovered on flagella of the pathogenic Gram-negative bacteria Campylobacter jejuni and Helicobacter pylori. C. jejuni, commonly found in animal feces, is a major cause of gastroenteritis in North America, whereas H. pylori has been reported to cause peptic ulcers and gastric cancer in some cases. Pseudaminic acid is found as an O-linked post-translational modification of the flagellar proteins of the bacteria. Mutant strains unable to
biosynthesize Pse have been demonstrated to be aflagellate, non-motile and non-pathogenic. Therefore, the presence of the Pse is absolutely required for the proper assembly of flagella. \(^{78}\)

The biosynthetic pathway of Pse has now been fully elucidated (Figure 1.31). \(^{79, 80}\) The first step of the biosynthesis is catalyzed by the dehydratase PseB, which forms UDP-2-diacetamido-2,6-dideoxy-\(\beta\)-L-arabino-hex-4-ulose (UDP-6-deoxy-4-keto-L-IdoNAc) from UDP-GlcNAc. This enzyme is a sugar nucleotide 4,6-dehydratase, yet it is unique among this family of enzymes in that it also promotes an inversion of configuration at C-5\(^{-}\). In the second step, a PLP-dependent aminotransferase PseC introduces an amino group at C-4\(^{-}\) to form UDP-4-amino-2,4,6-trideoxy-AltNAc. An acetyl group is introduced in the next step by the acetyltransferase enzyme PseH to give UDP-6-deoxy-AltdiNAc. The acetyltransferase enzyme employs acetyl coenzyme A (acetyl-CoA) as the acetyl source. Acetyl-CoA contains a primary thiol group which is attached to the acetyl group being transferred via a thioester linkage. The next step in the biosynthesis of Pse is the hydrolysis of the UDP linkage by PseG, which forms 2,4-diacetamido-2,4-6-trideoxy-L-altrose (6-deoxy-AltdiNAc). The additional three carbons are finally introduced by a reaction between PEP and 6-deoxy-AltdiNAc, which is catalyzed by Pse synthase (PseI). Just like with sialic acid, the activated version of Pse is formed by the action of a CTP-dependent synthetase (PseF) which gives cytidine monophosphate pseudaminic acid (CMP-Pse).
1.9.3 Legionaminic Acid Biosynthesis

Legionaminic acid (5,7-diamino-3,5,7,9-tetraddeoxy-D-glycero-D-galacto-non-2-ulosonic acid or Leg) is a nine-carbon α-keto acid that is related in structure to sialic acid (NeuAc) (Figure 1.32). It was first discovered in the lipopolysaccharide (LPS) of *Legionella pneumophila*, which is a facultative intracellular parasite that is the cause of Legionnaires' disease. The disease, which can cause a fatal pneumonia, was first recognized in a large outbreak at an American Legion convention in Philadelphia in 1976. After the initial discovery of legionaminic acid, the
incorrect L-glycero-D-galacto configuration was assigned to it. It was only after legionaminic acid was prepared synthetically that the correct configuration of D-glycero-D-galacto, which is the same as sialic acid, was determined.\(^8^4\)

![Legionaminic Acid](image1.png)

\[
\text{Legionaminic Acid (Leg)}
\]

![X=O, N,N'-diacetyllegionaminic acid](image2.png)

\[
\text{X=O, N,N'-diacetyllegionaminic acid (Leg5Ac7Ac)}
\]

![X=NH, 5-N-acetimidoyl-7-N-acetyllegionaminic acid](image3.png)

\[
\text{X=NH, 5-N-acetimidoyl-7-N-acetyllegionaminic acid (Leg5Am7Ac)}
\]

![Sialic Acid](image4.png)

\[
\text{Sialic Acid}
\]

\[
\text{N-acetylneuraminic acid (NeuNAc)}
\]

Figure 1.32 Structure of legionaminic acid derivatives and sialic acid.

LPS is an immunogenic glycolipid that makes up the outer surface of the outer membrane of Gram-negative bacteria.\(^8^5\) LPS consists of the three covalently linked domains: lipid A, the core region and the \(O\)-chain polysaccharide (Figure 1.33). The \(O\)-chain polysaccharide in the LPS of \(L.\ pneumophila\) is composed of a repeating homopolymer of \(\alpha-(2\rightarrow4)\)-linked 5-N-acetimidoyl-7-N-acetyllegionaminic acid \(\text{(Leg5Am7Ac, Figure 1.32)}\).\(^8^5\), \(^8^7\) The \(O\)-chain polysaccharide is a repeating oligosaccharide polymer of 1-40 units and its primary role appears to be protective.
Recently, Leg5Am7Ac was discovered in Campylobacter coli where it was found as a posttranslational modification on the flagellar proteins. Campylobacter species require the flagella for motility and pathogenicity. It was determined that the flagellin proteins are heavily glycosylated with derivatives of legionaminic acid and pseudaminic acid. Without these posttranslational modifications, the bacteria were not able to assemble functional flagellae. Legionaminic acid derivatives have also been found in the O-antigens of Pseudomonas fluorescens, Vibrio alginolyticus and several other bacteria.

None of the enzymes involved in the biosynthesis of legionaminic acid have yet been identified. However, it is reasonable to assume that the biosynthesis of legionaminic acid would involve similar transformations as those in the biosynthesis of sialic and pseudaminic acids. At the onset of this thesis work we proposed that the biosynthesis would begin with the action of a hydrolyzing 2-epimerase that would convert UDP-\(N,N'\)-diacetylbacillosamine (UDP-Bac2,4diNAc) to 2,4-diacetamido-2,4,6-trideoxymannose (6-deoxyMandiNAc, Figure 1.34).
This would be distinct from a pathway in which epimerization and hydrolysis occurs in two separate steps. The subsequent reaction between 6-deoxyMandiNAc and PEP would be catalyzed by legionaminic acid synthase to produce \(N,N'\)-diacetyllegionaminic acid (Leg5Ac7Ac). Finally the \(\alpha\)-keto acid would be activated as CMP-Leg5Ac7Ac by the corresponding synthetase. Recently, a 30 kb lipopolysaccharide locus was identified in the \textit{L. pneumophila} genome which appears to contain the genes for the biosynthesis of \(N,N'\)-diacetyllegionaminic acid.\(^9\) Based on the homology of the encoded proteins to those involved in bacterial sialic acid biosynthesis (NeuA-C), tentative functions can be assigned to several of these genes.

The proposed starting material UDP-Bac2,4diNAc has been identified in \textit{Campylobacter jejuni} where it is used in the N-linked protein glycosylation system.\(^92\)-\(^94\) It was found that the N-
linked glycans play a major role in host adherence, invasion and colonization. The biosynthesis of UDP-Bac2,4diNAc in *C. jejuni* has been fully characterized, and the pathway involves three enzymes that require UDP-GlcNAc as the starting material (Figure 1.35). The first enzyme is UDP-GlcNAc 4,6-dehydratase (PglF) which catalyzes the formation of UDP-6-deoxy-4-keto-GlcNAc. The product is then subjected to the reactions catalyzed by PLP-dependent aminotransferase (PglE) and acetyltransferase (PglD) in order to generate the product UDP-Bac2,4diNAc. Genes encoding for homologs of the enzymes PglD-F also found in the genome of *L. pneumophila* and, as a result, it is reasonable to assume that UDP-Bac2,4diNAc is used as the precursor for legionaminic acid by this organism.95

![Figure 1.35 Biosynthesis of UDP-Bac2,4diNAc in Campylobacter jejuni.](image-url)
1.10 Branched-Chain Sugar Nucleotides

The branched--chain sugars can be divided into two groups based on their biogenesis. Group I consists of sugars bearing methyl or two-carbon side chains, while Group II consists of hydroxymethyl-branched and formyl-branched sugars. Group I sugars are generally formed by the coupling reaction between a nucleotide sugar and a one- or two-carbon unit from the appropriate donors. Both the hydroxymethyl-branched and formyl-branched sugars are generally formed by a rearrangement of the sugar chain so that one of the carbon atoms is reoriented to form the branch.

1.10.1 Methylation Reaction in the Biosynthesis of Mycarose.

*Streptomyces fradiae* has been reported to produce the macrolide antibiotic tylosin. L-Mycarose (Figure 1.36), as well as D-mycaminose and D-mycinose sugars, form an important part of the antibiotic structure. The genes responsible for the biosynthesis of the Group I sugar, L-mycarose, have recently been identified. One of the genes encodes the TylC3 enzyme which is believed to be responsible for the C-methyltransferase reaction. The enzyme uses S-adenosylmethionine (SAM) as the source of the methyl carbon and delivers it onto the C-3” position of dTDP-2,6-dideoxy-4-ketoaltrose. The reaction is believed to be initiated by a deprotonation at the C-3” position. The enolate intermediate then attacks the methyl group of SAM. The overall reaction occurs with the inversion of configuration at C-3” (Figure 1.36).
1.10.2 **Biosynthesis of UDP-D-Apiose.**

UDP-D-apiose is the activated form of D-apiose and it belongs to Group II of the branched-sugar nucleotides (Figure 1.37). Apiose was initially discovered in 1901 in the flavone glycoside apiin from parsley.\(^\text{100}\) The actual structure of apiose as 3-C-hydroxymethyl-D-erythrose was elucidated much later.\(^\text{101}\) D-Apiose is the only branched carbon sugar that is found in the plant cell wall as a monosaccharide.\(^\text{102}\) In higher plants, D-apiose is only present in the pectic polysaccharide rhamnogalacturonan-II (RG-II), which is a component of the primary cell wall. RG-II has a structural function, but nowadays is also considered to be a highly complex organelle.\(^\text{103}\) RG-II is believed to be the only polysaccharide associated with the primary cell wall which contains boron, which is an important micronutrient for higher plants.\(^\text{104}\) The two vicinal hydroxyl groups of the furanose ring of D-apiose are ideally positioned to form a cyclic
diester with borate (Figure 1.37). This linkage serves as an attachment point for two highly complex side chains to the homogalacturonan backbone. The D-apiose sugars are linked to the rest of the homogalacturonan backbone of RG-II via D-galacturonate residues.\textsuperscript{105} Even small structural changes in the RG-II structure, such as replacing the L-fucose residue with structurally similar L-galactose, reduced the formation of borate cross-links. This led to development of dwarf plants and abnormal leaf growth in \textit{Arabidopsis thaliana}, indicating that plant growth heavily depends on wall pectic polysaccharide formation.\textsuperscript{106}

![Chemical Structure](image)

Figure 1.37 D-Apiose involvement in formation of borate cross-link between two different homogalacturonan chains.

Biochemical studies on the synthesis of D-apiose initially focused on the enzyme from parsley (\textit{Petroselinum hortense}).\textsuperscript{107} It has been established that UDP-D-apiose is produced from UDP-D-glucuronic acid (UDP-GlcA) by an NAD\textsuperscript{+}-dependent UDP-D-apiose/UDP-D-xylene
Initially it was believed that there were two separate enzymes responsible for production of each UDP-sugar. However, despite repeated improvements in purification procedures of the enzymes from the cell cultures of parsley, the enzymatic activity for apiose and xylose synthesis could not be separated. This is a bifunctional enzyme that catalyzes the conversion of UDP-glucuronic acid into the mixture of UDP-D-apiose and UDP-D-xylose (Figure 1.38). UDP-D-apiose/UDP-D-xylose synthase belongs to the SDR superfamily of enzymes.

![Figure 1.38 The UDP-D-apiose/UDP-D-xylose synthase reaction.](image)

The detailed mechanism of this conversion still remains unclear, however previous investigations into the mechanism of UDP-D-apiose/UDP-D-xylose synthase from both duckweed (*Lemna minor*) and parsley (*Petroselinum hortense*) have established that UDP-D-apiose is formed from UDP-D-glucuronic acid (UDP-GlcA). This occurs in a single enzymatic reaction which leads to the decarboxylation of the substrate, followed by rearrangement of the carbon skeleton and ring contraction. The initial steps of the reaction are believed to be the same as those catalyzed by UDP-glucuronic acid decarboxylase (Section 1.8),
which begins with NAD+-dependent oxidation of C-4” followed by decarboxylation (Figure 1.39). It is believed that the UDP-4-ketoxylose intermediate can either be converted to UDP-D-xylose (UDP-Xyl) or UDP-D-apiose (UDP-Api). A reduction at C-4” by NADH would lead to the formation of UDP-Xyl, whereas rearrangement followed by a reduction would give UDP-Api.111, 112 The mechanism for the rearrangement that produces the D-apiose skeleton is still unknown, however, it could involve a retro-aldol cleavage between C-2” and C-3” to give an open chain enol(ate) intermediate, followed by aldol condensation between C-2” and C-4” to form a contracted ring structure (Figure 1.39).113 This transformation is very similar to the mechanism proposed in the biosynthesis of streptose.114, 115 Finally, UDP-Api is formed by a reduction of the aldehyde at C-3” by NADH.
Previous biochemical studies support this mechanism. When the shoots of young parsley plants were fed with UDP-[3,4-14C]-GlcA, it was discovered that the C-3'' of apiose must originate from either C-3'' or C-4'' of glucose. Later, the incubation of UDP-[3''-14C]-GlcA with UDP-d-apiose/UDP-d-xylose synthase from cell-suspension cultures of parsley demonstrated that the branch hydroxymethyl carbon (C-3'') of d-apiose comes from the C-3'' of the UDP-GlcA (Figure 1.40).
Figure 1.40 Fate of the labeled C-3\(^\text{\textsuperscript{14}}\) of UDP-[3\(^\text{\textsuperscript{14}}\)C]-glucuronic acid during the reaction catalyzed by UDP-D-apiose/UDP-D-xylose synthase.

Subsequent studies focused on determining the fate of the C-4\(^\text{\textsuperscript{3}}\) hydrogen. UDP-[U-\(\text{\textsuperscript{14}}\)C, 4-\(\text{\textsuperscript{3}}\)H]-GlcA was incubated with UDP-D-apiose/UDP-D-xylose synthase isolated from *Lemna minor* (Figure 1.41).\(^{117}\) The result of the experiment demonstrated that the tritium label from C-4\(^\text{\textsuperscript{3}}\) of the starting material was exclusively found at the branch carbon of the UDP-Api. This result can be explained by a hydride transfer process involving an NAD\(^+\) cofactor. The tritium label is initially transferred to the enzyme-bound NAD\(^+\) cofactor to form NAD\(^{3}\)H and UDP-4-ketoglucuronic acid. The enzyme-bound cofactor would then transfer the label to the branch carbon after the carbon rearrangement had occurred.

Figure 1.41 Fate of the labeled 4\(^\text{\textsuperscript{3}}\)H of UDP-[U-\(\text{\textsuperscript{14}}\)C, 4-\(\text{\textsuperscript{3}}\)H]-GlcA during the reaction catalyzed by UDP-D-apiose/UDP-D-xylose synthase.

In order to study the stereochemistry involved in the reduction of the C-3\(^\text{\textsuperscript{\text{\textsuperscript{3}}}\)}\) aldehyde of the UDP-3-aldehydo-apiose, the enzyme was incubated with UDP-[4\(^\text{\textsuperscript{3}}\)H]-GlcA. The resulting
mixture was first treated with acid and then sodium borohydride to form apiitol (Figure 1.42).\textsuperscript{112} The treatment of apiitol with sodium periodate gave glycolic acid. It was then oxidized by with glycolate oxidase, which is specific for the pro-R hydrogen.\textsuperscript{118} The final analysis of the product from the enzymatic transformation demonstrated that the tritium label from UDP-[4\textsuperscript{\textasciitilde}H]-GlcA is transferred only to the pro-R position on the C-3\textsuperscript{\textasciitilde} methylene of apiose. Similar experiments were performed with UDP-[5\textsuperscript{\textasciitilde}H]-GlcA as a substrate for UDP-D-apiose/UDP-D-xylose synthase to show that the enzymatic transformation occurs with an inversion of configuration at C-4\textsuperscript{\textasciitilde} of UDP-Api.\textsuperscript{112}

\begin{center}
\includegraphics[width=0.7\textwidth]{figure1.42.png}
\end{center}

Figure 1.42 Determination of stereochemistry in the conversion of UDP-[4\textsuperscript{\textasciitilde}H]-GlcA to UDP-[3\textsuperscript{\textasciitilde}H]-apiosyl.\textsuperscript{48}

Finally, to prove the existence of the UDP-4-keto-xylose intermediate, the enzymatic reaction between UDP-[U\textsuperscript{14C}]-GlcA and UDP-D-apiose/UDP-D-xylose synthase was quenched.
by the addition of NaB$_3$H$_4$. It was shown that after extensive purification over 90% of the tritium label was found at C-4" of UDP-Xyl.$^{112}$

While the mechanism shown in Figure 1.39 is consistent with all of these observations, an alternative mechanism has also been proposed (Figure 1.43).$^{117}$ Following the initial oxidation at C-4" and decarboxylation at C-5", a ring contraction occurs via deprotonation of the C-3" hydroxyl group and an alkyl migration to form a UDP-3-aldehydo-apiose intermediate. This is then reduced by the NADH cofactor to form UDP-Api. Similar rearrangements can be found in the non-mevalonate pathway for the biosynthesis of isoprene units$^{119}$ and branched chain amino acids.$^{120}$
Figure 1.43 Carbon migration mechanism for the UDP-D-apiose/UDP-D-xylose synthase reaction.
1.11 Project Goals

The aim of this thesis is to study two different biosynthetic pathways which are involved in the production of two unique sugar nucleotides. My studies elucidating the biosynthetic pathway of CMP-N,N'-diacetyllegionaminic acid are presented in Chapter 2. Mechanistic studies on the enzyme responsible for the formation of UDP-D-apiose are the focus of Chapter 3.

Chapter 2 will focus on identifying the biosynthetic pathway of CMP-N,N'-diacetyllegionaminic acid for the first time. It is believed that the biosynthesis of legionaminic acid will involve similar transformations as those in the biosynthesis of sialic acid (Section 1.9). The first goal of Chapter 2 is to identify and study the first enzyme of the biosynthetic pathway. If the first enzyme is indeed a hydrolyzing 2-epimerase, then the mechanism of the enzyme can be investigated using isotopic incorporation experiments. The second enzyme is proposed to be N,N'-diacetyllegionaminic acid synthase (NeuB), a potential PEP-condensing synthase homologous to the NeuAc synthase. The second goal of Chapter 2 is to identify the synthase and analyze its mechanism using isotopically labeled PEP. Finally, the last part of Chapter 2 will focus on identifying the synthetase that generates CMP-N,N'-diacetyllegionaminic acid. Since N,N'-diacetyllegionaminic acid is found on the flagellum of Campylobacter coli, the presence of the acid is essential for motility and pathogenicity of the organism, elucidation of the biosynthetic pathway could provide the basis for the development of new antibiotic drugs.

UDP-Ap biosynthesis is the main focus of Chapter 3. It is believed that UDP-Ap is biosynthesized from UDP-glucuronic acid (Section 1.10). Just recently, the candidate gene for UDP-D-apiose/UDP-D-xylose synthase was identified in Arabidopsis thaliana and was functionally expressed in Escherichia coli. The first goal of the project was to establish the
activity of the enzyme using NMR spectroscopy and to characterize the expected branched chain product. The catalytic competency of the potential UDP-4-ketoxylose intermediate to produce both UDP-Xyl and UDP-Api will be tested. The second goal of the chapter is to provide evidence for one of the two potential mechanisms that could be used by UDP-D-apiose/UDP-D-xylose synthase to perform carbon skeleton rearrangements. Several substrate analogs, such as UDP-2-deoxy-2-fluoro-D-glucuronic acid, UDP-3-deoxy-3-fluoro-D-glucuronic acid and UDP-[U-\(^{13}\)C]-3-deoxy-3-fluoro-D-glucuronic acid, were prepared and tested with the enzyme in order to help elucidate the mechanism.
Chapter Two

Biosynthesis of CMP-$N,N'$-Diacetyllegionaminic Acid
2.1 Introduction

The focus of this chapter will be on the identification of the three enzymes involved in the biosynthesis of CMP-N,N'-diacetyllegionaminic acid (CMP-Leg5Ac7Ac) and the elucidation of their chemical mechanisms. The inability to produce the proposed starting material, UDP-N,N'-diacetylbacillosamine (UDP-Bac2,4diNAc), hampered any previous investigation on the biosynthetic pathway. As mentioned in Chapter 1, the recent discovery of the enzymes involved in the biosynthesis of UDP-Bac2,4diNAc in C. jejuni allows for the efficient preparation of the desired starting material from UDP-GlcNAc. The preparation of UDP-Bac2,4diNAc will be discussed in the first part of this chapter. The chapter will then focus on the identification of, and mechanistic studies on, a hydrolyzing 2-epimerase similar to that involved in the biosynthesis of sialic acid. Experiments will be presented that show the enzyme converts UDP-Bac2,4diNAc into 2,4-diacetamido-2,4,6-trideoxymannose (6-deoxyMandiNAc) and will demonstrate that the elimination of UDP occurs through a C-O bond cleavage process. It will also be demonstrated that the epimerase reaction proceeds with a net retention of configuration at C-1.

The second enzyme to be identified in this chapter is N,N'-diacetyllegionaminic acid synthase (NeuB) that catalyzes the formation of the N,N'-diacetyllegionaminic acid (Leg5Ac7Ac) and phosphate from 6-deoxyMandiNAc and PEP. The isolation and characterization of the enzymatically produced acid confirms its stereoconfiguration as being identical to previously chemically synthesized derivatives. It will also be shown that the reaction catalyzed by N,N'-diacetyllegionaminic acid synthase proceeds via a C-O bond cleavage mechanism. The final section of this chapter will demonstrate that CMP-N,N'-diacetyllegionaminic acid is biosynthesized from Leg5Ac7Ac by the action of CMP-Leg5Ac7Ac synthetase.
2.2 Preparation of UDP-Bac2,4diNAc.

In order to test the hypothesis that Leg5Ac7Ac was biosynthesized from UDP-Bac2,4diNAc, it was first necessary to prepare this starting material. This was accomplished in a three-step chemoenzymatic synthesis (Figure 2.1) that utilizes two recombinant *C. jejuni* enzymes (PgL and PgL). The acetyltransferase (PgL) was also prepared but not used in the synthesis due to the efficient chemical acetylation of primary amines using acetic anhydride.

![Chemoenzymatic synthesis of UDP-Bac2,4diNAc](image)

Figure 2.1 Chemoenzymatic synthesis of UDP-Bac2,4diNAc
2.2.1 Expression and Purification of PglF, PglE and PglD enzymes

The plasmids used in the over-expression of PglF (Cj1120c), PglE (Cj1121c) and PglD (Cj1123c) were donated by Dr. Ian C. Schoenhofen. *E. coli* cells, transformed with appropriate plasmids, were grown in the presence of ampicillin. Over-expression was induced by the addition of IPTG (isopropyl β-D-1-thiogalactopyranoside) due to the presence of a lac operon upstream from *pglF, pglF* or *pglD* in the corresponding plasmids. Normally, in the absence of lactose or IPTG, a repressor binds to the lac operon, preventing the T7 RNA polymerase from binding to a T7 promoter region on the plasmid. IPTG binds to the repressor, causing the repressor to release from the lac operon, and allowing T7 RNA polymerase to bind and express the gene. Cells were lysed and the soluble fraction was loaded onto an affinity chromatography column containing immobilized nickel (Ni²⁺). The N-terminal hexahistidine tags on the enzymes bind with Ni²⁺ allowing the cell lysate to be washed from the column. The tagged enzyme was eluted with 500 mM imidazole and the appropriate fractions were dialyzed against a phosphate buffer before being flash-frozen with liquid nitrogen in the presence of 10% glycerol. The enzymes can be stored for up to nine months without significant loss of activity. The purity of the resulting proteins was determined to be >90% by SDS-PAGE analysis (Figure 2.2).
2.2.2 Chemoenzymatic synthesis of UDP-Bac2,4diNAc

The first step in the synthesis of UDP-Bac2,4diNAc involved treating UDP-GlcNAc with PgIF. The progress of the reaction was followed by monitoring the appearance of the UDP-6-deoxy-4-keto-GlcNAc using ESI-mass spectrometry (m/z 587 [M-H]⁻). Once the reaction was completed, the resulting mixture was used for the next transformation without further purification.

Since PgIE is a PLP-dependent aminotransferase enzyme, L-glutamate and PLP were added to the reaction mixture. The second reaction was initiated by adding PgIE and after several hours of incubation it was determined that >95% of the UDP-6-deoxy-4-keto-GlcNAc was converted to UDP-6-deoxy-4-amino-GlcNAc (m/z 589 [M-H]⁻). Both enzymes were removed by centrifugal ultrafiltration and the resulting sugar was purified by ion-exchange chromatography.
In the biosynthetic pathway, the transformation of UDP-Bac2NAc to UDP-Bac2,4diNAc is catalyzed by the acetyltrnsferase enzyme PglD in the presence of acetyl-CoA (Figure 1.35). This cofactor is used by a variety of acyltransferase enzymes to transfer an acetyl group to various acceptors including amines, alcohols, carbon nucleophiles and other thiol groups. During the preparation of UDP-Bac2,4diNAc the acetyltrnsferase enzyme PglD was available to me, however acetylation of the 4-amino group was performed chemically. Chemical acetylation is a high yielding process which does not require the use of expensive starting reagents like acetyl-CoA. The acetylation reaction was initiated by the addition of acetic anhydride to a solution of UDP-6-deoxy-4-amino-GlcNAc and methanol. Analysis by negative ESI-masspectrometry showed that after 24 hours all of the starting material was converted to UDP-Bac2,4diNAc (m/z 631 [M-H]). The UDP-Bac2,4diNAc was purified by ion exchange chromatography, and it was characterized by $^1$H NMR spectroscopy (Figure 2.3). The observed NMR chemical shifts are in agreement with previously published results.92
Figure 2.3 $^1$H NMR spectrum of UDP-Bac2,4diNAc (400 MHz, D$_2$O, 25 °C). * = triethylammonium
2.3 Identification and Mechanistic Studies on the Hydrolyzing UDP-Bac2,4diNac 2-Epimerase from *Legionella pneumophila*

Previous studies on LPS biosynthesis in *L. pneumophila* had identified a *neuC* homolog that encodes an enzyme required in the biosynthesis of legionaminic acid.\(^9\) The protein product of the gene shares 30% and 28% sequence identity with the NeuC responsible for sialic acid biosynthesis in *Neisseria meningitidis* and *E. coli* K1, respectively. Previous work on the bacterial biosynthesis of sialic acid in *Neisseria meningitidis* and *E. coli* K1 indicated that NeuC is a hydrolyzing UDP-\(N\)-acetylglucosamine 2-epimerase.\(^{50,121}\) This result suggests that the *L. pneumophila* NeuC could also be a hydrolyzing 2-epimerase that uses UDP-Bac2,4diNac as a potential substrate.

2.3.1 Expression and Purification of the Hydrolyzing UDP-Bac2,4diNac 2-epimerase

An expression plasmid for the *neuC* gene of *L. pneumophila* (lpg0753) was prepared by our collaborators in Dr. Martin Young’s laboratory at NRC. The recombinant plasmid was overexpressed by induction with 1.0 mM IPTG. The resulting cells were harvested by centrifugation and stored in pellet form at -80 °C. Following cell lysis, the crude protein extract was loaded onto an affinity chromatography column containing immobilized nickel (\(\text{Ni}^{2+}\)). The enzyme was eluted using a 500 mM imidazole buffer. Following buffer exchange, the purified protein was either used directly or stored at -80 °C in a buffer containing 10% glycerol. The purity of the protein was determined to be greater than 90% by SDS-PAGE analysis (Figure 2.4).
2.3.2 Testing the Activity of the Hydrolyzing UDP-Bac2,4diNAc 2-epimerase

To test the activity of the hydrolyzing UDP-Bac2,4diNAc 2-epimerase (NeuC), a sample of UDP-Bac2,4diNAc was incubated with the enzyme and the reaction was monitored by $^{31}$P NMR spectroscopy. Before the addition of NeuC, signals at -11.08 ppm and -12.89 ppm were observed that correspond to the diphosphate group of UDP-Bac2,4diNAc (Figure 2.5A). After the addition of enzyme and incubating for 15 minutes, a new set of signals appeared at -7.54 ppm and -10.32 ppm, corresponding to free UDP (Figure 2.5B). The sample was spiked with UDP to confirm the identity of the signals. This result indicated that a hydrolysis reaction was taking place. The progress of the reaction was also followed by both negative and positive ESI-mass spectrometry. It was found that all of the starting material UDP-Bac2,4diNAc ($m/z$ 631 [M-H]$^-$) was converted to UDP ($m/z$ 403 [M-H]$^-$) along with a compound whose mass corresponded to 2,4-diacetamido-2,4,6-trideoxymannose (6-deoxyMandiNAc, $m/z$ 269 [M+Na]$^+$).
Figure 2.5 $^{31}$P NMR spectra monitoring the reaction of UDP-Bac2,4diNAc with NeuC. 
A) before the addition of NeuC and B) after the addition of NeuC (121.5 MHz, D$_2$O, 25 °C).

After the UDP was removed by ion exchange chromatography, the sugar product was
determined to be a 1:1 mixture of anomers of 2,4-diacetamido-2,4,6-trideoxymannose (6-
deoxyMandiNAc) by $^1$H NMR spectroscopy (Appendix Figure A.1). The small $J_{H1,H2}$ values
observed in both of the anomers (0.91 Hz and 1.47 Hz) indicated that the acetamido group at C-2
was in an axial position, and therefore an inversion of stereocenter had occurred at C-2. The $^1$H
NMR data obtained were in complete agreement with those reported in the literature.$^{122}$ These
results confirm that the *L. pneumophila* NeuC homolog is a hydrolyzing UDP-Bac2,4diNac 2-epimerase.

After demonstrating the activity of hydrolyzing UDP-Bac2,4diNac 2-epimerase, the specificity of the enzyme was tested. The reaction catalyzed by the enzyme is very similar to that of the hydrolyzing UDP-GlcNAc 2-epimerase, so it is reasonable to assume that it is able to catalyze the conversion of UDP-GlcNAc to ManNac and UDP. Therefore, a sample of UDP-GlcNAc was incubated with hydrolyzing UDP-Bac2,4diNac 2-epimerase at 37 °C for two days and the progress of the reaction was followed by $^{31}$P NMR spectroscopy. No reaction was observed, indicating that hydrolyzing UDP-Bac2,4diNac 2-epimerase is specific for UDP-Bac2,4diNac and is involved only in the biosynthesis of the Leg5Ac7Ac.

### 2.3.3 Kinetic Characterization of the Hydrolyzing UDP-Bac2,4diNac 2-epimerase.

A continuously coupled UDP assay was used to determine the kinetic parameters of the UDP-Bac2,4diNac 2-epimerase. The assay involved the use of pyruvate kinase and lactate dehydrogenase enzymes (Figure 2.6). The first coupling enzyme, pyruvate kinase, takes phosphoenolpyruvate and UDP to form UTP and pyruvate. The second coupling enzyme reduces pyruvate to lactate while consuming one equivalent of NADH. This is monitored by UV spectroscopy as a decrease in absorbance at 340 nm ($\varepsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$). The reaction followed Michaelis–Menten kinetics and the kinetic parameters (Figure 2.7) obtained with UDP-Bac2,4diNac as a substrate were a $k_{\text{cat}}$ of $59.1 \pm 1.6 \text{ s}^{-1}$, a $K_M$ of $36.5 \pm 3.2 \mu\text{M}$ and $k_{\text{cat}}/K_M$ of $1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ (in 50 mM NaH$_2$PO$_4$ pH 7.5, 10 mM MgCl$_2$, 2 mM PEP and 0.2 mM NADH).
2.3.4 Stereochemical Analysis and Solvent Isotope Incorporation

The first mechanistic study conducted on the NeuC enzyme was devised to determine the first formed product of the enzymatic reaction. Since the produced sugar could readily undergo mutarotation, the enzymatic formation of one anomer of 6-deoxyMandiNAc would quickly lead...
to a mixture of anomers. In order to determine which anomer is actually produced by the enzyme, a sample of the substrate was incubated with a large amount of enzyme and examined immediately using $^1$H NMR spectroscopy. A dilute buffer was employed to minimize the possibility of buffer-catalyzed mutarotation. Moreover, the above experiment will allow simultaneous monitoring for any incorporation of solvent-derived deuterium into a non-exchangeable position of the product.

The enzymatic conversion of UDP-Bac2,4diNAc to UDP and 6-deoxyMandiNAc was monitored using $^1$H NMR spectroscopy in a buffer prepared using D$_2$O. The initial $^1$H NMR spectrum showed the anomic proton signal of UDP-Bac2,4diNAc at 5.45 ppm in the absence of enzyme (Figure 2.8, t= 0 min). This signal appears as a doublet of doublets due to coupling to both the β-phosphorus atom and H-2’. A relatively large amount of enzyme was added to the sample and a new spectrum was recorded immediately to minimize the possibility of mutarotation. The NMR spectral time course revealed the appearance of signals corresponding to 6-deoxyMandiNAc and UDP, with full conversion to products after 120 minutes. The chemical shifts for the H-1’’ protons of the 6-deoxyMandiNAc anomers appeared at 5.06 ppm (downfield signal) and 4.90 ppm (upfield signal). After only two minutes, the spectrum showed a non-equilibrium mixture of anomers in 10:1 ratio favoring the anomer with the downfield signal (Figure 2.8, t= 2 min). After twelve hours, the spectra showed full conversion to products and the non-enzymatic mutarotation had equilibrated the anomers at a 1:1 ratio (Figure 2.8, t= 12 h). It is clear that the anomer displaying the downfield chemical shift (5.06 ppm) is the true product of the enzymatic reaction and the anomer displaying the upfield chemical shift (4.90 ppm) is formed by mutarotation.
In order to determine the identity of the first formed anomer, the anomeric signals of 6-deoxyMandiNAc had to be assigned. For sugars with an axial proton at C-2, the H-1 to H-2
coupling constant ($J_{H1,H2}$) values can serve this purpose ($J_{H1,H2}$ for the $\alpha$-anomer is larger than the $\beta$-anomer). The same principle cannot be applied to sugars with an equatorial proton at C-2 because the coupling constants are similar for both anomers. Bock and Pedersen reported that the C-1 to H-1 coupling constant ($J_{C1,H1}$) values in a $^1$H-coupled $^{13}$C NMR spectrum are 10 Hz larger in the $\alpha$-anomer than in the $\beta$-anomer for a variety of d-monosaccharides.\textsuperscript{124} I therefore ran a two-dimensional heteronuclear NMR experiment (HMQC) on a sample of 6-deoxyMandiNAc and established that the first formed product has a C-1$^{13}$C signal at 92.93 ppm and the second formed product had C-1$^{13}$C signal at 92.86 ppm. Due to the small amounts of product available, the $J_{C1,H1}$ could not be directly obtained from a $^1$H-coupled $^{13}$C NMR experiment. Instead, an equilibrated sample of 6-deoxyMandiNAc was used in an HMQC NMR experiment, with the $^1$H-$^{13}$C coupling constant retained in the $^1$H dimension (Figure 2.9).
The results showed that the $^{13}$C NMR signal at 92.93 ppm exhibits a $J_{C_1,H1}$ value of 173 Hz and therefore is assigned as the $\alpha$-anomer, whereas the $^{13}$C NMR signal at 92.86 ppm exhibits a $J_{C_1,H1}$ value of 163 Hz and is assigned to the $\beta$-anomer. The observation that 6-deoxy-$\alpha$-MandiNAc is the first formed product indicates that the stereochemical course of the reaction proceeds with a net retention of configuration at C-1".

Figure 2.9 HMQC experiment with 6-deoxyMandiNAc (400 MHz, D$_2$O, 25 °C).
This experiment also shows that the reaction proceeds with solvent isotope incorporation at C-2 and that the product formed is actually [2-\(^2\)H]-6-deoxy-\(\alpha\)-MandiNAc. The absence of an H-2 signal in the \(^1\)H NMR spectrum of the product obtained from the reaction run in D\(_2\)O (not shown in Figure 2.8) proved that deuterium incorporation occurred at C-2”. This was also shown by the appearance of the anomeric proton signals as singlets due to a negligible coupling constant to the C-2 deuterium atom and loss of coupling to the \(\beta\)-phosphorus atom of UDP (Figure 2.8). Moreover, the isotope incorporation was further supported by positive ESI-mass spectrometric analysis of the produced [2-\(^2\)H]-6-deoxyMandiNAc product. The sample prepared in deuterated water was one mass unit larger than the product produced in H\(_2\)O. The incubation of UDP-Bac2,4diNAc or 6-deoxyMandiNAc in the same deuterated solvent without the presence of the NeuC enzyme did not lead to the incorporation of deuterium in either sugar. Earlier studies on other hydrolyzing 2-epimerases had also demonstrated the incorporation of solvent-derived deuterium label, indicating that deprotonation occurs at the C-2” position, followed by reprotonation with a solvent-derived deuterium atom.  

2.3.5 Test for C-O vs. P-O Bond Cleavage

The next mechanistic study was designed to test if the loss of UDP proceeds via a C-O or P-O bond cleavage process. A C-O bond cleavage mechanism would involve the elimination of UDP followed by the attack of water at C-1” (Figure 2.10 A). A P-O bond cleavage mechanism would involve the nucleophilic attack of water at the phosphorus atom of the \(\beta\)-phosphate and displacement of the sugar (Figure 2.10 B). A subsequent epimerization of the free sugar would give the product. If the incubation of the hydrolyzing 2-epimerase with UDP-Bac2,4diNAc in H\(_2\)^{18}O led to the incorporation of the \(^{18}\)O-label into 6-deoxyMandiNAc, the reaction would
involve the cleavage of the C-O bond. If the $^{18}$O-label is found in the UDP, the reaction occurs via a P-O bond cleavage mechanism (Figure 2.10).

![Diagram](image)

**Figure 2.10 C-O vs P-O bond cleavage experiment. A) Products from C-O bond cleavage mechanism. B) Products from P-O bond cleavage mechanism.**

The NeuC enzyme and UDP-Bac2,4diNAc were incubated in a phosphate buffer (pH 7.5) containing 50% $^{18}$H$_2$. The reaction was run until completion and the products were investigated using negative (UDP) and positive (6-deoxyMandiNAc) ESI-mass spectrometry. The mass spectral analysis of the isolated UDP demonstrated the absence of the $^{18}$O label. However, the resulting 6-deoxyMandiNAc showed a 1:1 ratio of $m/z$ 269 [M+Na]$^+$ and $m/z$ 271 [M+Na+2]$^+$, indicating that the $^{18}$O-label had been incorporated into the sugar (Figure 2.11). Since the $^{18}$O-label would be incorporated at the anomeric position of 6-deoxyMandiNAc, it is possible that this label could non-enzymatically exchange with solvent, giving a false positive result. In order to confirm that the label was incorporated during the enzymatic reaction and not through non-
enzymatic wash in, a sample of unlabeled 6-deoxyMandiNAc was incubated in buffer prepared with $^{18}$O-labeled water. The progress of the experiment was followed by positive ESI-mass spectrometry and, while a slow incorporation of $^{18}$O into the sugar was observed, the rate was orders of magnitude slower than observed in the enzymatic reaction.

![Diagram of enzymatic reaction](image)

Figure 2.11 ESI-mass spectra of C-O vs. P-O bond cleavage experiment with NeuC. A) Negative ESI-mass spectrum of UDP-Bac2,4diNAc before addition of NeuC. B) Positive ESIMS spectrum of product of the enzymatic reaction carried out in 100% H$_2$O. C) Positive ESIMS spectrum of products of the enzymatic reaction carried out in 50% H$_2$O and 50% H$_2^{18}$O.

Based on the fact that UDP-Bac2,4diNAc 2-epimerase catalyzes the conversion of UDP-Bac2,4diNAc to 6-deoxyMandiNAc through the cleavage of both the C-2'–H bond and the C-1'–O bond, it is reasonable to assume that the reaction proceeds via a 6-deoxy-2,4-
diacetamidoglucal intermediate (Figure 2.12). This is very similar to the UDP-GlcNAc 2-epimerase enzyme which is involved in the biosynthesis of sialic acid.\textsuperscript{49, 125} It has also been demonstrated that this hydrolyzing epimerase reaction proceeds with net retention of configuration at C-1. This means then that the enzyme mechanism is initiated by the anti-elimination of UDP, followed by the syn-hydration of 6-deoxy-2,4-diacetamidoglucal, forming 6-deoxy-\(\alpha\)-MandiNAc (Figure 2.12). The hydrolyzing UDP-GlcNAc 2-epimerase also catalyzes a net-retention of configuration at C-1 and the second step involves the syn-addition of water.

![Figure 2.12 Mechanism of the reaction catalyzed by the hydrolyzing UDP-Bac2,4diNAc 2-epimerase.](Image)

2.4 Identification and Mechanistic Studies on \(N,N'\)-Diacytylegionaminic Acid Synthase (NeuB)

Previous studies have identified a neuB homolog (lpg0752) in \textit{L. pneumophila} that is clustered with the neuC and neuA homologs and was believed to encode an \(N,N'\)-diacetyllegionaminic acid synthase.\textsuperscript{91} The protein is 61% identical in sequence to sialic acid synthase from \textit{N. meningitidis}. 

![Image]
2.4.1 Expression and Purification of N,N'-Diacetyllegionaminic Acid Synthase.

The *neuB* gene was cloned by our collaborators in the laboratory of Dr. Martin Young at NRC. Recombinant plasmids containing the *neuB* gene were transformed into *E. coli* and over-expressed to give both N-terminal and C-terminal His-tagged proteins. The resulting proteins were purified using a nickel column, and then subjected to a buffer exchange using centrifugal protein filters. In both cases, high levels of over-expression were achieved; however most of the resulting protein was insoluble and remained in the cell pellet during purification. As a result, only a very small amount of soluble protein (4 μg/L) could be prepared in this way. It is likely that during over-expression most of the protein misfolded creating inclusion bodies. Factors such as the temperature of induction, the pH of the cultivation medium or the changes in amino acid sequence due to addition of His-tag region all could have had an effect on formation of inclusion bodies. Varying the temperature of induction and pH of the medium reduced the amount of the protein produced but did not improve its solubility. As a result, our collaborators also expressed *neuB* as a maltose-binding protein fusion, MalE-NeuB. Maltose binding protein is a useful affinity tag that has been shown to increase the expression levels and solubility of the resulting tagged proteins in many instances. The recombinant plasmid was transformed into *E. coli* and over-expressed by induction with IPTG. The resulting protein was purified using an amylase column, eluting with buffer containing maltose. The purified protein was used directly in further studies or stored in a buffer containing 10% glycerol at -80 °C. The protein was determined to be greater than 90% pure by SDS-PAGE analysis (Figure 2.13).
Very high amounts of soluble protein (54 mg/L) could be prepared in this way. Unfortunately, it was later discovered that MalE-NeuB possesses only very low levels of synthase activity. Since MalE-NeuB was designed to contain a protease cleavage site between the two fusion partners, it was subjected to a thrombin treatment. Thrombin is a serine protease that converts soluble fibrinogen into insoluble strands of fibrin.\textsuperscript{128,129} The thrombin cleavage site (Leu-Val-Pro-Arg-Gly-Ser) had been incorporated into the MalE-NeuB sequence and NeuB produced after cleavage was designed to possess an N-terminal hexahistidine tag. The MalE-NeuB enzyme was incubated with thrombin at 25 °C, and the reaction was followed by SDS-PAGE (not shown). The results indicated that full cleavage was reached after twelve hours of incubation. The resulting mixture was loaded onto a nickel column and NeuB was eluted with 500 mM imidazole buffer. SDS-PAGE analysis showed that NeuB protein was still intact and did not degrade during the thrombin cleavage. However, no synthase activity was detected either before or after the His-tag purification column. As a result, for all remaining studies the MalE-NeuB version of the \(N,N'\)-diacetyllegionaminic acid synthase was used.
2.4.2 Test for \(N,N'-\text{Diacetyllegionaminic Acid Synthase (NeuB)}\) Activity

Once the MalE-NeuB enzyme was prepared, its activity was tested by incubation with 6-deoxyMandiNAc, PEP, and MgCl\(_2\) in Tris-DCI buffer prepared using D\(_2\)O. Due to the extremely low activity of the fusion protein, very high concentrations were used (typically 50 mg/mL) and it was difficult to force the reaction to completion. The reaction catalyzed by the \(N,N'\)-diacetyllegionaminic acid synthase is shown in Figure 2.14. The progress of the reaction containing a 2:1 mixture of PEP to 6-deoxyMandiNAc was monitored using \(^{31}\)P NMR spectroscopy (Figure 2.15). Before the addition of MalE-NeuB, the \(^{31}\)P NMR spectrum showed a single phosphorus signal at -3.19 ppm belonging to the phosphate group of PEP (Figure 2.15A). After the addition of MalE-NeuB and incubation for 18 hours, a new signal appeared at 0.03 ppm corresponding to inorganic phosphate (Figure 2.15B). Control reactions lacking enzyme or 6-deoxyMandiNAc did not produce any phosphate under similar conditions, indicating that the reaction is not due to a phosphatase impurity. Moreover, the incubation of MalE-NeuB with ManNAc and PEP did not produce any phosphate, indicating that NeuB does not possess sialic acid synthase activity.
Figure 2.14 Reaction catalyzed by \(N,N'\)-diacetyllegionaminic acid synthase (NeuB) and Fischer projections of 6-deoxyMandiNAc and \(N,N'\)-diacetyllegionaminic acid (open chain)
Figure 2.15 $^{31}$P NMR spectra monitoring the reaction of PEP and 6-deoxyMandiNAc with MalE-NeuB. A) Before the addition of MalE-NeuB and B) 18 h after the addition of MalE-NeuB (121.5 MHz, D$_2$O, 25 °C).

The MalE-NeuB enzymatic reaction was also followed by $^1$H NMR spectroscopy. The most efficient way to follow the enzymatic conversion was to focus on the 1.7 to 2.4 ppm region of the spectrum (Figure 2.16). Before the addition of MalE-NeuB, the $^1$H NMR spectrum shows peaks corresponding to the acetamido protons for the $\beta$-anomer and $\alpha$-anomer of 6-deoxyMandiNAc at 2.03 and 2.12 ppm, and 2.04 and 2.08 ppm, respectively (Figure 2.16A). After the addition of MalE-NeuB, the signals corresponding to 6-deoxyMandiNAc were found to be replaced by signals consistent with those expected for Leg5Ac7Ac (Figure 2.16B).$^{84}$ A single anomer of Leg5Ac7Ac is the main product formed in the reaction as indicated by the appearance of only one new pair of acetamido methyl signals at 1.99 and 2.01 ppm. This is assigned to the $\beta$-anomer which is expected to be more stable due to the equatorial carboxylate group. A similar preference is seen with sialic acid.$^{130}$ The C-3” methylene protons of Leg5Ac7Ac appear at 1.83
ppm and 2.23 ppm. The doublet of doublets at 2.23 ppm corresponds to the H-3 equatorial proton (H-3eq) and is characterized by a large $J_{3eq,3ax}$ value (13.1 Hz) and small $J_{3ax,4}$ value (4.7 Hz). The doublet of doublets that appears as a triplet at 1.83 ppm corresponds to the H-3 axial proton (H-3ax). This signal is split by a strong geminal coupling to the H-3eq proton ($J_{3ax,3eq} = 13.1$ Hz) and a strong coupling to the H-4 proton ($J_{3ax,4} = 12.2$ Hz). The large $J_{3ax,4}$ value indicates that H-3ax and H-4 have a trans-diaxial relationship and that a newly formed hydroxyl group occupies the equatorial position. This demonstrates that MalE-NeuB catalyzes the addition of PEP to the si-face of the open chain aldehyde of 6-deoxyMandiNAc to form the (S)-configuration at C-4. This is the same stereospecificity as displayed by sialic acid synthase and pseudaminic acid synthase.49,131
Figure 2.16 Partial ¹H NMR spectra monitoring the incubation of PEP and 6-deoxyManDiNAc with MalE-NeuB A) Before the addition of MalE-NeuB and B) after the addition of MalE-NeuB (400 MHz, D₂O, 25 °C).

Since MalE-NeuB exhibits very low activity of ∼4.4 x 10⁻³ μmol min⁻¹mg⁻¹, no attempt was made to kinetically evaluate the reaction. Similar activity studies were performed on both N- and C-terminal His-tagged versions of the NeuB enzyme. It was found that the C-terminal His-tagged enzyme was about 10-fold more active and the N-terminal His-tagged was about 50-fold more active than MalE-NeuB. These low levels of activity likely do not represent the true activity of NeuB in vivo. It is possible that modification of either terminus results in an enzyme with dramatically reduced activity. Perhaps more likely is the possibility that NeuB requires the presence of another unidentified protein or cofactor for full activity. Similar to sialic acid
synthase, Leg5Ac7Ac synthase requires the presence of a divalent cation in order for the reaction to occur. Performing the enzymatic incubation in the presence of 1.0 mM EDTA instead of MgCl₂ did not lead to formation of inorganic phosphate or Leg5Ac7Ac. This will be discussed further in Section 2.4.4.

2.4.3 Isolation and Characterization of N,N'-Diacetyllegionaminic Acid

Previous synthetic studies assigned the D-glycero-D-galacto configuration to Leg5Ac7Ac. In order to confirm that MalE-NeuB produces Leg5Ac7Ac and not an epimer, it was necessary to fully characterize the product. Enzymatically produced Leg5Ac7Ac was isolated by first removing MalE-NeuB using centrifugal protein filters and then passing the solution through a single anion-exchange chromatographic column. The isolated product was then characterized using ESI-mass spectrometry and ¹H NMR spectroscopy. The ESI-mass spectral analysis of Leg5Ac7Ac showed one major peak at m/z 333 [M-H]. The ¹H NMR spectrum of Leg5Ac7Ac was in complete agreement with that of the synthetic material previously reported in the literature (Figure 2.17).
Figure 2.17 $^1$H NMR spectrum of Leg5Ac7Ac (400 MHz, D$_2$O, 25 °C). * = triethylammonium
2.4.4 Potential Synthase Mechanisms and Test for C-O vs P-O Bond Cleavage

Enzymes that catalyze the condensation between a sugar carbonyl and PEP have been proposed to follow one of two mechanisms: a C-O bond cleavage mechanism or a P-O bond cleavage mechanism.\textsuperscript{132} The first step of the C-O bond cleavage mechanism is the attack of C-3 of the PEP at the carbonyl carbon of the open chain aldehyde of 6-deoxyMandiNAc (Figure 2.18A). This attack would be facilitated by the presence of the divalent cation activating the aldehyde carbon. An oxocarbonium ion intermediate would be formed as a result of the PEP attack. A tetrahedral intermediate is then formed by the attack of a water molecule onto the oxocarbonium ion. Upon the collapse of the tetrahedral intermediate and release of phosphate, the open chain of the $N,N'$-diacetyllegionaminic acid is generated, which subsequently cyclizes to the pyranose form in solution. The C-O bond cleavage mechanism is employed by both sialic acid and pseudaminic acid synthase. It has also been observed with 2-keto-3-deoxy-D-manno-2-octulosonic acid 8-phosphate synthase and 2-keto-3-deoxy-D-arabino-2-octulosonic acid 7-phosphate synthase.\textsuperscript{133,134} The P-O bond cleavage mechanism is initiated by the attack of water on to the phosphate group of the PEP (Figure 2.18B). This attack would result in the formation of the enolate anion of pyruvate and the release of phosphate. The formed enolate anion would then attack the carbonyl carbon on the open chain form of 6-deoxyMandiNAc to form the open chain of the Leg5Ac7Ac, which again would cyclize to the pyranose form in solution. The P-O bond cleavage mechanism is employed by pyruvate kinase\textsuperscript{135} and PEP carboxykinase\textsuperscript{136} where catalysis is thought to proceed with a nucleophilic attack at the phosphate group.

In order to distinguish between a C-O and a P-O bond cleavage mechanism, an incubation of [2-$^{18}$O]-PEP with PEP-condensing synthases is generally used (see $^{18}$O-labeled atoms in Figure 2.18).\textsuperscript{132,134} If MalE-NeuB operates via a C-O bond cleavage mechanism, then the
reaction with [2-\(^{18}\text{O}\)]PEP would lead to the formation of \(^{18}\text{O}\)-labeled phosphate. On the other hand, if the enzyme employs the P-O bond cleavage mechanism, then \(^{18}\text{O}\)-labeled Leg5Ac7Ac should be formed. Similar labeling studies on both sialic acid synthase and pseudaminic acid synthase had shown that a C-O bond cleavage mechanism is employed by both of these enzymes.\(^{49,131}\)

\[\text{Figure 2.18 Proposed C-O vs. P-O bond cleavage mechanisms for MalE-NeuB. A) C-O bond cleavage mechanism and B) P-O bond cleavage mechanism.}\]

\[\text{[2-\(^{18}\text{O}\)]-PEP was generously donated by a former group member Dr. Wayne Chou.}^{131}\]

The purity of the compound was determined by an ESI- mass spectral analysis. It was found that the [2-\(^{18}\text{O}\)]-PEP had 54% \(^{18}\text{O}\) label incorporation. The substitution of \(^{16}\text{O}\) for \(^{18}\text{O}\) in a singly bonded position to phosphorus results in a small upfield shift in the \(^{31}\text{P}\) NMR signal of the labeled phosphorus atom,\(^{137}\) therefore, \(^{31}\text{P}\) NMR spectroscopy was used to monitor the enzymatic incubation. MalE-NeuB was incubated with 20 mM [2-\(^{18}\text{O}\)]-PEP and 12 mM 6-deoxy MandiNAc in a buffer containing 1.0 mM MgCl\(_2\). The initial \(^{31}\text{P}\) NMR spectrum of the 54%
labeled [2-¹⁸O]-PEP showed two phosphorus signals at -3.01 ppm and -3.03 ppm corresponding to unlabeled PEP and ¹⁸O-labeled PEP, respectively (Figure 2.19A). After several hours of incubation, two new phosphorus signals appeared. The signal at 0.16 ppm is attributed to the unlabeled phosphate and the signal at 0.14 ppm to the ¹⁸O-labeled phosphate (Figure 2.19B). The ratio of ¹⁶O to ¹⁸O was similar to that in the [2-¹⁸O]-PEP, indicating that the ¹⁸O label was fully retained in the phosphate produced. The results of this experiment show that the reaction proceeds through the C-O bond cleavage mechanism.

Figure 2.19 ³¹P NMR spectra monitoring the incubation of MalE-NeuB with [2-¹⁸O]PEP and 6-deoxyMandiNAc A) Before the addition of MalE-NeuB and B) after the addition of MalE-NeuB (121.5 MHz, D₂O, 25 °C).
2.5 CMP-\(\text{N,N'}\)-Diacetylegionaminic Acid Synthetase

Previous studies had identified a \textit{neuA} homolog (lpg0751) in \textit{L. pneumophila} that is believed to encode a CMP-Leg5Ac7Ac synthetase\(^91\). The protein is 54\% and 50\% identical in sequence to the CMP-NeuAc synthetases from \textit{E. coli} and \textit{N. meningitidis}, respectively.

2.5.1 Identification, Expression and Purification of CMP-\(\text{N,N'}\)-Diacetylegionaminic Acid Synthetase

The \textit{neuA} gene was cloned by our collaborators in the laboratory of Dr. Martin Young at NRC. This recombinant plasmid was over-expressed in \textit{E. coli} by induction with IPTG. The resulting cells were harvested, and following lysis, the crude cell lysate was loaded onto an affinity chromatography column containing immobilized nickel (Ni\(^{2+}\)) and the enzyme of interest was eluted using 500 mM imidazole buffer. Following buffer exchange, the protein was either used directly or stored at -80 °C in a buffer containing 10\% glycerol. The protein was determined to be greater than 90\% pure by SDS-PAGE analysis (Figure 2.20).
Figure 2.20 SDS-PAGE gel showing NeuA purification. Lane 1: molecular weight standards; lane 2: crude cell lysate before over-expression; lane 3: crude cell lysate after incubating with ITPG for 5 h; lane 4: purified NeuA. Molecular weight standards BSA (29 kDa) and carbonic anhydrase (66 kDa).

2.5.2 Activity of CMP-\(N,N'\)-Diacetylegionaminic Acid Synthetase

Once the NeuA enzyme was prepared, its activity was tested by incubation with Leg5Ac7Ac and cytidine-5'-triphosphate (CTP) in Tris-HCl buffer containing MgCl₂. The reaction catalyzed by CMP-Leg5Ac7Ac synthetase is shown in Figure 2.21. Due to the low activity of Leg5Ac7Ac synthase (MalE-NeuB) only small amounts of pure Leg5Ac7Ac were prepared. As a result, the progress of the NeuA reaction was only followed by negative ESI-mass spectrometry. Before the addition of NeuA, the ESI-mass spectrum showed a single peak at \(m/z\) 333[M-H]⁻ belonging to the starting material Leg5Ac7Ac. After the addition of NeuA and incubation for two hours, a new signal appeared at \(m/z\) 660 [M+Na-2H]⁻ belonging to CMP-Leg5Ac7Ac. This observation was used to confirm that the \(L. pneumophila\) NeuA possessed CMP-Leg5Ac7Ac synthetase activity. No further characterizations of the product or studies were performed on the NeuA enzyme.
2.6 Conclusions

Three candidate genes coding for the proteins involved in the biosynthesis of CMP-Leg5Ac7Ac were cloned by our collaborators in Dr. Martin Young’s laboratory. I have characterized the corresponding enzymes for the first time and thus have elucidated the biosynthetic pathway of CMP-Leg5Ac7Ac. Several experiments were carried out with the enzymes in order to elucidate their mechanisms of action.

I first demonstrated that the NeuC homolog was a hydrolyzing 2-epimerase that catalyzed the conversion of hydrolyzing UDP-Bac2,4diNAc into 6-deoxyMandiNAc and UDP. The reaction is analogous to the UDP-GlcNAc 2-epimerase of sialic acid biosynthesis.49 The incubation of UDP-Bac2,4diNAc in a deuterated buffer with the hydrolyzing 2-epimerase led to the formation of 6-deoxy-α-[2-2H]MandiNAc. The fact that the α-anomer was formed first showed that the reaction proceeds with the retention of configuration at C-1. The observation that the reaction proceeds with solvent-derived isotope incorporation at C-2" indicates that C-2" is deprotonated and reprotonated with a solvent-derived isotope during the course of the reaction. The reaction carried out in H₂¹⁸O buffer demonstrated that the loss of UDP occurs through a C-O bond cleavage mechanism. Together, these results indicate that the first step of the reaction consists of an anti-elimination of UDP. The second step of the reaction involves a syn-hydration.
of the 6-deoxy-2,4-diacetamidoglucal intermediate. This is in agreement with the mechanism proposed for the UDP-GlcNAc 2-epimerase enzyme that is involved in the biosynthesis of sialic acid.

The second enzyme in the biosynthesis of CMP-Leg5Ac7Ac is a PEP-dependent Leg5Ac7Ac synthase. The incubation of 6-deoxyMandiNAc with PEP in the presence of MalE-NeuB led to the formation of Leg5Ac7Ac and phosphate. The resulting Leg5Ac7Ac had a D-glycero-D-galacto configuration which is in agreement with previously published results. The use of [2-18O]-PEP in the MalE-NeuB reaction resulted in the formation of 18O-labeled phosphate, indicating that MalE-NeuB utilizes a C-O bond cleavage mechanism. The mechanism is believed to involve an initial attack of the C-3 of PEP onto the open chain aldehyde of 6-deoxyMandiNAc forming an oxocarbenium ion intermediate, followed by the attack of water and formation of a tetrahedral intermediate (Figure 2.18A). The tetrahedral intermediate then collapses, releasing phosphate and forming the open chain form of Leg5Ac7Ac which cyclizes in the solution to form predominantly the β-anomer.

Finally, the last enzyme in the biosynthetic pathway, NeuA, was identified as CMP-\(N,N'\)-diacetyllegionaminic acid synthetase and converted Leg5Ac7Ac into CMP-Leg5Ac7Ac in the presence of CTP and MgCl\(_2\). The biosynthesis of CMP-Leg5Ac7Ac from UDP-Bac2,4diNAc in Legionella pneumophila is shown in Figure 2.22.
Figure 2.22 Biosynthesis of CMP-Leg5Ac7Ac in *Legionella pneumophila.*
2.7 Future Directions

The three enzymes involved in the biosynthesis of CMP-Leg5Ac7Ac have been identified and investigated. Further studies could focus on a more detailed investigation of their chemical mechanisms and on understanding the reasons for the low activity exhibited by Leg5Ac7Ac.

In order to strengthen the proposal of an elimination-hydration mechanism for the hydrolyzing 2-epimerase reaction, it would be useful to incubate synthetically prepared 6-deoxy-2,4-diacetamidoglucal intermediate with the enzyme. This would test whether the intermediate is catalytically competent to serve as a substrate for the second step of the reaction (hydration). A similar experiment was performed on the hydrolyzing UDP-GlcNAc 2-epimerase. Obtaining an X-ray crystal structure of the UDP-Bac2,4diNAc hydrolyzing 2-epimerase would help to identify key residues in the active site of the epimerase and give insight into their role during the enzymatic reaction. In case of the bacterial UDP-GlcNAc hydrolyzing 2-epimerase several mutants were prepared and investigated. The Asp131Asn mutant catalyzed the formation and a release of 2-acetamidoglucal intermediate into solution. It is believed that Asp131 is involved in the hydration of the glycal intermediate. Similar studies can be performed on the UDP-Bac2,4diNAc hydrolyzing 2-epimerase with the corresponding Asp136 residue to determine if it plays the same role in the catalytic cycle.

The Leg5Ac7Ac synthase requires full kinetic characterization. It would be necessary to prepare a more catalytically active version of the enzyme. If the low activity of NeuB is due to N- or C-terminal modification, then a native enzyme must be obtained. One way to prepare a native version of the enzyme is to make a version of the N- or C-terminal His-tag which can be cleaved during the purification process. Another approach would involve the preparation of the
NeuB enzyme without any modification to the amino acid sequence and the purification of the enzyme using several chromatographic steps. However, the low activity of NeuB could be due to a requirement for another protein or cofactor, and to test for this the protein could be co-incubated with proteins encoded by nearby genes in the operon. The activity could also be tested in the presence of crude cell lysate of *Legionella pneumophila*. The stereochemical course of the addition of PEP to 6-deoxyMandiNAc during the reaction needs to be addressed. The separate incubations of NeuB with *Z-[3-\(^{2}\text{H}\)]-PEP and *E-[3-\(^{2}\text{H}\)]-PEP in the presence of 6-deoxyMandiNAc would help to determine the stereochemistry of the addition.\(^{138,139}\) I have demonstrated that the reaction occurs through a C-O bond cleavage mechanism, however, the stereoconfiguration of the tetrahedral intermediate is still unknown (Figure 2.23). Several tetrahedral intermediate analogs can be prepared to help determine the stereoconfiguration at C-2. These compounds would also be expected to serve as potent inhibitors of the enzyme. They could be prepared as a mixture of epimers with different stereoconfiguration at C-2. Upon incubation of the mixture with synthase, it would be expected that the enzyme would bind most tightly to the epimer that bears the same stereoconfiguration as the normal intermediate. The stereoconfiguration of the tetrahedral intermediate could be elucidated by analyzing the resulting enzyme-inhibitor complexes with X-ray crystallography. Finally, the UDP-Bac2,4diNAc that was used in these experiments was prepared with *Campylobacter jejuni* enzymes. It is necessary to locate the genes encoding these three enzymes in the *L. pneumophila* genome in order to demonstrate that the bacterium is capable of generating the substrate for the pathway described in this thesis.
Figure 2.23 Potential tetrahedral intermediate analogs for NeuB.
2.8 Experimental

2.8.1 Materials and General Methods

All chemicals and enzymes, unless otherwise noted, were purchased from Sigma-Aldrich and were used without further refinement. $^{18}$O-enriched H$_2$O (95%) was purchased from Cambridge Isotope Laboratories. All the buffer solutions were prepared using distilled water. $^1$H NMR spectra were obtained on Bruker AV300/AV400 NMR spectrometers. $^{13}$C NMR spectra were obtained on Bruker AV300/AV400 NMR spectrometers at a field strength of 75 MHz or 100 MHz, respectively. Proton-decoupled $^{31}$P NMR spectra were recorded either on a spectrometer at a field strength of 121.5 MHz or 162 MHz, respectively. ESI-mass spectrometry was performed on a Bruker Esquire LC mass spectrometer. Chelex$^\text{®}$ 100 resin (200-400 mesh, Na$^+$ form), AG$^\text{®}$ 1-X8 resin (100-200 mesh, formate form), and Bio-Gel$^\text{®}$ P-2 resin were purchased from Bio-Rad Laboratories. DE-52 resin (DEAE Cellulose) was purchased from Whatman. Amicon Ultra-4 centrifugal protein filters (4 mL, 10 000 MWCO) were purchased from Millipore. All proteins were handled at 4 °C unless otherwise stated. Protein concentrations were determined according to the method of Bradford$^{140}$ with bovine serum albumin as the standard. Protein purity was determined using SDS-PAGE gel electrophoresis and visualized using coomassie blue stain according to the method of Laemmli.$^{141}$ Protein molecular masses were determined using BSA (66 kDa) and carbonic anhydrase (29 kDa) as mass standards.
2.8.2 Cloning of *Legionella pneumophila* neuA, neuB and neuC.

The cloning of neuA, neuB and neuC was performed in the laboratory of Dr. Martin Young at NRC. The neuA, lpg0751, neuB, lpg0752, neuC, lpg0753, genes were obtained from the genomic DNA of *L. pneumophila* subsp. *pneumophila* ATCC 33152™ by PCR amplification with Phusion DNA polymerase (New England Biolabs Inc.) according to the manufacturer’s instructions. The primers used incorporated either *NdeI* or *SalI* cloning sites (underlined) and a 6X His tag in either the 5P or 3P PCR primer. The primer pairs used were: (neuA5PHis) 5’-CTAGCTAGCTAGCATATGCAATCCATCACCATACCATCAGAATATTGGCAGTAATCC-CGGC-3’ (forward) and 5’-CTAGCTAGCTAGGTGACCTATTATATACTAGACCTCTTGGTTTAATTCC-3’ (reverse), (neuA3PHis) 5’-CTAGCTAGCTAGCATATGAGAATATTGGCAGTAATCCCGGC-3’ (forward) and 5’-CTAGCTAGCTAGGTGACTTATTAGTGATGTGGTGATGGTGATGGTGATGGTGATGGTGATGTGCAGTAATCC-3’ (reverse), (neuB5PHis) 5’-CTAGCTAGCTAGCATATGAGAATATTGGCAGTAATCCCGGC-3’ (forward) and 5’-CTAGCTAGCTAGGTGACCTATTATATGTTCCCATAACAAAGTGTACACCGC-3’ (reverse), (neuB3PHis) 5’-CTAGCTAGCTAGCATATGAGAATATTGGCAGTAATCCCGGC-3’ (forward) and 5’-CTAGCTAGCTAGGTGACCTATTATATGTTCCCATAACAAAGTGTACACCGC-3’ (reverse), (neuC5PHis) 5’-CTAGCTAGCTAGCATATGAGAATATTGGCAGTAATCCCGGC-3’ (forward) and 5’-CTAGCTAGCTAGGTGACCTATTATATGTTCCCATAACAAAGTGTACACCGC-3’ (reverse), (neuC3PHis) 5’-CTAGCTAGCTAGCATATGAGAATATTGGCAGTAATCCCGGC-3’ (forward) and 5’-CTAGCTAGCTAGGTGACCTATTATATGTTCCCATAACAAAGTGTACACCGC-3’ (reverse). The PCR products were gel purified and cloning sites
were generated by double digestion with NdeI and SalI restriction enzymes according to the
manufacturer's suggested protocol (New England Biolabs Inc.). The genes were cloned into
NdeI/SalI digested plasmid pCWori and the constructs maintained in Escherichia coli AD202.

2.8.3 Over-expression and Purification of L. pneumophila NeuA, NeuB and NeuC.

The recombinant neuC plasmid was transformed into E.coli BL21 (DE3) competent cells
which were incubated in 10 mL Luria-Bertani (LB) medium containing 50 mg/L ampicillin at 37
°C/225 rpm for 10 h. The overnight culture was then poured into 500 mL of LB medium
containing 50 mg/L ampicillin and shaken at 37 °C/225 rpm until an OD600 of 0.6 – 1.0 had been
reached. Cultures were induced with 1 mM isopropyl β-D-galactopyranoside (IPTG). After
incubation for 5 h at 37 °C, the cells were harvested by centrifugation and stored as a pellet at
—80°C. The pellets were resuspended in 10 mL of a phosphate buffer (20 mM, pH 8.0) containing
2 mM dithiothreitol (DTT), 1 mg/L of aprotinin, and 1 mg/L pepstatin A at 4 °C. The cells were
subsequently lysed by passage through a French Pressure cell at 20 000 psi. The lysate was
centrifuged at 6 000×g for 1 h, passed through 0.45 μm and 0.22 μm filters, and loaded onto a
column containing 10 mL of Chelating Sepharose Fast Flow resin (Pharmacia Biotech), which
was previously charged with 100 mM NiSO4 and washed with sodium phosphate buffer (20 mM,
ph 8.0, containing 0.5 M NaCl and 5 mM of imidazole). The purification process was monitored
by a Flow Thru UV monitor Spectrometer at 280 nm. Nonspecifically bound proteins were
washed away by applying buffers containing first 5 mM and then 125 mM imidazole. Finally,
bound enzyme was eluted using a 500 mM imidazole buffer. The fractions containing the
desired enzyme were combined and concentrated using Amicon Ultra Centricons (Millipore)
before flash freezing with liquid N2 in the presence of 10% glycerol. Similar procedures were
used for NeuB and NeuA. For NeuA, the fractions eluted with 500 mM imidazole were dialyzed
against 20mM Tris-HCl buffer (pH 7.5) containing 200 mM NaCl before being concentrated and flash-frozen with liquid N₂ in the presence of 10% glycerol.

2.8.4 Sub-cloning of L. pneumophila neuB for Fusion Protein, Overexpression and Purification.

The cloning of neuB was performed in the laboratory of Dr. Martin Young. The NeuB5PHis clone, verified by sequence analysis, was double digested with NdeI and SalI restriction enzymes. The liberated insert was gel purified and sub-cloned into pCWOri⁺ containing the E. coli malE gene with a downstream thrombin cleavage recognition sequence that was cloned as a BamH1 to NdeI fragment. The constructs were maintained in E. coli AD202, and cells bearing positive clones were identified by colony PCR using specific malE and neuB primers and by restriction mapping. The resulting malE-neuB plasmid was transformed into E. coli BL21 (DE3) competent cells, which were then grown at 37 °C in a 2YT media supplemented with 50 mg/L ampicillin. Over-expression of the fusion protein was induced by the addition of IPTG to a final concentration of 1 mM at an A₆₀₀ of 0.5, and growth continued for a further 6 h. Cells were harvested by centrifugation at 10 000 x g for 15 min, resuspended in a 20 mM Tris-HCl buffer pH 7.5 containing 1 mg/L of aprotinin, and 1 mg/L pepstatin A, and lysed by passage through a French Pressure cell at 20 000 psi. The cell lysate was clarified by centrifugation at 27 000 x g for 30 min and the cell debris was discarded. The total membrane and soluble protein fractions were obtained from clarified cell extracts by centrifugation at 10 000 x g for 60 min. Following the adjustment to 200 mM NaCl and 1 mM EDTA the soluble protein fraction was passed through a 20 mL amylose resin (New England Biolabs Inc.) column previously equilibrated with 200 mM NaCl, 20 mM Tris-HCl pH 7.5 and 1 mM EDTA. The column was washed with 3 column volumes of an equilibration buffer and bound protein was
eluted with an equilibration buffer containing 10 mM maltose. Fractions containing protein of interest, as judged by SDS-PAGE, were pooled and dialyzed against 200 mM NaCl, 20 mM Tris-HCl buffer pH 7.5. Glycerol (10%) was added to the solution and aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

2.8.5 Over-expression and Purification of His-Tagged PgLF and PgLE

The plasmids pNRC40.1 and pNRC41.3 were used in the over-expression of pgLF (Cj1120c) and pgLE (Cj1121c) to give His-tagged proteins, as described previously. In each case the appropriate plasmid was transformed into E. coli BL21 (DE3) competent cells which were incubated in 10 mL Lucia-Bertani (LB) medium containing 50 mg/L ampicillin at 37 °C/225 rpm for 10 h. The culture was then added to 500 mL of LB medium containing 50 mg/L ampicillin and shaken at 37 °C/225 rpm until an OD600 of 0.6 had been reached. The culture was allowed to continue to grow for 5 h after 70 mg/L of IPTG was added. Cells were harvested by centrifugation and the resulting pellet was stored at -80 °C. The pellet was then resuspended in 10 mL of a phosphate buffer (10 mM, pH 7.0) containing 2 mM dithiothreitol (DTT), 1 mg/L of aprotinin, and 1 mg/L pepstatin A. The cells were lysed by passage through a French Pressure cell at 20,000 psi. The lysate was centrifuged at 10,000 × g for 1 h and passed both through a 0.45 μm and a 0.22 μm filters. A column containing 10 mL of Chelating Sepharose Fast Flow resin (Pharmacia Biotech) was charged with 20 mL of 100 mM NiSO4, washed with 20 mL of distilled H2O and 30 mL of a sodium phosphate buffer (10 mM, pH 7.0, containing 0.5 M NaCl and 5 mM of imidazole). The lysate was loaded onto the column and eluted with the same buffer containing increasing amounts of imidazole in a step-wise fashion (5 mM, 125 mM and 500 mM). Fractions containing protein of interest, as judged by SDS-PAGE, were pooled and dialyzed against a 20 mM phosphate buffer (pH 7.0) before being flash frozen with liquid N2 in
the presence of 10% glycerol.

2.8.6 Chemo-enzymatic Synthesis of UDP-Bac2,4diNAc

UDP-Bac2,4diNAc was prepared under conditions slightly modified from those described previously. A purified sample of PglF (5 mg) was added to 50 mL of a phosphate buffer (10 mM, pH 7.0) containing 500 mg of UDP-N-acetylglucosamine disodium salt and 200 μM NAD⁺. The solution was incubated for 6 h at 37 °C and the reaction progress was followed by a negative ESI-mass spectrometry. Almost all of the starting material was converted to a UDP-4-keto-sugar. To synthesize the UDP-4-amino sugar, the following components were added to the mixture: 5 mg of purified PglE, L-glutamate (15 mM final concentration) and pyridoxal 5’-phosphate (PLP) (100 μM final concentration). The solution was incubated for 4 h at 37 °C and the reaction progress was monitored by a negative ESI-mass spectroscopy. It was determined that > 95% of the UDP-4-keto sugar (m/z 587, [M-H]⁻) was converted to UDP-Bac2Ac (m/z 589, [M-H]⁻) during this time. Enzymes were then removed by centrifugal ultrafiltration and the resultant filtrate was loaded onto a 220 mL column of DEAE cellulose (DE-52, Whatman Inc.) and eluted with a linear gradient of a 0 to 0.5 M triethylammonium bicarbonate buffer. The A₂₅₄ of the eluant was monitored, and UV-active fractions were analyzed by a negative ESI-mass spectroscopy. Those containing UDP-Bac2Ac were lyophilized to dryness. The lyophilized sugar (295 mg) was stirred with 1.2 mL acetic anhydride in 25 mL methanol at room temperature for 24 h. Negative ESI-MS showed that the starting material was completely converted to UDP-Bac2,4diNAc (m/z 631, [M-H]⁻) during this time. After removal of the solvent under reduced pressure, the product was loaded onto a DE-52 anion exchange column and subjected to linear gradient elution as described above. After lyophilization, the product was dissolved in 10 mL H₂O and lyophilized again. This procedure
was repeated twice more to yield 151 mg (28%) of the UDP-Bac2,4diNAc as its triethylammonium salt. $^1$H and $^{31}$P NMR spectra matched those in the literature. $^{92}$ $^1$H NMR (D$_2$O): $\delta$ 7.97 (d, 1H, $J_{5,6} = 8.1$ Hz, H-6); 5.97 (d, 2H, $J_{5,6} = 8.1$, $J_{1',2'} = 4.4$ Hz, H-5, H-1'); 5.48 (dd, 1H, $J_{1',2'} = 3.3$ Hz, $J_{1',1'} = 7.0$ Hz, H-1''); 4.38. (m, 2H, H-2', H-3'); 4.28. (m, 1H, H-4'); 4.18-4.23 (m, 2H, H-5'); 4.05 (m, 1H, H-5''); 4.02 (m, 1H, H-2''), 3.79 (t, 1H, $J_{2',3'} = J_{3',4'} = 10.2$ Hz, H-3''); 3.69 (t, 1H, $J_{4',5'} = 10.2$ Hz, H-4''); 2.06 (s, 3H, H$_3$CONH); 2.03 (s, 3H, H$_3$CONH); 1.20 (d, 1H, $J_{5',6'} = 6.1$ Hz, H-6''). $^{31}$P NMR (D$_2$O): $\delta$ −10.84 (d, $J_{Pa, Pp} = 20.1$ Hz), −12.67 (d, $J_{Pa, Pp} = 20.3$ Hz). ESIMS: m/z 631 [M-H]

2.8.7 Characterization of Hydrolyzing 2-Epimerase Activity

2.8.7.1. NeuC Homolog Activity Assay

A glycerol stock solution of the NeuC homolog (70 µg) was subjected to a buffer-exchange with a 25 mM phosphate buffer (pH 7.5, 150 µL final volume) using centrifugal ultrafiltration. This was added to a solution of UDP-Bac2,4diNAc (3.0 mg) dissolved in 850 µL H$_2$O (1.0 mL final volume) and $^{31}$P NMR and positive ion ESI-mass spectra were acquired at timed intervals. Once the reaction was complete, the enzyme was removed by centrifugal ultrafiltration and the resultant filtrate was passed through a column (15 mL) of AG-1X8 resin (100-200 mesh, formate form) and eluted with water to remove the UDP. The flow through was lyophilized to dryness and redissolved in D$_2$O for spectral analysis. Material prepared in this fashion was also analyzed by a 2-D heteronuclear NMR experiment (HMQC) both with, and without, the $^1$H-$^{13}$C coupling constant retained in the $^1$H dimension. This was done in order to establish the identity of the H-1'' signals for each of the anomers (vide infra). $^1$H NMR (D$_2$O): $\alpha$-anomer, $\delta$ 5.10 (d, 1H, $J_{1',2'} = 1.2$ Hz, H-1); 4.30 (dd, 1H, $J_{2',3'} = 4.5$ Hz, H-2''); 4.06 (dd, 1H,
$J_{3',4''} = 10.5$ Hz, H-3’); 3.97 (m, 1H, H-5’); 3.80 (t, 1H, $J_{4',5''} = 10.4$ Hz, H-4’); 2.08 (s, 3H, H$_3$CONH); 2.04 (s, 3H, H$_3$CONH); 1.19 (d, 3H, $J_{5',6''} = 6.4$ Hz, H-6’); $\beta$-anomer, 8 4.96 (d, 1H, $J_{1',2''} = 1.3$ Hz, H-1’); 4.47 (dd, 1H, $J_{2',3''} = 4.3$ Hz, H-2’); 4.47 (dd, 1H, $J_{2',3''} = 4.3$ Hz, H-2’); 3.67 (t, 1H, $J_{4',5''} = 10.2$ Hz, H-4’); 3.67 (m, 1H, H-5’); 2.12 (s, 3H, H$_3$CONH); 2.03 (s, 3H, H$_3$CONH); 1.23 (d, 3H, $J_{5',6''} = 6.3$ Hz, H-6’). ESIMS: m/z 269 [M+Na$^+$].

2.8.7.2. NeuC Kinetic Studies

Enzyme kinetics were measured using a continuous coupled assay for UDP formation.$^{123}$ Each cuvette contained a 50 mM NaH$_2$PO$_4$ buffer (pH 7.5), 10 mM MgCl$_2$, 2 mM PEP, 0.2 mM NADH, 20 units of lactate dehydrogenase, 18 units of pyruvate kinase, and UDP-Bac2,4diNAc (varying from 25 to 1000 μM) at a total volume of 800 μL. The concentrations of stock UDP-sugar solutions were determined by measuring $A_{260}$ ($\varepsilon = 9890$ M$^{-1}$ cm$^{-1}$). Enzymatic reactions were initiated by the addition of a 20 μL of 0.05 mg/mL enzyme solution (final concentration 2.0 nM). Rates were measured by monitoring the decrease in $A_{340}$ at 37 °C. Kinetic parameters were determined by fitting initial velocities to the Michaelis-Menten equation using GraFit 4.0. No detectable background release of UDP was observed in the absence of the added NeuC homolog.

2.8.7.3. Stereochemistry and Solvent Deuterium Isotope Incorporation Studies

A glycerol stock solution of NeuC (70 μg) was subjected to a buffer-exchange with a 25 mM phosphate /D$_2$O buffer (pD 7.4, 200 μL final volume) using centrifugal ultrafiltration. This was added to a solution of UDP- Bac2,4diNAc (3 mg) dissolved in 800 μL D$_2$O (1 mL final volume) and $^1$H NMR spectra were acquired at timed intervals during the incubation at 25 °C.
The isotope incorporation also was monitored by positive ESI-mass spectrometry as a function of time.

2.8.7.4. **Metal Dependency of NeuC**

Two aliquots of a solution containing UDP-Bac2,4diNAc (3 mg per aliquot) in a 25 mM phosphate buffer (pH 7.5) were prepared. One aliquot received MgCl₂ and the other EDTA tetrasodium salt, each at a 10 mM final conc. at a total volume of 990 μL. The NeuC homolog (70 μg) was added to each sample and the mixtures were incubated for 2 h at room temperature. The progress of the reactions was monitored by a ³¹P NMR spectroscopy with integration of the diphosphate signals.

2.8.8 **Test for C-O vs. P-O Bond Cleavage Mechanism**

A solution of a 25 mM phosphate buffer (pH 7.5, 1.60 mL) was prepared from 50% H₂¹⁶O and 50% H₂¹⁸O (95% isotopic enrichment) and divided into two aliquots. To one aliquot UDP-Bac2,4diNAc (2.0 mg) was added, while 6-deoxyMandiNAc (1.0 mg) was added to the other. The NeuC homolog (70 μg) was added to each sample and the mixtures were incubated at room temperature. The isotope incorporation was monitored by both positive (sugar detection) and negative (UDP detection) ESI-MS as a function of time. The extent of incorporation into 2,4,6-trideoxy-diacetamidomannose was calculated from the ratio of peaks at m/z 269 (¹⁶O, [M+Na]+) and m/z 271 (¹⁸O, [M+2+Na]+).
2.8.9 Characterization of \(N,N^\prime\)-Diacetyllegionaminic Acid Synthase

A solution containing 6-deoxyMandiNAc (12 mM) and PEP (20 mM) in a deuterated Tris/DCl buffer (700 \(\mu\)L, 10 mM, prepared in using D_2O, pD 7.4) was placed in an NMR tube. Initial \(^1\text{H}\) and proton-decoupled \(^3\text{P}\) NMR spectra were taken. The solution was removed from the tube and mixed with 5 mg of the NeuB homolog and 1 mM MgCl\(_2\) in the same deuterated buffer (1 mL total volume). After incubation of the reaction mixture for 20 h at 25 °C, Chelex-100 resin (~20 mg, previously rinsed with D_2O) was added, and the solution was incubated for an additional hour at room temperature. The resulting mixture was analyzed directly by \(^1\text{H}\) and \(^3\text{P}\) NMR spectroscopy. In order to determine the activity of the NeuB homolog under initial velocity conditions, a solution containing 2,4-diacetamido-2,4,6-trideoxymannose (6.4 mM), PEP (20 mM), MgCl\(_2\) (1 mM), and the NeuB homolog (6.5 mg) in Tris/DCl buffer prepared using D_2O (10 mM, pD 7.4, 1.0 mL total volume) was placed into NMR tube and immediately monitored by \(^1\text{H}\) NMR spectroscopy. Spectra were taken every 5 minutes for a period of one hour while incubating at 25 °C. The conversion rate was calculated by comparing the integrals of the signals due to the acetamido methyl protons of both 2,4-diacetamido-2,4,6-trideoxymannose anomers (2.03, 2.04, 2.08 and 2.12 ppm) to those of the product \(N,N^\prime\)-diacetyllegionaminic acid (1.99 and 2.01 ppm). The rate of the reaction was determined by using the data that was accumulated during the first 15% of the reaction.

2.8.10 Isolation and Characterization of the \(N,N^\prime\)-Diacetyllegionaminic Acid

The NeuB homolog was removed from the enzymatic activity test reactions by centrifugal ultrafiltration and the resulting filtrate was loaded onto a 15 mL column of Dowex-
AG1 X8 resin (formate form, 100-200 mesh, Bio-Rad) pre-equilibrated with water. A stepwise gradient of 0-1.0 M formic acid in water with 0.1 M increments (50 mL per increment) was used to elute the product. *N,N''*-Diacetyllegionaminic acid eluted from the column in the 0.2 M and 0.4 M fractions which were concentrated *in vacuo* and then lyophilized. *N,N''*-Diacetyllegionaminic acid was characterized using a $^1$H NMR and negative ESI-MS mass spectrometry and the spectroscopic data was identical to that reported for the synthetically produced material.$^{84,122}$ $^1$H NMR (D$_2$O) $\beta$-anomer $\delta$ 1.17 (d, 3H, $J_{5r,5''}$ = 6.2 Hz, H-9$''$), 1.84 (dd, 1H, $J_{3ax''},4''$ = 11.9 Hz, $J_{3ax''},3eq''$ = 13.1 Hz, H-3ax$''$), 1.99 (s, 3H, H$_3$CONH); 2.01 (s, 3H, H$_3$CONH), 2.26 (dd, 1H, $J_{3eq'3ax'}$ = 13.1, $J_{3eq'3ax'}$ = 4.7 Hz, H-3eq$'$), 3.64 (dd, 1H, $J_{4''5''}$ = 10.3 Hz, $J_{5''6''}$ = 10.5 Hz, H-5$''$), 3.75-3.8 (m, 2H, H-7$''$, H-8$''$), 3.93 (ddd, 1H, $J_{3eq'4'}$ = 4.8, $J_{3ax'4'}$ = 11.9, $J_{4,5}$ = 10.3 Hz, H-4$''$), 4.23 (dd, 1H, $J_{5'',6''}$ = 10.5, $J_{6'',7''}$ = 1.9 Hz, H-6$''$). ESIMS: $m/z$ 333 [M-H]$^-$.

2.8.11 Test for C-O vs. P-O Bond Cleavage Mechanism

$[2^{-18}$O$]$-PEP disodium salt was prepared as described previously$^{142,143}$ and found to have a 54% incorporation of the isotopic label as indicated by $^3$P NMR and mass spectral analysis. A solution containing 6-deoxyMandiNAc (12 mM) and $[2^{-18}$O$]$-PEP (20 mM) in Tris/DCI buffer (700 µL prepared using D$_2$O, 10 mM, pH 7.4) was placed in an NMR tube. Chelex-100 resin (~20 mg) was added and an initial proton-decoupled $^3$P NMR spectrum was obtained using a previously reported procedure.$^{131}$ The Chelex resin was removed by decanting the solution, and the solution was mixed with 5 mg of the NeuB homolog and 1.0 mM MgCl$_2$ in 250 µL of the same Tris/DCI buffer. The reaction was incubated for 20 h at 25 °C and a solution of EDTA tetrasodium salt in D$_2$O (10 mM final concentration) was added to the reaction. Another proton-
decoupled $^{31}$P NMR spectrum was acquired with the same parameters: $^{31}$P NMR 0.16 (s, P-$^{16}$O), 0.14 (s, P-$^{18}$O), -3.01 (s, P-$^{16}$O, PEP), -3.03 (s, P-$^{18}$O, PEP).

2.8.12 Characterization of CMP-$N,N'$-Diacytylegionaminic Acid Synthetase Activity

A glycerol stock solution of the NeuA homolog (containing 1.25 mg of enzyme) was subjected to a buffer-exchange with a 20 mM Tris-HCl buffer, pH 7.5. The enzyme was added to a solution containing 500 μM $N,N'$-diacytylegionaminic acid, 1.5 mM cytidine 5'-triphosphate disodium salt and 1.0 mM MgCl$_2$ in the same buffer (final volume 1.0 mL). The reaction was incubated at 25 °C and the progress was monitored using a negative ion ESI-mass spectrometry. After 2 h, all of the starting material with $m/z$ 333 ([M-H]$^-$) was converted to a CMP-$N,N'$-diacytylegionaminic acid product with $m/z$ 660 ([M+Na-2H]$^-$).
Chapter Three

Mechanistic Studies on UDP-\(d\)-Apiose Synthase
3.1 Introduction

The focus of this chapter will be on the elucidation of the chemical mechanism of UDP-D-apiose/UDP-D-xylose synthase (AXS1), the enzyme that converts UDP-D-glucuronic acid (UDP-GlcA) into UDP-D-apiose (UDP-Api) in plants (Figure 1.38). As discussed in Section 1.10.2, D-apiose is a critical component of the rhamnogalacturonan-II (RG-II) that, in plant cell walls crosslinks the homogalacturonan backbone via the formation of a borate ester. The lack of UDP-D-apiose/UDP-D-xylose synthase resulted in RG-II deficiency and cell death. The work in this chapter was prompted by a report that a candidate gene AXS1 from Arabidopsis thaliana, encoding a potential UDP-D-apiose/UDP-D-xylose synthase, had been identified by comparing the sequence to that of known UDP-D-glucurionate decarboxylases. It was also reported that expression of this gene in E. coli led to production of an active UDP-D-apiose/UDP-D-xylose synthase.

The first part of this chapter describes the over-expression of the AXS1 in E. coli in order to generate a hexa-histidine tagged version of UDP-D-apiose/UDP-D-xylose synthase. The purified enzyme was then incubated with UDP-GlcA and the progress of the enzymatic incubation was followed by NMR spectroscopy and ESI-mass spectrometry to fully characterize the reaction products. Further studies were designed to mechanistically investigate the UDP-D-apiose/UDP-D-xylose synthase. A potential intermediate in the UDP-D-apiose/UDP-D-xylose synthase-catalyzed reaction, namely UDP-4-ketoxylose, was prepared enzymatically. The reaction between AXS1 and UDP-4-ketoxylose supported the notion that it is the intermediate of the enzymatic transformation, however the rate of the reaction was extremely low. Finally, UDP-2-deoxy-2-fluoro-D-glucuronic acid, UDP-3-deoxy-3-fluoro-D-glucuronic acid and UDP-[U-¹³C]-3-deoxy-3-fluoro-D-glucuronic acid were prepared. Incubations with these three substrate
analogs provided further evidence for the retro-aldol mechanism of the carbon skeleton rearrangement carried out by UDP-D-apiose/UDP-D-xylose synthase.

3.2 Cloning, Expression and Purification of UDP-D-Apiose/UDP-D-xylose Synthase

A search of the *Arabidopsis thaliana* genome sequence revealed two genes, *AXS1* (At2g27860) and *AXS2* (Atg1g08200) which could potentially encode UDP-D-apiose/UDP-D-xylose synthase. Both of the gene products are SDR enzymes that contain the conserved Gly-X₃-Gly-X-Gly sequence responsible for the binding of NAD⁺ and the conserved catalytic triad responsible for promoting hydride transfer. It was demonstrated that *AXS1* and *AXS2* share 96% sequence identity and that *AXS2* encodes a functional UDP-D-apiose/UDP-D-xylose synthase that is identical to the enzyme encoded by *AXS1*. As a result, we chose to focus our studies on the enzyme that is encoded by *AXS1*.

3.2.1 Preparation of cDNA from *Arabidopsis thaliana*

Compared to prokaryotic and other eukaryotic genomes, plant genomes tend to be larger and more complex. Generally, they are 10 to 100 times larger than those of other eukaryotes. The major part of the plant genome is composed of introns, which are the non-coding intervening sequences of DNA. The minor part is composed of exons which are amino acid coding regions that are eventually translated to give proteins. Both introns and exons can be transcribed into mRNA by action of an RNA polymerase. The transcribed mRNA undergoes splicing which removes the introns. This is carried out by small nuclear ribonuclear proteins and produces mature mRNA. Since mature mRNA serves as a template for protein synthesis, it is isolated from plant tissues to obtain complementary DNA (cDNA). To obtain mRNA from *A. thaliana*, several
leaves of the plant were flash frozen, ground and then treated with a lysis buffer. The plant leaves were generously donated by Dr. Reinhard Jetter, and Ortwin Guhling provided invaluable assistance in performing mRNA and cDNA preparation. The RNeasy® kit (Qiagen) was used to isolate the mRNA from the leaves. To obtain cDNA, the isolated mRNA was subjected to the action of a reverse transcriptase enzyme, which is a DNA polymerase that will use either an RNA or DNA strand as a template (Figure 3.1). A short oligonucleotide complementary to the poly-A tail at the 3’ end of the mRNA is first hybridized to the RNA to act as a primer for the reverse transcriptase, which then copies RNA into a complementary DNA chain. The mRNA is degraded by the action of Ribonuclease H (RNaseH) that specifically degrades the RNA in the RNA:DNA hybrid, without affecting DNA. The single strand of cDNA prepared in this way is devoid of both upstream and downstream regulatory sequences, and of introns. Therefore, cDNA from eukaryotes can be translated into functional proteins when expressed in bacteria.
Figure 3.1 Preparation of cDNA from plant mRNA

3.2.2 Expression and Purification of AXS1

The next step in the cloning process was to amplify the AXS1 sequence from the cDNA mixture by PCR and insert it into an expression vector. The preparation of an expression plasmid bearing an AXS1 that encodes a His-tagged enzyme was completed using the Novagen Xa ligation-independent cloning (LIC) kit. The AXS1 sequence was PCR-amplified from Arabidopsis thaliana cDNA to create a double-stranded piece of DNA containin AXS1 sequence sandwiched between two specifically designed oligomers 15 and 17 nucleotides long. The PCR product was then incubated with T4 DNA polymerase and dGTP. Under these conditions, T4 polymerase hydrolyzes nucleotides from the 3'-ends of the PCR product until the first guanosine
is reached. As a result of the polymerase action, the DNA strands are no longer perfectly matched. They contain 12- or 15-base single stranded overhangs that are generally referred to as “sticky” ends. These “sticky” ends are complementary to the overhangs in a commercially available pET-30 vector, such that the target insert anneals with the vector to give a doubly nicked plasmid. The annealed mixture was transformed into competent *E. coli* cells where ligation occurs to form a plasmid. At this point the plasmid was purified, and the *AXS1* insert was sequenced at the Nucleic Acid Protein Service (NAPS) unit to verify that no errors were introduced during the PCR and DNA polymerase treatments.

The plasmid was then transformed into *E. coli* competent cells for protein over-expression. The cells were grown at 37 °C in LB medium containing kanamycin. The prepared plasmid contains a gene encoding for kanamycin resistance, so only those cells bearing the plasmid would grow in the above medium. Following induction with IPTG, the cells were harvested, lysed, and the soluble fraction was loaded onto an affinity chromatography column containing immobilized nickel (Ni$^{2+}$). The hexahistidine tag on AXS1 binds with Ni$^{2+}$, and the remainder of the cell lysate can be washed from the column. AXS1 was eluted with 500 mM imidazole, and the AXS1 containing fractions were dialyzed into a buffer at pH 8.0. As previously reported, AXS1 was found to be unstable and a significant loss of activity was observed within 48 h, or upon storage at 4 °C, -20 °C or -80 °C. The presence of up to 2 mM of 1,4-dithiothreitol did not have an effect on the stability or the activity of AXS1. As a result, the enzyme was prepared freshly prior to use in all experiments. AXS1 prepared in this fashion typically yielded ~250 mg per liter of culture. SDS-PAGE analysis (Figure 3.2) of the purified protein revealed a single band at ~46 kDa, consistent with the predicted molecular weight and indicating a purity of greater than 90%.
Proper folding of the recombinant protein was assessed by the observation of a tightly bound NADH cofactor using UV/Vis spectroscopy. It has been reported that recombinant AXS1 from *A. thaliana* was isolated with NAD⁺ tightly bound, and that it showed detectable levels of activity without the addition of exogenous cofactor. The UV spectrum of the recombinant AXS1 prepared in this study showed a strong absorbance band at 355 nm indicative of bound NADH with a chromophore that is slightly red-shifted from that of free NADH (Figure 3.3). Sodium borohydride was added to convert any bound NAD⁺ into NADH. However, only a 10% increase in absorbance was observed, indicating that the bound cofactor was present as a 9:1 ratio of NADH:NAD⁺.
Figure 3.3 Partial UV spectra of 147 \( \mu \text{M} \) AXS1 in 20 mM Tris-HCl pH 8.0 before and after the addition of sodium borohydride (to give 1.2 mM sodium borohydride).

For further studies with the enzyme, the catalytically relevant oxidized form of the tightly bound cofactor was desired. Several attempts were made to generate a sample of the enzyme containing an NAD\(^+\) cofactor. As reported previously, an extended incubation of AXS1 with NAD\(^+\) did not lead to the exchange of the cofactor bound to the active site.\(^{111}\) Most of the SDR family enzymes that employ transient oxidation mechanisms bind tightly to the cofactor and only release it upon protein denaturation.

In order to facilitate exchange of the bound cofactor, the enzyme was incubated in a 100 mM Tris-HCl buffer containing 8 M urea for 30 minutes. It was anticipated that this would result in partial denaturation of the protein and in cofactor release. The resulting mixture was then dialyzed against a buffer solution containing 20 mM Tris-HCl and 1 mM NAD\(^+\) over a period of
Following buffer exchange and concentration, the UV spectrum of reconstituted AXS1 was measured and showed no significant absorbance at wavelengths above 310 nm. Upon the addition of sodium borohydride, absorbance bands at 354 nm and 425 nm were observed (spectra not shown), which were consistent with the generation of tightly bound NADH. The 354 nm band corresponds to the biologically relevant 1,4-reduction product, whereas the 425 nm band corresponds to the 1,2- and 1,6-reduction products, as previously reported.

Unfortunately, the reconstituted AXS1 showed no enzymatic activity when incubated with UDP-D-glucuronic acid despite repeated attempts. The reason for the lack of activity in an apparently folded protein is not clear, but the loss of the activity observed upon storage of AXS1 is consistent with the loss that occurred upon extended incubation/dialysis in these studies. As a result, the initially isolated AXS1 having 90% of the cofactor in the wrong oxidation state was used in all experiments described below.

### 3.3 Testing the Activity of AXS1

#### 3.3.1 Monitoring the Activity of AXS1 Using $^1$H and $^{31}$P NMR Spectroscopy

In the previous studies, only limited amounts of AXS1 were available, and the product UDP-Api had never been directly characterized (only degradation products were characterized to support apiose formation). With the recombinant AXS1 enzyme in hand, I could test its activity by incubating it with UDP-GlcA, in phosphate buffer at pH 8.0 prepared in D$_2$O and containing 1 mM NAD$^+$. Even though the enzyme already has a tightly bound cofactor, it was reported that addition of NAD$^+$ to the incubation mixture helps to improve enzyme stability. The progress of the reaction was monitored using $^1$H NMR spectroscopy (Figure 3.4). The initial $^1$H NMR spectrum showed the anomeric proton signal of UDP-GlcA at 5.49 ppm in the absence of
enzyme (Figure 3.4, t= 0 min). This signal appears as a doublet of doublets due to coupling to the β-phosphorus atom and H-2". Initial attempts to observe activity by adding 2-5 mg of enzyme showed only partial reaction even after extended incubation times, indicating that the recombinant enzyme had very low levels of activity. Low activity levels of recombinant AXS1 were also reported in the literature. The low activity of AXS1 could be due to N-terminal modification introduced during the preparation of the recombinant His-tagged protein. It is also conceivable that the true substrate is not UDP-GlcA, but a structurally related sugar nucleotide. However this is unlikely, since no other branched chain sugar nucleotides are known. A more likely scenario is that AXS1 prepared using E. coli lacks the post-translational modifications necessary for optimal activity. Of course, the fact that only 10% of the protein contains cofactor in the correct oxidation state exacerbates this problem. A further difficulty is that UDP-Api is known to decompose upon extended incubation, and therefore it was not possible to detect it in this manner. In order to circumvent these problems, a relatively large amount of enzyme (150 mg) was added to the sample and a new spectrum was recorded immediately. The NMR spectral time course revealed the appearance of signals corresponding to UDP-Api and UDP-Xyl, with full conversion to products after 15 minutes (Figure 3.4). A new signal at 5.58 ppm is tentatively assigned to the UDP-Api anomeric proton. The signal is a doublet of doublets, but shows up as a triplet due to the similar coupling constants between $J_{H-1"-H-2"}$ and $J_{H-1"-P}$. The other doublet of doublets at 5.42 ppm is assigned to UDP-Xyl by its comparison to a sample of commercially available material. This demonstrates that the same enzyme catalyzes the formation of both UDP-Api and UDP-Xyl and allowed for the first spectral characterization of UDP-Api. Control reactions lacking either the enzyme or UDP-GlcA did not produce any products under similar conditions. The ratio of the two products was found to be approximately 1:1. It has been reported
that the ratio of UDP-API to UDP-Xyl varies with the ions present in the solution, with phosphate buffer having the highest ratio of UDP-API to UDP-Xyl. As a result, all of the experiments were performed in phosphate buffer with the exception of those monitored by \(^{31}\text{P}\) NMR spectroscopy. After ten hours, the spectrum showed full decomposition of UDP-API, while no degradation was observed for UDP-Xyl (Figure 3.4, t= 10 h). Since the recombinant AXS1 shows extremely low activity (\(\sim 6.7 \times 10^{-3} \text{ mol min}^{-1}\text{mg}^{-1}\)) no kinetic studies were performed.
Figure 3.4 $^1$H NMR spectra monitoring the reaction of UDP-GlcA with the AXS1 at different time intervals (400 MHz, D$_2$O, 25 °C).

Previous studies had shown that UDP-Api is not stable at a weakly alkaline pH and has a half life of about 97 minutes.$^{110,150}$ The sugar decomposes non-enzymatically to $\alpha$-D-apio-D-furanosyl 1,2-cyclic phosphate and uridine 5'-monophosphate (UMP, Figure 3.5). The formation of 1,2 cyclic phosphates at alkaline pH has also been observed with UDP-glucose$^{151}$ and TDP-6-deoxy-L-talose$^{152}$; however, the rates of cyclization were considerably slower. The newly formed
five-membered phosphorus-containing ring prefers to adopt a planar conformation. In both furanose and pyranose rings the C-O bonds at C-1" and C-2" must be in an eclipsing conformation to accommodate the formation of the phosphorus-containing ring. In certain envelope conformations of furanosides the eclipsing of these bonds is already established and, as a result, furanosides can readily undergo such a cyclization. In the case of a pyranoses, the ring must distort, partially or completely, from the more stable "chair" conformation to the less stable "half-chair" conformation, and the cyclization is slow.\(^\text{110}\) In the case of UDP-ApI the energy barrier for formation of a 1,2 cyclic phosphate compound is low, therefore rapid decomposition of the sugar is observed.

![UDP-D-apiose](image)

**Figure 3.5 Non-enzymatic decomposition of the UDP-ApI to α-D-apio-D-furanosyl 1,2-cyclic phosphate and UMP.**

The reaction catalyzed by AXS1 was also monitored by \(^{31}\text{P} \) NMR spectroscopy (Figure 3.6). Before the addition of AXS1, the \(^{31}\text{P} \) NMR spectrum showed two doublets belonging to the phosphate groups of UDP-GlcA. After the addition of the enzyme and incubation for 15 minutes, all of the starting material was converted to UDP-ApI and UDP-Xyl, which displayed \(^{31}\text{P} \) NMR signals that were similar in chemical shift to each other and to those of UDP-GlcA. After extended incubation (10 h), the UDP-ApI non-enzymatically decomposed to α-D-apio-D-furanosyl 1,2-cyclic phosphate (19 ppm) and UMP (3 ppm). UMP and UDP-Xyl standards were
spiked into the reaction mixture to confirm the identity of these signals. A control reaction lacking AXS1 did not lead to the formation of the products.

Figure 3.6 ³¹P NMR spectra monitoring the enzymatic conversion of UDP-GlcA to UDP-Api and UDP-Xyl. UDP-Api decomposes to UMP and 1,2 cyclic phosphate after overnight incubation (121.5 MHz, D₂O, 25 °C). * Signal corresponds to NAD⁺.

ESI-mass spectrometry was also used to monitor the enzymatic reaction and showed the conversion of UDP-GlcA to UDP-Api and UDP-Xyl. The initial ESI-mass spectrum, taken before the addition of AXS1, showed a signal at m/z 579 [M-H]⁻. After the addition of AXS1 a
new peak appeared at $m/z$ 535 [M-H]$^-$, corresponding to both UDP-API and UDP-Xyl. Upon extended incubation the decomposition product UMP ($m/z$ 323 [M-H]$^-$) could be detected.

### 3.3.2 Preparation of UDP-[U-$^{13}$C]-D-Glucuronic Acid

In order to provide more direct evidence for the formation of carbon dioxide and a branched-chain sugar in the reaction catalyzed by AXS1, a fully $^{13}$C-labeled hexose skeleton was used in the reaction (UDP-[U-$^{13}$C]-glucuronic acid). The advantage of this substrate is that the reaction can be directly monitored by $^{13}$C NMR spectroscopy, and only the hexose skeleton (or products thereof) will be observed. From the changes in the $^{13}$C-$^{13}$C coupling patterns, the rearrangement or cleavage of the skeleton will become evident. UDP-[U-$^{13}$C]-glucuronic acid was prepared using enzymatic synthesis, starting with commercially available and relatively inexpensive [U-$^{13}$C]-D-glucose (Figure 3.7). [U-$^{13}$C]-D-glucose was first converted to UDP-[U-$^{13}$C]-D-glucose by the actions of hexokinase, phosphoglucomutase and UDP-glucose pyrophosphorylase. This enzymatic synthesis has been used previously with unlabeled glucose$^{153}$ and the activity of all three enzymes was described earlier in this thesis (Section 1.1) as one of the biosynthetic pathways leading to sugar-nucleotides. The transformation begins with phosphorylation at C-6 of [U-$^{13}$C]-D-glucose by hexokinase and ATP. Then phosphoglucomutase reversibly transfers the phosphate group from C-6 to C-1, generating the α-anomer of [U-$^{13}$C]-D-glucose 1-phosphate. The second enzyme requires the presence of glucose-1,6-diphosphate in order to stay active. Finally, the action of UDP-glucose pyrophosphorylase and UTP couples UMP to the resultant [U-$^{13}$C]-glucose 1-phosphate to generate the product UDP-[U-$^{13}$C]-D-glucose. An alternative method for preparing sugar nucleotide diphosphates involves chemical synthesis using phosphoromorpholidates, however, the one-step enzymatic preparation is much simpler and less expensive.$^{154}$
Upon completion of the reaction, all of the enzymes were removed using centrifugal ultrafiltration and the resulting sugar was purified by ion-exchange and size exclusion chromatography. The purified UDP-[U-13C]-D-glucose was then oxidized with UDP-glucose dehydrogenase (UDPGlcDH) and NAD⁺. Unlike the first three enzymes which were purchased from Sigma, the UDPGlcDH was prepared in the laboratory. The plasmid pGAC147, which was used previously in this laboratory to express UDPGlcDH from Group A Streptococci, was transformed into E. coli. The cells were grown in the presence of chloramphenicol, and expression was induced by IPTG. After 3 hours of growth, the cells were harvested by centrifugation and stored in the pellet form at -80 °C. When the enzyme was needed, the cells were lysed and the insoluble cell debris was removed by ultracentrifugation. The resulting crude UDPGlcDH was used in the synthesis without further purification. The reaction of UDPGlcDH
with UDP-[U-\textsuperscript{13}C]-glucose and NAD\textsuperscript{+} was monitored by negative ESI-mass spectrometry to ensure that all of the starting material was converted to UDP-[U-\textsuperscript{13}C]-glucuronic acid (m/z 585 [M-H]). The resulting sugar was purified by ion-exchange chromatography to give the triethylammonium salt of UDP-[U-\textsuperscript{13}C]-D-glucuronate in 74% overall yield.

3.3.3 Testing the Activity of AXS1 Using UDP-[U-\textsuperscript{13}C]-Glucuronic Acid

The purified UDP-[U-\textsuperscript{13}C]-glucuronic acid was incubated with AXS1 in a sealed NMR tube and the progress of the reaction was followed by \textsuperscript{13}C NMR spectroscopy (Figure 3.8). The \textsuperscript{13}C NMR signals in the region of the spectrum between 85 and 180 ppm were particularly informative for observing this conversion. The initial \textsuperscript{13}C NMR spectrum, recorded before the addition of AXS1, presented a signal at 176 ppm corresponding to the carboxylate group. This signal appeared as a doublet due to a coupling between C-5" and C-6" of the sugar. A doublet corresponding to C-1" of UDP-[U-\textsuperscript{13}C]-glucuronic acid was found at 94.5 ppm (Figure 3.8, t= 0 min). After incubation with AXS1, a new signal appeared as a singlet at 160 ppm corresponding to carbon dioxide (Figure 3.8, t= 40 min). If the NMR tube was left open overnight, the signal at 160 ppm disappeared. The spectrum also showed the appearance of two new doublets at 95.2 and 98.2 ppm belonging to C-1" of UDP-Api and C-1" of UDP-Xyl, respectively. This further confirms that AXS1 is able to catalyze the formation of both UDP-Api and UDP-Xyl.
Figure 3.8 ¹³C NMR spectra monitoring the enzymatic conversion of UDP-[U-¹³C]-GlcA to UDP-[U-¹³C]-Api and UDP-[U-¹³C]-Xyl (150 MHz, D₂O, 25 °C).

In order to confirm that the signals attributed to UDP-[U-¹³C]-Api were associated with a sugar containing a branched carbon chain, the ¹³C NMR spectra of UDP-[U-¹³C]-Api and UDP-[U-¹³C]-Xyl were fully assigned. The first step was to assign the signals due to UDP-[U-¹³C]-Xyl. This was done by allowing a reaction to proceed to completion and to incubate long enough for the complete conversion of UDP-[U-¹³C]-Api into the cyclic phosphate breakdown product. The UMP and the α-D-apio-D-furanosyl 1,2-cyclic phosphate were removed by ion exchange chromatography, and the isolated UDP-[U-¹³C]-Xyl was characterized by ¹³C NMR spectroscopy (spectra not shown). Once the signals that belong to UDP-[U-¹³C]-Xyl were assigned, the signals due to UDP-[U-¹³C]-Api in the mixture of sugar nucleotides could also be assigned (Figure 3.9). The signal for the key tertiary C-3"""" carbon of UDP-[U-¹³C]-Api appears
downfield at 76 ppm as a doublet of doublets of doublets due to the coupling constants between $J_{C-2'',C-3''}$, $J_{C-3'',C-4''}$ and $J_{C-3'',C-5''}$.

![Chemical structures](image1.png)

**Figure 3.9** $^{13}$C NMR spectrum of UDP-[$U-^{13}$C]-Api and UDP-[$U-^{13}$C]-Xyl (150 MHz, D$_2$O, 25 °C).

The enzymatic reaction was also monitored by ESI-mass spectrometry and showed the conversion of UDP-[$U-^{13}$C]-GlcA to UDP-[$U-^{13}$C]-Api and UDP-[$U-^{13}$C]-Xyl. The initial ESI-mass spectrum, taken before the addition of AXS1, showed a peak at $m/z$ 585 [M-H]$^-$. After the addition of AXS1 a new peak appeared at $m/z$ 540 [M-H]$^-$, corresponding to both UDP-[$U-^{13}$C]-Api and the UDP-[$U-^{13}$C]-Xyl.
3.4 Catalytic Competence of UDP-4-ketoxylose

3.4.1 Preparation of UDP-4-ketoxylose

The formation of both UDP-API and UDP-Xyl is thought to occur through a common intermediate, UDP-4-ketoxylose (Figure 1.39). Several studies had already been performed on UDP-API synthase from *Petroselinum hortense* to provide indirect evidence for the existence of the 4-keto intermediate, however, this compound has never been isolated from the enzymatic reaction. In order to test whether the UDP-4-ketoxylose is catalytically competent to act as an intermediate in the AXS1 reaction, I wished to incubate it with the enzyme and demonstrate that it could be converted into a mixture of UDP-Xyl and UDP-API. This process would result in an overall reduction of the UDP-sugar, and therefore would require a stoichiometric amount of the enzyme in the NADH-containing form. Since 90% of our recombinant AXS1 contained bound NADH, it could be used directly in these studies. This would also provide us with a method for generating the NAD\(^+\)-containing form of the enzyme, without the need for denaturation. The chemical synthesis of UDP-4-ketoxylose would certainly pose a difficult challenge, given the variety of sensitive groups in the molecule. Fortunately, this compound is known in nature, and thus an enzymatic synthesis could be employed.

UDP-4-ketoxylose is found as an intermediate in the biosynthesis of UDP-N-formyl-4-amino-4-deoxy-L-arabinose (Figure 3.10). This sugar-nucleotide is used in the modification of normal lipopolysaccharide structures in *E. coli* and *Salmonella typhimurium*. It was found that the addition of *N*-formyl-4-amino-4-deoxy-L-arabinose to lipid A caused the bacteria to be resistant to polymixin and cationic antimicrobial peptides. The biosynthesis begins with the conversion of UDP-glucose to UDP-GlcA by the action of UDP-glucose dehydrogenase.
UDP-GlcA is then oxidatively decarboxylated by the bifunctional enzyme ArnA to produce the UDP-4-ketoxylose. This activity of ArnA is very similar to that of UDP-GlcA decarboxylase (Section 1.8) with the exception that it consumes NAD\(^+\) and does not reduce the C-4” carbonyl following decarboxylation. This 4-keto sugar is subsequently converted to a 4-amino sugar by the pyridoxal phosphate-dependent enzyme ArnB. The biosynthesis is completed by the second activity of ArnA that transfers a formyl group to give UDP-N-formyl-4-amino-4-deoxy-L-arabinose. Ultimately, this sugar is incorporated into lipid A by the action of several more enzymes.\(^{158}\)

![Diagram of biosynthesis of UDP-4-amino-4-deoxy-L-arabinose](image)

Figure 3.10 Biosynthesis of UDP-4-amino-4-deoxy-L-arabinose in polymyxin-resistant *E. coli* and *Salmonella typhimurium*. 
The *arnA* gene was amplified from genomic *E. coli* K-12 W3110 DNA by PCR and cloned using a procedure similar to the one described in Section 3.2.2. The resulting plasmid encoding an N-terminal hexahistidine-tagged ArnA was then transformed into competent *E. coli*. The cells were grown in kanamycin-containing media and protein expression was induced by the addition of IPTG. The cells were harvested, lysed, and the soluble fraction was purified using nickel affinity column chromatography. SDS-PAGE analysis of the purified protein revealed a single band at ~77 kDa. That is consistent with the expected, previously reported molecular weight, and indicated a purity of greater than 80% (Figure 3.11).\(^{155}\)

![Figure 3.11 SDS-PAGE gel showing ArnA purification. Lane 1: molecular weight standards; lane 2: Cell lysate (before IPTG); lane 3: Crude cell lysate 5 h after IPTG induction; Lane 4: Purified ArnA. Molecular weight standards BSA (29 kDa) and carbonic anhydrase (66 kDa).](image)

UDP-GlcA was incubated with NAD\(^+\) and purified ArnA in phosphate buffer (pH 7.5). The progress of the reaction was monitored by negative ESI-mass spectrometry, or more specifically, by monitoring the appearance of the unhydrated UDP-4-ketoxylose sugar (\(m/z\) 533 [M-H]) as well as the hydrated UDP-4-ketoxylose (\(m/z\) 551 [M-H]) that are formed in a 15 to 1 ratio as estimated by the intensity of the ESI-mass spectra signals. In solution, UDP-4-ketoxylose would
readily and reversibly react with water to form a hydrated version of the ketone (Figure 3.12 top). Similarly, in the presence of a high concentration of MeOH, a hemiacetal can be reversibly formed (Figure 3.12 bottom), as will be apparent in the later experiments. Once the reaction was complete, the enzyme was removed by centrifugal ultrafiltration and the resulting sugar was purified by ion exchange chromatography. The UDP-4-ketoxylose was characterized by $^1$H NMR spectroscopy (Appendix Figure A.2) and the observed values are in agreement with previously published results.\textsuperscript{155} According to the integration of the $^1$H NMR spectra signals, the UDP-4-ketoxylose is found predominantly (50:1) as the hydrated compound in D$_2$O.

![Figure 3.12 Reaction of UDP-4-ketoxylose with H$_2$O and MeOH.](image)

3.4.2 Reaction of UDP-4-ketoxylose with AXS1

The proposed reaction between the reduced form of AXS1 and UDP-4-ketoxylose is shown in Figure 3.13A. In the first attempt at monitoring this reaction, a large excess (100-fold) of UDP-4-ketoxylose was incubated with AXS1 and changes in the oxidation state of the enzyme bound cofactor (90% NADH at $t=0$) were followed by UV spectroscopy (Figure 3.13B). Before addition of the substrate, there was a large absorption peak at 355 nm belonging to the tightly
bound NADH. Having most of the AXS1 enzyme in the NADH form is beneficial for this experiment because the proposed reaction is a net reduction and requires a stoichiometric amount of the reducing agent. No changes were observed immediately following the addition of ketose, indicating that, if the reaction does occur, it is extremely slow. However, after incubating the mixture for 20 hours at 37 °C, there was a significant decrease in the NADH band (~70%) due to reduction of the 4-keto sugar that generates NAD⁺. The excess UDP-4-ketoxylose was removed using centrifugal ultrafiltration, and NaBH₄ was added to the enzyme solution. Immediately after the addition of NaBH₄, the intensity of the band at 355 nm increased, and a new band at 425 nm appeared. The 425 nm shoulder is due to 1,2- and 1,6-reduction products, which are known to form upon reduction of NAD⁺ with NaBH₄. As shown by the UV spectrum, not all of the NADH cofactor bound to AXS1 was oxidized to NAD⁺ even though a large excess of substrate was added. As a control reaction, incubation of AXS1 in phosphate buffer alone was also monitored by UV spectrometry. No decrease in the intensity of the NADH band was observed. The fact that the reaction did not proceed to completion was attributed to a very slow rate of the reaction, combined with a loss of enzyme activity upon prolonged incubation. This kinetic argument is favoured over the thermodynamic argument, since it is expected that reduction of a cyclohexanone functionality by NADH should be a favourable process. This is because the presence of an sp² center in a cyclohexyl ring introduces torsional strain into the molecule. Similar arguments explain why the cyclohexanone is largely hydrated in solution.
A second experiment was designed in order to detect the sugar nucleotide products of the reaction. In this case a 1:1 molar ratio of enzyme to substrate was employed in order to maximize the percentage of ketone converted to product. The progress of the reaction was followed by diluting aliquots into a 9:1 mixture of H₂O:MeOH and analyzing the products by negative ESI-mass spectrometry (Figure 3.14). Before the addition of AXS1, there were three major peaks that
belonged to the starting material UDP-4-ketoxylose. The signal at \( m/z \) 533 corresponds to the unhydrated version of starting material, the signal at \( m/z \) 551 can be assigned to the hydrated version and the signal at \( m/z \) 565 can be assigned to the MeOH adduct of the ketone. After AXS1 was added, the mixture was incubated for 20 hours at 37 °C before an ESI-mass spectrum was recorded. A new signal had appeared at \( m/z \) 535 [M-H]⁻ that could correspond to either UDP-Xyl or UDP-Api. Since UDP-Api is not stable under prolonged incubation times, the signal was tentatively assigned to UDP-Xyl. The signal at \( m/z \) 540 belongs to an NAD⁺ fragment. A weak signal corresponding to UMP was observed at \( m/z \) 323 [M-H]⁻, however, no signal corresponding to the cyclic 1,2 phosphate species was detected. Thus, it was not possible to definitively conclude that UDP-Api had been formed during the reaction. Given that only a fraction of the ketone had been converted to products (10-20%), and that the sample was inherently diluted due to the stoichiometric nature of the reaction, the absence of the signal could easily be attributed to a poor signal-to-noise ratio.
The observation that incubation of the enzyme with UDP-4-ketoxylose results in production of NAD\(^+\) and UDP-Xyl supports the notion that the UDP-4-ketoxylose is a true intermediate in the enzymatic reaction. It is somewhat surprising that the rate of the reaction between the intermediate and the enzyme is extremely slow; however, this is not unprecedented. There are several possible reasons for such low turnover of the intermediate. The conformation and/or protonation state of the free enzyme in solution may differ considerably from that which normally binds the intermediate during the course of the enzymatic reaction. This could introduce a significant kinetic barrier towards formation of the productive enzyme-intermediate complex. Alternatively the intermediate may adopt a different form (such as the hydrate of a ketone) when free in solution. The latter case was observed when testing the intermediate of the
enzyme ribulose bisphosphate carboxylase that catalyzes the formation of two molecules of 3-phosphoglycerate from ribulose 1,5-bisphosphate and CO$_2$. The intermediate in the reaction is 2-carboxy-3-keto-d-arabinitol 1,5-bisphosphate (CKABP).$^{160}$ The intermediate was found to exist as a hydrate in the active site of the enzyme, whereas the keto form of CKABP is predominant in solution. It was reported that feeding of CKABP to the activated enzyme resulted in a significantly lower turnover rate as compared to the natural substrate.$^{160}$ It is thought that the keto-CKABP is not a kinetically competent intermediate, because the keto version is not the enzyme-bound intermediate formed during the reaction. The rate of conversion of exogenously added intermediate into product is thought to depend on the rate of the hydration of the ketone in the enzyme active site.$^{161}$ In the case of AXS1 it is believed that the keto version of the UDP-4-ketoxylose is bound to the enzyme.$^{102, 111}$ However, in solution the UDP-4-ketoxylose is predominantly in the hydrated form.$^{155}$ Therefore, the rate of conversion of UDP-4-ketoxylose by AXS1 may depend on the rate of dehydration that generates the keto form of the sugar (either in the active site or in solution). Moreover, the free enzyme would normally contain NAD$^+$ prior to binding substrate for catalysis. In this case, the free enzyme responsible for reducing UDP-4-ketoxylose must contain an NADH cofactor, and this may alter the normal protonation state/conformation of the active site. It is possible that, when the enzyme is bound to the NADH form of the cofactor, the active site is “clamped down” so that UDP-4-ketoxylose cannot enter it. The barrier towards adopting a productive conformation/protonation state may slow the reaction considerably.

### 3.5 Testing UDP-Xylose as a Potential Substrate

Previous studies had indicated that UDP-Xyl could not be converted into UDP-Api by UDP-D-apiose/UDP-D-xylose synthase.$^{111, 112, 162}$ This seems counterintuitive since all the steps
past the decarboxylation in either of the proposed mechanisms (Figure 1.39 and Figure 1.43) would be expected to be reversible. In order to try to resolve this discrepancy, I decided to perform a reaction between UDP-Xyl and AXS1. The progress of the reaction was monitored by \(^{31}\text{P}\) and \(^{1}\text{H}\) NMR spectroscopy. After 24 hours of incubation, no UDP-Api or its decomposition products were detected. One possible explanation for the observed results is that the normal product ratio of UDP-Xyl to UDP-Api (approx. 1:1) represents a kinetic ratio of products and does not reflect the relative thermodynamic stabilities of the sugar-nucleotides. If UDP-Xyl is thermodynamically much more stable than UDP-Api, than the conversion of UDP-Xyl to UDP-Api would occur more slowly than the conversion of UDP-Api to UDP-Xyl, and the product would not be expected to accumulate. The extremely low activity of recombinant AXS1 is likely the reason why the interconversion of UDP-Xyl to UDP-Api cannot be detected.

### 3.6 Attempted Synthesis of UDP-D-Apiose

The inability to observe the AXS1-catalyzed conversion of UDP-Xyl into UDP-Api may be due to the lower thermodynamic stability of the latter compound. If this were the case, the conversion of UDP-Api into UDP-Xyl should be much more readily detected. Therefore, I attempted to undertake the first chemical synthesis of UDP-Api in order to obtain a pure sample of this substrate. It had been reported that UDP-Api has a relatively short half life at 37 °C and pH 8, however, only 6% of a sample of UDP-Api degraded upon storage for 120 days at -20 °C and pH 6.4.\(^{110}\) Moreover, the total synthesis of UDP-galactofuranose, a sugar-nucleotide that has a similar stereochemical orientation at C-1" and C-2" as UDP-Api (Figure 3.15), had been reported.\(^{163}\) Both sugars are susceptible to the formation of a 1,2-cyclic phosphate decomposition product, yet a 35% yield of UDP-galactofuranose was reported for the coupling
reaction of the corresponding sugar 1-phosphate and activated 5'-UMP. This suggests that the synthesis of UDP-Api is possible.

![UDP-Api and UDP-galactofuranose](image)

**Figure 3.15** Structures of UDP-Api and UDP-galactofuranose.

A key aspect in preparing UDP-Api is that the open chain form of d-apiose can cyclize into four possible closed-chain configurations (Figure 3.16). There are three potential ways to keep the sugar in the *erythro* form during the synthesis. The first approach is to protect the anomic hydroxyl group so the open-chain form cannot be produced. The second method involves protecting the C-3' hydroxyl group so that it cannot be involved in cyclization. Finally the third method is to protect the C-2 and C-3 hydroxyl groups as a *cis*-fused 5-membered acetal. All of these methods will avoid the formation of the *threo* products in which the C-2 and C-3' hydroxyl groups bear a *trans* relationship, and are used at some point in the following synthetic scheme.
Commercially available D-xylose was converted to the known compound 4 in five steps (Figure 3.17).\textsuperscript{164,165} An initial protection with 2-methoxypropene gave aldehyde 1 which, upon treatment with formaldehyde and base, yielded diol 2 in an aldol-Canizzaro type reaction. Treatment of 2 with dilute acetic acid selectively removed the cyclic acetal from C-4 and C-5, and the addition of sodium meta-periodate in the presence of base led to the formation of the aldose, which rapidly cyclized to the designed erythro-furanose form 3. The anomeric position was then protected as a methyl glycoside by treatment with trimethyl orthoformate in the presence of pyridinium p-toluenesulfonate to give 4 as the mixture of two anomers with the $\beta$-anomer being the major product. Benzyla
tion of the mixture gave the fully protected compound 5 and the removal of the methoxy group followed by acetylation led to the formation of known compound 6 as a mixture of two anomers.\textsuperscript{166} Unfortunately, the selective anomeric deacetylation proved to be problematic. It was found, that once the acetyl group at C-1 had been removed, a
migration of the acetyl group from C-2 to C-1 readily occurred. As a result, the isolated material was a mixture of deacetylated products and 7 could not be obtained in the pure form.
Figure 3.17 Proposed synthetic route for the formation of UDP-Api.
In order to avoid the deacetylation step, an alternative synthetic route was designed (Figure 3.18). The major modification consisted in the use of benzyl protecting groups for all of the hydroxyl groups in order to eliminate the chance of migration. Compound 3 was treated with sodium hydride followed by benzyl bromide to form 11, which was subsequently deprotected with formic acid to give a previously reported compound 12.\textsuperscript{166} The remaining benzyl group at C-3' prevented the formation of any L-threo-furanose structures. All of the hydroxyl groups were then protected using a treatment with benzyl alcohol/HCl, followed by benzyl bromide/NaH to form the fully protected apiose 13 as a mixture of anomers. Generally, benzyl protecting groups are cleaved by hydrogenolysis, however, anomeric benzyl groups can be selectively removed using acid.\textsuperscript{166} Treatment of 13 with formic acid led to formation of 14 as an α/β mixture of anomers. The $^1$H NMR spectrum of the anomerically deprotected 14 is shown in the Appendix as Figure A.3. Sugar 14 was phosphorylated using a phosphoramidite reagent followed by an oxidation to generate 15. The progress of the reaction was monitored by positive ESI-mass spectrometry. All of the starting material was converted to 15 ($m/z$ 703 [M+Na$^+$]), however, all attempts to purify the resulting sugar using flash chromatography only led to its decomposition. The use of base-treated silica gel or alumina was also unsuccessful and resulted in the partial or full decomposition of 15 into a variety of different products. It was consequently decided that 1-phosphoapiose derivatives are too unstable to synthesize and no further attempts towards the synthesis of UDP-Api were made.
Figure 3.18 Alternative synthetic route for preparation of anomerically deprotected apiose.

3.7 Fluorinated Analogs of UDP-GlcA

3.7.1 Introduction

Fluorinated substrate analogs are widely used as mechanistic probes and inhibitors in order to gather information about the catalytic mechanisms of enzymes. They have been employed to investigate various types of reactions including anionic, cationic, oxidation-reduction, and radical-based processes. One advantage of fluorine lies in the fact that it can mimic a hydroxyl group due to its similar size and electronegativity. It can also act as a leaving group in reactions involving carbanionic intermediates and as a destabilizing group in reactions involving carbocationic intermediates. One example of the use of fluorinated analogs was reported in the case of CDP-D-glucose 4,6-dehydratase from Yersinia pseudotuberculosis. The generally accepted mechanism for the reaction catalyzed by 4,6-dehydratase is discussed in Section 1.6.1.
Researchers prepared CDP-6-deoxy-6,6-difluoroglucose\cite{170} and found it to be an irreversible inhibitor, which bound covalently to the active site of the enzyme. It is thought that CDP-6-deoxy-6,6-difluoroglucose is a suicide substrate operating by the mechanism shown in Figure 3.19. After the initial oxidation, a deprotonation at C-5\textsuperscript{\prime} would lead to the elimination of HF, giving a fluoro-enone intermediate. Subsequently, NADH would deliver a hydride to C-6\textsuperscript{\prime}, generating an enolate ion that would eliminate the second fluoride to give an \(\alpha,\beta\) unsaturated ketone. Since the active site now bears NAD\(^+\), this intermediate could not be reduced further. Instead, it would react with an active site nucleophile, covalently modifying the active site. This proposed mechanism is supported by the facts that fluoride ions were released in solution during the reaction and that mass spectral analysis indicated the formation of a covalent adduct between the enzyme and the inhibitor.\cite{170}
Another example of the use of fluorinated substrate analogs to investigate an enzyme mechanism was reported with 1-deoxy-D-xylulose-5-phosphate (DXP) reductoisomerase, which is a NADPH-dependent enzyme that catalyzes the conversion of DXP to methyl-D-erythritol 4-phosphate (MEP, Figure 3.20A).\textsuperscript{171} There are two proposed mechanisms for the conversion of DXP to MEP and they closely mirror the suggested mechanisms for the rearrangement catalyzed by AXS1. The first mechanism involves an α-ketol rearrangement (alkyl migration) to produce the methylerythrose phosphate as an intermediate, which is then reduced by NADPH to produce the desired product (Figure 3.20B).\textsuperscript{171} The other proposed mechanism involves a retro-aldol rearrangement to form the common methylerythrose phosphate, which is then reduced to form
DXP (Figure 3.20C). In order to try to provide evidence for one of the mechanisms three fluorinated substrate analogs were prepared (Figure 3.20D). It was found that the 1-fluoro-DXP substrate analog acted as a poor substrate, while the 3-fluoro- and 4-fluoro-DXPs were noncompetitive inhibitors. The results suggest that both the C-3 and C-4 hydroxyl groups of DXP were crucial for the catalysis.
Figure 3.20 A) Reaction catalyzed by 1-deoxy-D-xylulose-5-phosphate (DXP) reductoisomerase. B) α-ketol rearrangement mechanism. C) Retro-aldol rearrangement mechanism. D) Fluorinated substrate analogs.

The reaction catalyzed by UDP-D-apiose/UDP-D-xylene synthase is proposed to occur either via a retro-aldol mechanism (Figure 1.39) or a carbon migration mechanism (Figure 1.43).
In an attempt to gain more insight into the mechanism of AXS1, I prepared UDP-2-deoxy-2-fluoro-D-glucuronic acid (UDP-2F-GlcA) and UDP-3-deoxy-3-fluoro-D-glucuronic acid (UDP-3F-GlcA, Figure 3.21) and examined their competence as substrates. For both fluorinated substrates it is expected that the corresponding fluorinated UDP-xylose products could be formed regardless of the mechanism employed, because neither the C-2" hydroxyl nor the C-3" hydroxyl plays a key role in UDP-xylose formation. A productive reaction will indicate that the fluorine substitution did not affect binding or the ability of the enzyme to catalyze the hydride transfer and decarboxylation steps. In the case of UDP-2F-GlcA, the formation of the UDP-2-deoxy-2-fluoroapiose would only be observed if the carbon migration mechanism was operative. This is because deprotonation of the C-2 hydroxyl is not required in this mechanism, yet is required for the retro-aldol process. In the case of UDP-3F-GlcA, no UDP-3-deoxy-3-fluoroapiose would be expected to form if either mechanism were at play. However, it may be possible to detect alternative ring-opened by-products formed by a retro-aldol ring opening that leads to the formation of unnatural intermediates.
Figure 3.21 Proposed results of the incubation of fluorinated analogs with AXS1.

3.7.2 Synthesis of UDP-2-Deoxy-2-fluoro-D-glucuronic acid

UDP-2-deoxy-2-fluoro-D-glucuronic acid (UDP-2F-GlcA, 22) was prepared following the synthetic route shown in Figure 3.22. Compound 20 was prepared according to a previously published procedure with minor modifications. The key transformation is the introduction of
the fluorine group at C-2\textsuperscript{\("\)}} which was accomplished with Selectfluor\textsuperscript{TM} (1-chloromethyl-4-fluoro-
1,4-diazoniabicyclo[2.2.2]octane bis(tetrafluoroborate)), a commonly used fluorinating reagent in carbohydrate chemistry.\textsuperscript{173} Treating the commercially available glycal 16 with Selectfluor\textsuperscript{TM} produced a mixture of the desired compound 17a and its C-2\textsuperscript{\("\)}} epimer 17b. Both sugars were acetylated and separated by flash chromatography to give 18 in 25\% yield over two steps. The inversion of configuration at C-1\textsuperscript{\("\)}} of 18 was accomplished in two steps by first generating the \(\alpha\)-glycosyl bromide using HBr, followed by a reaction with silver(I) acetate to produce the desired \(\beta\)-glycosylacetate 19 in 95\% yield. In order to introduce the \(\alpha\)-phosphate, a modification\textsuperscript{174} of the method of MacDonald\textsuperscript{175} was used. This procedure allowed the stereospecific preparation of the fully acetylated \(\alpha\)-phosphate that was immediately deacetylated with LiOH to give the free sugar 1-phosphate 20 in 51\% yield. The remaining steps were carried out enzymatically. Sugar 20 was dissolved in Tris-HCl buffer pH 7.8 containing UTP and UDP-glucose pyrophosphorylase. After completion of the reaction, 21 was purified using ion-exchange chromatography. The purified sugar nucleotide was then oxidized with UDP-glucose dehydrogenase and NAD\textsuperscript{+} to yield UDP-2F-GlcA 22 in 67\% yield. The final product was fully characterized by \(^1\)H NMR spectroscopy (Appendix Figure A.4) and negative ESI-mass spectrometry (\(m/z\) 581 [M-H]).
Figure 3.22 Synthesis of UDP-2-deoxy-2-fluoro-D-glucuronic acid 22.

3.7.3 Testing UDP-2F-GlcA as a Substrate Analog

Once the 2-fluoro analog 22 was prepared, it was incubated with AXS1 in phosphate buffer at pH 8.0 in the presence of 1 mM NAD$. The progress of the reaction was monitored by $^{19}$F NMR spectroscopy (Figure 3.23) and negative ESI-mass spectrometry. The initial $^{19}$F NMR spectrum, measured before the addition of AXS1, showed a single fluorine signal at -201.16 ppm assigned to the fluorine atom of UDP-2F-GlcA (Figure 3.23, $t=0$). After incubation with AXS1,
a new signal appeared at -201.37 ppm corresponding to either UDP-2-deoxy-2-fluoroxyllose (UDP-2F-Xyl, Figure 3.23, t= 12 h) or UDP-2-deoxy-2-fluoroapiose (UDP-2F-Api). After further incubation, almost all of the starting material was converted to the new product. These results were confirmed by the ESI-mass spectrometric analysis. Before the addition of enzyme, the mass spectrum showed a signal at m/z 581 [M-H]⁺ which belonged to the starting material UDP-2F-GlcA. After 24 hours of incubation with AXS1, all of the starting material was converted to a product with a MS signal at m/z 537 [M-H]⁺, which corresponded to either UDP-2F-Xyl or UDP-2F-Api. It should be noted that, unlike UDP-Api, UDP-2F-Api would not be expected to decompose under extended incubation due to the lack of the C-2" hydroxyl.

![Diagram](image)

**Figure 3.23 Partial^{19}F NMR spectra monitoring the reaction of UDP-2F-GlcA with AXS1 over a 24 h period (282.4 MHz, D₂O, 25 °C).**
In order to decide whether the newly formed product was UDP-2F-Xyl or UDP-2F-Api a sample was purified by ion exchange chromatography and analyzed by \(^1\)H NMR spectroscopy (Figure 3.24). If the isolated product was UDP-2F-Api then we would not expect to see J coupling between H-2'' and H-3'' because the tertiary C-3'' of UDP-2F-Api does not bear a hydrogen. Analysis by 2D COSY NMR spectroscopy (correlated spectroscopy) indicated that there is coupling between H-2'' and H-3'' (Appendix Figure A.5) in the isolated product, and the analysis of the J values was consistent with a sugar in the xylose configuration. This means that the carbon skeleton remained intact and the isolated product is indeed the UDP-2F-Xyl. No presence of UDP-2F-Api was detected either by \(^{31}\)P/\(^{19}\)F NMR spectroscopies or ESI-mass spectrometry. The activity of the enzyme towards the substrate analog 22 was estimated to be 100-fold lower than with the natural substrate.
A control reaction run without AXS1 did not lead to the formation of the same mixture. Moreover, a control reaction with the natural substrate UDP-GlcA led to the formation of both UDP-Api (or its cyclic decomposition product) and UDP-Xyl. The fact that UDP-2F-Xyl but not UDP-2F-Api was detected during the reaction between UDP-2F-GlcA and AXS1 (Figure 3.25) supports the retro-aldol mechanism for the carbon skeleton rearrangement (Figure 3.21). The presence of the 2-fluoro group does not prevent C-4" oxidation and decarboxylation, and therefore permits the formation of the UDP-2-deoxy-2-fluoro-4-ketoxylose intermediate. Once the intermediate is formed, it can be reduced by the enzyme to generate the UDP-2F-Xyl.
However the carbon skeleton rearrangement cannot be initiated because it requires the deprotonation of the hydroxyl group at C-2” to promote the retro-aldol ring opening process.

Figure 3.25 Proposed reaction between UDP-2F-GlcA 22 and AXS1.

If the alkyl migration mechanism were at play in this enzyme (Figure 1.43), one might expect that UDP-2F-Api could be still generated since the C-2” is not intimately involved in the rearrangement process. However, the results from this experiment do not completely rule out the alkyl migration mechanism, since it is possible that the presence of the fluorine at C-2” significantly reduces the rate of the alkyl migration relative to the rate of C-4” reduction. Therefore, the next set of experiments focused on UDP-3F-GlcA as a substrate analog.
3.7.4 Synthesis of UDP-3-Deoxy-3-fluoro-D-glucuronic acid

UDP-3-deoxy-3-fluoro-D-glucuronic acid (UDP-3F-GlcA, 28) was synthesized in seven steps from D-glucose (Figure 3.26). The synthesis of compound 27 had been previously reported. D-Glucose was first protected as the bis-acetonide 23 and then the C-3 hydroxyl was oxidized to give the ketone 24. Next, a stereospecific reduction of 24 gave 1,2:5,6-di-O-isopropylidene-α-D-allofuranose (25). It is thought that the access by reducing agent to the bottom face of the ring is hindered by the presence of the 1,2-isopropylidene group, therefore exclusive reduction from the top face of the ring is observed. The introduction of the fluorine atom at C-3 to give 26 was accomplished with the use of the fluorinating reagent, DAST (diethylaminosulfur trifluoride), which directly replaces a hydroxyl group with fluorine. DAST is relatively mild and can be used with acid-sensitive compounds. The reaction is believed to occur through a SN2-like displacement with an inversion of stereoconfiguration. 3-Deoxy-3-fluoro-D-glucose (27) was obtained after removal of both isopropylidene groups under acidic conditions. Compound 27 was then incubated with hexokinase, phosphoglucomutase, and UDP-glucose pyrophosphorylase followed by UDP-glucose dehydrogenase, to form sugar-nucleotide 28 in a 35% yield. The final product was purified by ion exchange chromatography and stored as its triethylammonium salt. It was fully characterized by $^1$H (Appendix Figure A.6), $^{31}$P and $^{19}$F NMR spectroscopy. The negative ESI-mass spectrum of UDP-3F-GlcA displayed a signal at $m/z$ 581 [M-H]$^-$. The $^{13}$C-labeled compound UDP-[U-$^{13}$C]-3-deoxy-3-fluoro-D-glucuronic acid (UDP-[U-$^{13}$C]-3F-GlcA) was prepared using exactly the same procedure, but starting with [U-$^{13}$C]-D-glucose. The $^{13}$C NMR spectrum of UDP-[U-$^{13}$C]-3F-GlcA is shown in Appendix Figure A.7.
Figure 3.26 Synthesis of UDP-3-deoxy-3-fluoro-D-glucuronic acid 28.

3.7.5 Testing UDP-3F-GlcA as a Substrate Analog

Initial studies showed that the reaction of compound 28 catalyzed by AXS1 was extremely slow and required near stoichiometric amounts of enzyme to reach completion. However, when a sample of pure compound 28 (3 mg) was incubated with AXS1 (125 mg) in phosphate buffer at pH 8.0, the enzymatic reaction could be monitored by $^{19}$F and $^{31}$P NMR
spectroscopies. The initial $^{19}$F NMR spectrum, taken before the addition of AXS1, showed a single signal at -200.9 ppm corresponding to the fluorine group of UDP-3F-GlcA (Figure 3.27, t= 0). After incubation with AXS1, a new signal appeared at -122.6 ppm (Figure 3.27, t= 12 h). Addition of sodium fluoride enhanced this signal, indicating that fluoride was released from 28. After further incubation, all of the starting material was consumed and only one peak corresponding to free fluoride ion was observed (Figure 3.27, t= 24 h). No spectral changes were seen in a control sample of 28 in absence of AXS1, indicating that the production of fluoride was enzyme- catalyzed. In the $^{31}$P NMR spectra (spectra not shown), the release of free UDP was observed at the same rate as the formation of the free fluoride in the $^{19}$F NMR spectra.

![Diagram of NMR spectra](image)

Figure 3.27 $^{19}$F NMR spectra monitoring the incubation of UDP-3F-GlcA 28 with AXS1 in 50 mM potassium phosphate pH 8.0 (282.4 MHz, D$_2$O, 25 °C).
The observation of the loss of both UDP and fluoride from compound 28 indicates that a serious structural perturbation of the sugar nucleotide has occurred. It is likely that the enzymatic reaction generated an unstable intermediate which underwent further decomposition in solution. In each experiment the ratio of UDP-3F-GlcA and catalytically active enzyme was around 10 to 1. This means that the enzyme must regenerate the NAD$^+$ with every catalytic cycle. The result of the incubation with UDP-3F-GlcA is somewhat surprising and it is difficult to explain the release of free fluoride and UDP in the solution with either of the proposed mechanisms.

Since only small amounts of product can be obtained in this fashion, and since it will likely be composed of an uncharged pentose (or fragments thereof), analysis of the fate of the glucuronic skeleton poses a serious problem. For this reason, I decided to incubate UDP-[U-^{13}C]-3F-GlcA with AXS1 and monitor the reaction by $^{13}$C NMR spectroscopy. This would have two advantages. Firstly, it would be possible to deduce whether decarboxylation had occurred from the appearance of CO$_2$. Secondly, it might be possible to determine whether the carbon skeleton had fragmented into smaller pieces from analysis of $^{13}$C-$^{13}$C coupling patterns. In the retro-aldol mechanism, a similar cleavage occurs and such an observation would support the mechanism. Therefore, AXS1 was incubated with UDP-[U-^{13}C]-3F-GlcA and the reaction was followed by $^{13}$C NMR spectroscopy (Figure 3.28). Before addition of the enzyme, there was a doublet at 175.1 ppm that can be assigned to the carboxylate group. After incubation for 10 hours, two new signals appeared as a doublet at 180 ppm and a singlet at 159.7 ppm. The singlet at 159.7 ppm was attributed to carbon dioxide, because it has been observed previously with the normal substrate and it disappeared when the reaction tube was left open overnight. The signal at 180 ppm must belong to a carbon that is still attached to the sugar ring since it appears as a doublet due to J coupling between adjacent isotopically labeled carbons. This is likely due to a C-4$^-$-
oxidized intermediate that has not yet been decarboxylated. Ultimately, all of the substrate underwent decarboxylation as the CO₂ signal at 159.7 ppm was the only peak in the downfield region of the spectrum (Figure 3.28 24 h). Unfortunately, the upfield region of the \(^{13}\)C NMR spectra displayed a complex mixture of signals in the 45-120 ppm region and did not provide any further information about the structure of the species that was created during the release of fluoride and UDP. It appears that more than one product is formed during the decomposition of the enzyme-generated intermediate. The progress of the reaction was also followed by ESI-mass spectrometry. Before the addition of enzyme, the mass spectrum showed a signal at \(m/z\) 587 which can be assigned to the starting material UDP-[U-\(^{13}\)C]-3F-GlcA. After 24 hours of incubation with AXS1, all of the starting material was consumed and only a peak at \(m/z\) 403 [M-H]\(^{-}\) that is assigned to UDP could be detected. No other information could be obtained from the mass spectrum. The activity of the enzyme towards 28 was estimated to be 100-fold less than in the case of the unfluorinated substrate. In summary, the treatment of UDP-3F-GlcA with AXS1 resulted in the catalytic formation of UDP, fluoride, CO₂ and unknown carbon species.
Figure 3.28 Partial $^{13}$C NMR spectra monitoring the reaction of UDP-[$U^{13}$C]-3F-GlcA with AXS1 over a 24 h period (150 MHz, D$_2$O, 25 °C).

One possible explanation accounting for the formation of the observed products is shown in Figure 3.29. The reaction is initiated by the oxidation at C-4" to form UDP-3-deoxy-3-fluoro-4-ketoglucuronic acid. A fragmentation of the carbon skeleton then occurs before a decarboxylation reaction takes place. This would mimic the ring-opening step of the normal retro-aldol mechanism (Figure 1.39). The open chain enol(ate) intermediate would not be able to close in a manner that generates the apiose skeleton because the fluoride at C-3" does not allow this aldol reaction to occur. Following a rapid enolization, the C-2" aldehyde could then be reduced to regenerate the NAD$^+$ cofactor and a reduced open chain intermediate that is released...
into solution. This reduction may mimic the reduction of the UDP-3-aldehydo-apiose intermediate that occurs during normal catalysis (Figure 1.39 and Figure 1.43). The proposed reduction at C-2" is favoured over a potential reduction at C-4" since it leaves the fluoromethyl ketone functionality intact and this could later be hydrolyzed non-enzymatically to generate fluoride. Once in solution, the reduced open-chain intermediate would rapidly undergo decarboxylation and decomposition of the carbon skeleton to form free fluoride ion, UDP, and an unknown carbon skeleton. If this decomposition pathway could be confirmed, it would support the retro-aldol mechanism for the carbon skeleton rearrangement, since that mechanism occurs via an open-chain intermediate. It is interesting to note that, unlike the case of the UDP-2F-GlcA, no evidence for the formation of the UDP-3-deoxy-3-fluoroxylose was obtained. It is expected that, if this product were formed, it would be stable and readily identified. The free sugar, 3-deoxy-3-fluoroxylose, is known to be stable in solution,\textsuperscript{181,182} so even if the hydrolysis of the glycosidic bond and UDP release occurred, no fluoride ion would be released. It is clear from the production of CO\textsubscript{2} that C-4" oxidation is occurring, however, an alternative step (such as retro-aldol ring opening) must take place before decarboxylation and reduction can generate UDP-3-deoxy-3-fluoroxylose.
Figure 3.29 Proposed reaction between UDP-3F-GlcA 28 and AXS1.
3.8 Conclusions and Future Directions

The gene encoding the protein involved in the biosynthesis of UDP-D-apiose was cloned from the cDNA of Arabidopsis thaliana and expressed in E. coli. It was found that 90% of the isolated AXS1 enzyme had the wrong form of the cofactor (NADH) bound in the active site. The incubation of UDP-D-glucuronic acid with AXS1 led to the formation of a mixture of UDP-D-apiose and UDP-D-xylose roughly in a 1:1 ratio and to the release of CO₂. The progress of the reaction was monitored by ¹H, ³¹P and ¹³C NMR spectroscopy. The specific activity of AXS1 was very low (~6.7 x 10⁻³ µmol min⁻¹mg⁻¹) possibly because of absence of any post-translational modifications. As previously reported, UDP-D-apiose decomposed under the incubation conditions to α-D-apio-D-furanosyl 1,2-cyclic phosphate and UMP.¹¹⁰

UDP-4-ketoxylose was prepared using the ArnA enzyme that is involved in the biosynthesis of UDP-N-formyl-4-amino-4-deoxy-L-arabinose. The competence of UDP-4-ketoxylose to act as a potential intermediate in the AXS1 reaction was tested. It was found that about 10-20% of UDP-4-ketoxylose was converted to either UDP-Xyl or UDP-Api during the incubation with AXS1. The slow rate of conversion could be attributed to the fact that UDP-4-ketoxylose exists preferentially in the hydrated form in solution whereas in the active site the keto form serves as the true intermediate. Attempts to generate UDP-Api from UDP-Xyl using AXS1 were unsuccessful, presumably due to the instability of the former compound and the low enzyme activity.

Three substrate-based analogs were prepared and tested with AXS1. It was found that the reaction of UDP-2F-GlcA with AXS1 led to the formation of only UDP-2F-Xyl. This result is consistent with the retro-aldol mechanism, because it requires the presence of the C-2” hydroxyl
group to initiate the ring opening. The only products that were identified from the incubation of UDP-3F-GlcA and AXS1 were UDP, free fluoride ion and CO₂. The structure of the resulting carbon skeleton could not be conclusively assigned based on the obtained data. It is postulated that the enzyme releases a reduced open-chain intermediate into solution, which then decomposes to UDP, free fluoride and the unidentified carbon species (Figure 3.29). This is interpreted as being consistent with a retro-aldol mechanism for the AXS1 reaction.

In the future, the conditions required to obtain a more active version of AXS1 need to be found. As mentioned in Chapter 2, one way is to try to prepare the AXS1 enzyme without any modification to the amino acid sequence (such as a His-tag) and to purify the enzyme using traditional chromatographic methods. Another approach would involve the use of cDNA from other plant sources such as *Lemna minor* and *Petroselinum hortense* to see if a more active version of the enzyme can be prepared. Finally, the preparation of recombinant AXS1 using a yeast or insect expression system would be attractive since it may result in the incorporation of post-translational modifications necessary for full activity. Obtaining an X-ray crystal structure of AXS1 would help to identify key residues in the active site of the enzyme and give insight into their role during the enzymatic reaction. A structure could guide site-directed mutagenesis studies that would help to elucidate the individual roles of the residues. For example, the mutations of a residue involved in the retro-aldol rearrangement should eliminate the ability of AXS1 to form UDP-Api but would not affect its ability to form UDP-Xyl.
3.9 Experimental

3.9.1 Materials and General Methods

The previous “Materials and General Methods” from Chapter 2 also applies in this Chapter with the following additions. The isotopically labeled glucose was purchased from Cambridge Isotopes Laboratories. Dry solvents were distilled fresh, using CaH₂ (CH₂Cl₂, pyridine, DMSO) or Na/benzophenone (THF) as drying agent. ¹⁹F NMR spectra were obtained on the Bruker AV300 spectrometer at 282.3 MHz. ¹³C NMR spectra were obtained on the Bruker AV600 spectrometer at 150 MHz.

3.9.2 Source and Cloning of AXSI

Approximately 50 mg of leaf material of Arabidopsis thaliana was kindly provided by Dr. Reinhard Jetter. It was placed in a mortar chilled with liquid nitrogen and ground for several minutes. The tissue powder was transferred into a microcentrifuge tube and the liquid nitrogen was allowed to evaporate. The mRNA of the plant was then isolated using the RNeasy kit (Qiagen).¹⁴⁶ The cDNA was prepared by Ortwin Guhling from Dr. Reinhard Jetter’s laboratory. The AXSI gene (GeneID: 817332) was amplified by PCR using Arabidopsis thaliana cDNA as a template. Oligonucleotide primers, synthesized by the NAPS Unit at UBC, included overhangs for ligation-independent cloning: 5'-GGTATTGAGGGTGCATGGCGAATGGAGCTAATAGAG-3' (forward sequence, AXSI) and 5'-AGAGGAGAGTTAGAGCCGGAGAGTTAGGAAGCC-3' (reverse sequence, AXSI). A general mixture in the PCR tube included: 5.0 μL of 10X PCR buffer (Invitrogen), 1.0 μL of 10 mM dNTP mix, 1.5 μL of 50 mM MgCl₂, ~0.5 μL of Arabidopsis thaliana cDNA, 25 pmol of each primer, 0.25 μL of 5 U/μL Taq polymerase, and distilled H₂O to a total volume of 50 μL. DNA was amplified
using an iCycler Thermal Cycler (Bio-Rad) according to the following cycles: one cycle of 3 min at 94 °C; thirty cycles of 60 s at 94 °C, 60 s at 55 °C, and 90 s at 72 °C; and followed by cooling to 4 °C. The PCR product was cloned into the pET-30 Xa/LIC vector (Novagen) using the ligation-independent cloning method according to the manufacturer's directions. The resulting plasmid was transformed into NovaBlue Gigasingles chemically competent *E. coli* cells (Novagen). The presence of the gene was confirmed by DNA sequencing.

3.9.3 Over-expression and Purification of AXS1

The recombinant *AXS1* plasmid was transformed into *E. coli* BL21 (DE3) competent cells which were incubated in 10 mL terrific broth (TB) medium containing 12 g/L of bacto tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 2.31 g/L of KH₂PO₄ 12.54 g/L of K₂HPO₄ and 35 mg/L kanamycin at 37 °C/225 rpm for 10 h. The overnight culture was then poured into 500 mL of LB medium containing 35 mg/L kanamycin and shaken at 37 °C/225 rpm until an OD₆₀₀ of 0.6 – 1.0 had been reached. Over-expression was induced with 1mM IPTG. After incubation for 20 h at 22 °C, the cells were harvested by centrifugation and stored as a pellet at -80 °C. The pellets were re-suspended in 10 mL of a phosphate buffer (20 mM, pH 8.0) containing 1 mM dithiothreitol (DTT), 1 mM NAD⁺, 1 mg/L of aprotinin, and 1 mg/L pepstatin A at 4 °C. The cells were subsequently lysed by passage through a French Pressure cell at 20 000 psi. The lysate was centrifuged at 6 000×g for 1 h, passed through 0.45 μm and 0.22 μm filters, and loaded onto a column containing 10 mL of Chelating Sepharose Fast Flow resin (Pharmacia Biotech), which was previously charged with 100 mM NiSO₄ and washed with sodium phosphate buffer (20 mM, pH 8.0, containing 0.5 M NaCl and 5 mM of imidazole). The purification process was monitored by a Flow Thru UV monitor Spectrometer at 280 nm and was carried out at 4 °C. Nonspecifically bound proteins were washed away by applying buffers containing first 5 mM, and then 125 mM
imidazole. Finally, bound enzyme was eluted using a 500 mM imidazole buffer. The fractions containing the desired enzyme were combined and concentrated using Amicon Ultra Centricons (Millipore) and stored on ice until required for use. Typical yields of purified protein were ~250 mg/L of cell culture.

### 3.9.4 Source and Cloning of arnA

The arnA gene (EcoGene Accession Number: EG14091) was amplified by PCR using *Escherichia coli* K-12 W3110 genomic DNA as a template. Oligonucleotide primers, synthesized by the NAPS Unit at UBC, included overhangs for ligation-independent cloning: 5'-GGTATTGAGGGTCGCATGAAAACCGTCGTTTTTGCT-3' (forward sequence, arnA) and 5'-AGAGGAGAGGTAGAGCTCATGATGGTTTATCCGTAAG-3' (reverse sequence, arnA). The plasmid containing arnA was prepared according to procedure described in Section 3.9.2.

### 3.9.5 Over-expression and Purification of ArnA

Over-expression and purification of ArnA was carried out as described for PglF (Section 2.8.5). Typical yields of purified protein were ~30 mg/L of cell culture. ArnA could be stored frozen in the storage buffer (25 mM Tris-HCl, pH 8.0, containing 10% glycerol) at −80 °C for several months without loss in activity.

### 3.9.6 Over-expression and Purification of UDP-Glucose Dehydrogenase

The pGAC147 plasmid was transformed into *E. coli* JM109 (DE3) cells and inoculated onto LB-agar plates containing 25 μg/mL chloramphenicol. After overnight incubation at 37 °C, a single colony was used to inoculate 500 mL of TYPG media (8 g bacto-tryptone, 8 g yeast extract, 2.5 g sodium chloride, 1.25 g dibasic potassium phosphate, and 2.5 g D-glucose) containing 25 μg/mL chloramphenicol. The cell culture was grown at 37 °C with vigorous
shaking (280 rpm) until an OD$_{600}$ of 0.7-1.0 IPTG (48 mg) was added at this point to a final concentration of 0.4 mM to induce the over-expression of UDP-glucose dehydrogenase. After 3 h of further growth, the cells were harvested by centrifugation for 15 min at 5000 rpm and the cell pellet was stored at −80 °C.

The cell pellet was later thawed rapidly with warm water and re-suspended in 10 mL cold 50 mM triethanolamine-HCl (Trien-HCl) buffer, pH 8.7, containing 2 mM DTT, 10% (v/v) glycerol, 1.5 mM phenylmethanesulfonyl fluoride, 1 mg/L pepstatin, and 1 mg/L aprotinin. The cells were subsequently lysed by passage through a French Pressure cell at 20 000 psi. The lysate was centrifuged at 6 000×g for 1 h, passed through 0.45 μm and 0.22 μm filters. The crude enzyme was used in reactions without further purifications.

3.9.7 Monitoring Enzyme Incubation by NMR Spectroscopy

The general procedure to monitor enzyme incubations by NMR spectroscopy is as follows. The freshly prepared AXS1 was exchanged into deuterated buffer (pD 7.4 potassium phosphate, D$_2$O) by successive centrifugal filtration steps to affect a 1000-fold dilution of the storage buffer. The enzyme solution was stored at 4 °C or on ice and used within a few hours. The substrate was dissolved with the same deuterated phosphate buffer and concentrations were measured by UV spectroscopy at 262 nm (molar absorptivity = 9890 M$^{-1}$ cm$^{-1}$ in 10 mM potassium phosphate, pH 7.0). Before the addition of enzyme, a spectrum of substrate only, which served as the zero time point, was measured. Incubations were initiated by addition of the enzyme to the NMR tube. Spectra were taken, initially every 5 min, then after progressively longer time intervals.
3.9.8 AXS1 Activity Test with UDP-GlcA

In order to determine the activity of AXS1 under initial velocity conditions, a solution containing UDP-GlcA (5.2 mM) and AXS1 (85 mg) in 10 mM phosphate buffer prepared using D$_2$O containing 1 mM NAD$^+$ (pD 8.0, 1.0 mL total volume) was placed into an NMR tube and immediately monitored by $^1$H NMR spectroscopy. Spectra were taken every two min for a period of one hour while incubating at 25°C. The conversion rate was calculated by comparing the integrals of the signals due to the anomeric proton of UDP-GlcA (5.49 ppm) to those of the products UDP-Api and UDP-Xyl (5.76 and 5.42 ppm). The rate of the reaction was determined from data accumulated during the first 15% of the reaction.

3.9.9 Enzymatic synthesis of UDP-[U-$^{13}$C]-Glucuronic Acid

The enzymatic coupling of UDP to [U-$^{13}$C]-glucose was carried out in an incubation mixture of the following composition: [U-$^{13}$C]-glucose (25 mg, 134 µmol), ATP (227 mg, 412 µmol), UTP (200 mg, 412 µmol), glucose-1,6-diphosphate (2.2 µg, 0.165 µmol), MgSO$_4$ (72 mg, 292 µmol), 70 mM Tris-HCl, pH 7.8, (40 mL), hexokinase (100 units), phosphoglucomutase (210 units), UDP-glucose pyrophosphorylase (25.5 units), and inorganic pyrophosphatase (33 units). The incubation was carried out at 30 °C for 5 h. The progress of the reaction was monitored by $^{31}$P NMR spectroscopy and ESI-mass spectrometry. Once the reaction was complete, the enzymes were removed by centrifugal ultrafiltration and the remaining solution was loaded onto a DE-52 column (220 mL) and eluted with a linear gradient of 0 - 400 mM triethylammonium bicarbonate buffer (800 mL total). The triethylammonium bicarbonate buffer was prepared by bubbling CO$_2$ into 400 mM triethylamine solution over 10 h. The separation was monitored by a UV detector (Spectra/Chrom Flow Thru UV monitor controller, Spectrum) with a 254 nm filter. The fractions containing UDP-[U-$^{13}$C]-glucose were lyophilized to dryness.
The sugar was then dissolved in 50 mM phosphate buffer (10 mL final volume), pH 8.0, containing 5 mM NAD⁺, 2 mM DTT, and a solution (5 mL) of freshly prepared UDP-glucose dehydrogenase, and incubated at 37 °C for 3 h. The progress of the reaction was followed by ESI-mass spectrometry. The sugar was once again purified by a DE-52 column (220mL) and lyophilized to dryness. \(^{13}\text{C NMR (D}_2\text{O)}: \delta 176 (d, J_{c-5', c-6'}= 57 \text{ Hz, C-6''}); 94.86 (d, J_{c-1', c-2'}=46 \text{ Hz, C-1''}); 70.5-73 (m, C-2'', C-3'', C-4'', C-5''). \(^{31}\text{P NMR (D}_2\text{O)}: \delta -10.51 (d, J_{P\alpha, P\beta}= 20.3 \text{ Hz}, -12.51 (m). ESIMS: } m/z 585 [M-H].

3.9.10 Studies with UDP-[U-\(^{13}\text{C}\)]-GlcA

The prepared AXS1 (85 mg) was added to a solution of UDP-[U-\(^{13}\text{C}\)]-GlcA (3.0 mg) dissolved in 850 \mu L phosphate buffer solution pH 8.0 and 1 mM NAD⁺ (1.0 mL final volume). The reaction was incubated at 25 °C. \(^{13}\text{C NMR and negative ESI-mass spectra were acquired at timed intervals. Once the reaction was complete, the mixture was allowed to incubate for another several hours before the enzyme was removed by centrifugal ultrafiltration. The resulting solution was loaded onto a DE-52 column (220 mL) and eluted with a linear gradient of 0 - 400 mM triethylammonium bicarbonate buffer (800 mL total). The triethylammonium bicarbonate buffer was prepared by bubbling CO₂ into 400 mM triethylamine solution over 10 h. The separation was monitored by an UV detector (Spectra/Chrom Flow Thru UV monitor controller, Spectrum) with a 254 nm filter. Fractions containing UDP-[U-\(^{13}\text{C}\)]-Xyl were lyophilized to dryness and analyzed by \(^{13}\text{C NMR. } \(^{13}\text{C NMR (D}_2\text{O)}: \delta 96.7 (d, J_{c-1', c-2'}= 43 \text{ Hz, C-1''}); 73.9 (m, C-3''); 72.6 (m, C-2''); 70.15 (t, J_{c-3', c-4'}= J_{c-4', c-5'} = 43 \text{ Hz, C-4''}); 61.3 (d, J_{c-4', c-5'} = 42.9 \text{ Hz, C-5''}). \(^{31}\text{P NMR (D}_2\text{O)}: \delta -10.71 (d, J_{P\alpha, P\beta}= 20.7 \text{ Hz}, -12.35 (m). ESIMS } m/z 540 [M-H].
3.9.11 Enzymatic Synthesis and Isolation of UDP-4-ketoxylose

UDP-4-ketoxylose was prepared under conditions slightly modified from those described previously. A purified sample of ArnA (7 mg) was added to 50 mL of a phosphate buffer (10 mM, pH 7.0) containing 50 mg of UDP-GlcA disodium salt and 150 mg of NAD. The solution was incubated for 4 h at 37 °C and the reaction progress was followed by negative ESI-mass spectrometry. It was determined that > 95% of the UDP-GlcA (m/z 581, [M-H]) was converted to UDP-4-ketoxylose (m/z 551, [M+H2O-H]) during this time. The enzyme was then removed by centrifugal ultrafiltration and the resultant filtrate was loaded onto a DE-52 column (220 mL) and eluted with a linear gradient of a 0 to 0.5 M triethylammonium bicarbonate buffer. The A254 of the eluant was monitored, and UV-active fractions were analyzed by a negative ESI-mass spectrometry. Those containing UDP-4-ketoxylose were lyophilized to dryness. The lyophilized sugar (29 mg, 65% yield) was dissolved in 10 mL H2O and lyophilized again. This procedure was repeated twice more to yield the UDP-4-ketoxylose as its triethylammonium salt. 1H and 31P NMR spectra matched those in the literature. 1H NMR (D2O): δ 7.85 (d, 1H, J5,6 = 8.1 Hz, H-6); 5.90 (d, 2H, H-5, H-1’); 5.48 (dd, 1H, J1”,2” = 3.3 Hz, J1”,p = 7.0 Hz, H-1’”); 4.28 (m, 2H, H-2’, H-3’); 4.20 (m, 1H, H-4’); 4.10-4.18 (m, 2H, H-5’); 3.84 (d, 1H, J5a”,5b”=12 Hz, H-5b”); 3.72 (d, 1H, J2”,3” = 10.1 Hz, H-3”); 3.58 (dt, 1H, J1”,2” = 3.3 Hz, J2”,3” = 10.1, J2”,p = 3.1 Hz, H-2”); 3.84 (d, 1H, J5a”,5b”= 12 Hz, H-5a”). 31P NMR (D2O): δ -10.24 (d, JPa,pβ= 20.7 Hz), -12.10 (d, JPa,pβ= 20.7 Hz). ESIMS: m/z 533 [M-H] and m/z 551 [M+H2O-H].

3.9.12 Test of UDP-4-ketoxylose as a Potential Intermediate for AXS1

AXS1 (3 mg) was added to a solution of a sample of UDP-4-ketoxylose (4.43 mg, 8.3 μmol) dissolved in a solution of sodium phosphate buffer pH 8.0 (1.0 mL final volume). The reaction was incubated at 25 °C and monitored by UV/Vis spectroscopy. The mixture was
allowed to incubate for 24 h before the keto substrate was removed by centrifugal ultrafiltration. The resulting enzyme solution was dissolved back to 1 mL and another UV/Vis spectrum was acquired. NaBH₄ (0.03 mg, 750 nmol) in sodium phosphate buffer pH 8.0 (20 μl) was added to the enzyme solution. The UV/Vis spectrum was acquired immediately. In another experiment, AXS1 (202 mg) was added to a solution of UDP-4-ketoxylose (3 mg, 5.6 μmol) in solution of phosphate buffer pH 8.0. The reaction was incubated at 25 °C and was monitored by negative ESI-mass spectrometry.

3.9.13 AXS1 Activity Test with UDP-Xyl

A sample of AXS1 (100 mg) was added to a solution containing UDP-Xyl (3.0 mg, 5.6 μmol) and 1 mM NAD⁺ in 10 mM sodium phosphate/D₂O buffer (pD 7.4, 1.0 mL final volume). The reaction was incubated at 25 °C. ¹H, and ³¹P NMR spectra were acquired at timed intervals. The mixture was allowed to incubate for another several hours before the enzyme was removed by centrifugal ultrafiltration. A final ¹H NMR spectrum was acquired.

3.9.14 Attempted Synthesis of UDP-API 10

3.9.14.1. Methyl 3'-benzyl-2,3-O-isopropylidene-β-D-erythro-apifuranose 5

Methyl 2,3-O-isopropylidene-β-D-erythro-apiofuranose 4 was synthesized as a mixture of anomers according to procedure described in literature. Compound 4 (100 mg, 0.489 mmol) dissolved in dry DMF (5 mL) was added to a flask containing sodium hydride (32 mg) in DMF (10 mL). After the addition, the mixture was allowed to cool on ice for 10 min. The solution of benzyl bromide (140 mg, 100 μL) was added slowly over the course of 10 min. The reaction was allowed to stir for 5 min before being tightly sealed. After stirring for 16 h, the mixture was poured into 10 mL of ice cold water and extracted with 4 x 20 mL of ethyl acetate.
The solvent was evaporated under reduced pressure and the resulting syrup was redissolved in ether and washed with brine to remove excess DMF. The ether was evaporated under reduced pressure to obtain a crude syrup. Purification by silica gel column chromatography (9:1 petroleum ether/ethyl acetate) gave the compound 5 as a 20:1 mixture of α:β anomers (101 mg, 0.0343 mmol) in 70% yield. $^1$H NMR (CDCl$_3$) δ 7.30 (m, 5H, Ph); 4.91 (s, 1H, H-1); 4.63 (d, 1H, $J_{\text{PhCH}-1, \text{PhCH}-2}$ = 12.4 Hz, PhCH$_2$); 4.57 (d, 1H, $J_{\text{PhCH}-1, \text{PhCH}-2}$ = 12.4 Hz, PhCH$_2$); 4.27 (s, 1H, H-2); 3.99 (d, 1H, $J$ = 10.1 Hz, H-4 or H-3'); 3.85 (d, 1H, $J$ = 10.1 Hz, H-4 or H-3'); 3.67 (d, 1H, $J$ = 10.1 Hz, H-3' or H-4'); 3.62 (d, 1H, $J$ = 10.1 Hz, H-3' or H-4'); 3.29 (s, 3H, OCH$_3$); 1.47 (s, 3H, C(CH$_3$)); 1.38 (s, 3H, C(CH$_3$)). ESIMS: $m/z$ 317 [M+Na]$^+$. 

3.9.14.2. 3'-Benzyl-1,2,3-O-triacetyl-D-erythro-apifuranose 6

Compound 5 (100 mg, 0.034 mmol) was added to a mixture of water and Dowex 50 H$^+$ (300 mg, wet). The resulting mixture was heated at reflux for 16 h. The resin was removed by filtration and water evaporated under reduced pressure. The residue was dissolved in a mixture of pyridine (10 mL) and 4-dimethylaminopyridine (20 mg) and cooled on ice for 20 min. Acetic anhydride (0.356 mL, 3.78 mmol) was added, and the resulting mixture was stirred for 16 h at rt. The solvent was evaporated under reduced pressure, and the resulting syrup was dissolved in CHCl$_3$ (20 mL) and washed with 0.1 N HCl, NaHCO$_3$ and brine. The chloroform was evaporated under reduced pressure to obtain a crude syrup. Purification by silica gel column chromatography (3:7 petroleum ether/ethyl acetate) gave the compound 6 (76 mg, 0.021 mmol) in a 61% yield over 2 steps. Compound 6 was isolated a mixture of two anomers with β being the dominant form. $^1$H NMR spectrum matched that in the literature. ESIMS: $m/z$ 389 [M+Na]$^+$. 

170
3.9.14.3.  3'-Benzy1-D-erythro-apifuranose 12

The known compound 3\textsuperscript{164} was converted to known compound 11 with a modification of a previously reported procedure.\textsuperscript{166} Compound 3 (100 mg, 0.526 mmol) dissolved in dry DMF (5 mL) was added to a flask containing sodium hydride (40 mg) in DMF (10 mL). After the addition, the mixture was allowed to cool on ice for 10 min. A solution of benzyl bromide (287 mg, 194 μL) was added slowly over the course of 10 min. The reaction mixture was allowed to stir for 5 min before being tightly sealed. After stirring for 16 h, the mixture was poured into 10 mL of ice cold water and extracted with 4 x 20 mL of ethyl acetate. The solvent was evaporated under reduced pressure and the resulting syrup was redissolved in ether and washed with brine to remove excess DMF. The ether was evaporated under reduced pressure to obtain a crude syrup. The syrup was purified by silica gel column chromatography (9:1 Petroleum ether/ethyl acetate). Compound 11 (100 mg, 0.027 mmol) was dissolved in 80% formic acid (25 mL), and heated at 60 °C for 2 h. The solvent was evaporated under reduced pressure. Traces of acid were removed by co-evaporating with toluene under reduced pressure. Compound 12 (53.8 mg, 0.024 mmol, 83% yield) was isolated as 8:1 mixture of α:β anomers. \textsuperscript{1}H NMR (CHCl\textsubscript{3}) of 12-α: δ 7.32 (m, 5H, Ph); 5.24 (d, 1H, J\textsubscript{1,2}= 4.56 Hz, H-1); 4.56 (s, 2H, PhCH\textsubscript{2}); 4.01 (d, 1H, J=10.0 Hz, H-4 or H-3'); 3.93 (d, 1H, J\textsubscript{1,2}= 4.56 Hz, H-2); 3.86 (d, 1H, J=10.0 Hz, H-4 or H-3'); 3.53 (d, 1H, J= 9.5 Hz, H-3' or H-4); 3.50 (d, 1H, J= 9.5 Hz, H-3' or H-4); 1.6 (s, OH). ESIMS: m/z 263 [M+Na]\textsuperscript{+}.

3.9.14.4.  1,2,3,3'-Tetrabenzy1-D-erythro-apifuranose 13

Compound 12 (100 mg, 0.042 mmol) was dissolved in benzyl alcohol (25 mL) and a catalytic amount of HCl (300 μL). The mixture was heated to 70 °C and stirred for 16 h, at which time benzyl alcohol was removed by distillation at reduced pressure. The resulting syrup
dissolved in dry DMF (5 mL) was added to a flask containing sodium hydride (40 mg) in DMF (10 mL). After the addition, the mixture was allowed to cool on ice for 10 min. The solution of benzyl bromide (287 mg, 194 µL) was added slowly over the course of 10 min. The reaction was allowed to stir for 5 min before being tightly sealed. After stirring for 16 h, the mixture was poured into 10 mL of ice cold water and extracted with 4 x 20 mL of ethyl acetate. The solvent was evaporated under reduced pressure and the resulting syrup was redissolved in ether and washed with brine to remove excess DMF. The ether was evaporated under reduced pressure to obtain a crude syrup. Purification by silica gel column chromatography (7:1 petroleum ether/ethyl acetate) gave the compound 13 as a 12:1 mixture of α:β anomers (139 mg, 0.027 mmol) in 65% yield over three steps. ¹H NMR (CHCl₃) of 13-α: δ 7.37-7.21 (m, 20H, Ph); 5.25 (d, 1H, J₁₂= 2.41 Hz, H-1); 4.76-4.44 (m, 8H, PhCH₂); 4.17 (d, 1H, J= 10.3 Hz, H-4 or H-3'); 4.09 (d, 1H, J= 10.3 Hz, H-4 or H-3'); 3.69 (d, 1H, J= 2.4 Hz, H-2); 3.61 (d, 1H, J= 10.2 Hz, H-3' or H-4); 3.68 (d, 1H, J= 10.2 Hz, H-3' or H-4). ESIMS: m/z 533 [M+Na]⁺.

3.9.14.5. 2,3,3'-Tribenzyl-D-erythro-apifuranose 14

Compound 13 (100 mg, 0.020 mmol) was dissolved in 80% formic acid (25 mL), and heated at 60 °C for 2 h. The solvent was evaporated under reduced pressure. Traces of acid were removed by co-evaporating with toluene under reduced pressure. Compound 14 was isolated as a 10:1 mixture of α:β anomers (75 mg, 0.018 mmol, 91% yield). ¹H NMR (CHCl₃) of 14-α: δ 7.35-7.20 (m, 15H, Ph); 5.25 (d, 1H, J₁₂= 2.39 Hz, H-1); 4.71-4.41 (m, 6H, PhCH₂); 4.17 (d, 1H, J= 10.3 Hz, H-4 or H-3'); 4.08 (d, 1H, J₁₂= 10.3 Hz, H-4 or H-3'); 4.01 (d, 1H, J= 2.4 Hz, H-2); 3.68 (d, 1H, J= 10.1 Hz, H-3' or H-4); 3.68 (d, 1H, J= 10.2 Hz, H-3' or H-4). ESIMS: m/z 443 [M+Na]⁺.
3.9.14.6. Dibenzyl 2,3,3'-tribenzyl-\textit{D-erythro}-apifuranose phosphate 15

This procedure is taken from the literature\textsuperscript{184} and slightly modified. Dibenzyl \textit{N,N}-diethylphosphoramidite (2.50 mmol, 880\mu L) was added to a solution of compound 14 (100 mg, 0.023 mmol) and 1,2,4-triazole (72 mg, 4 mmol) in 5.0 mL dry CH\textsubscript{2}Cl\textsubscript{2}. The progress of the reaction was followed by ESI-mass spectrometry. After stirring at rt for 2 h all of the starting material was converted to product and 30 mL diethyl ether was added. The organic solution was washed with saturated sodium bicarbonate solution (3 x 15 mL) and NaCl brine (3 x 10 mL). The organic layer was dried over sodium sulphate and evaporated to give an oil under reduced pressure. The oil was dissolved in 9 mL THF and cooled to −78 °C, then 30\% H\textsubscript{2}O\textsubscript{2} (1.9 mL) was added and the solution was allowed to warm to rt over 2 h. Diethyl ether (30 mL) was added and the organic solution was washed with saturated sodium bicarbonate solution (3 x 15 mL) and NaCl brine (3 x 10 mL). The organic layer was dried over sodium sulphate and evaporated to give an oil under reduced pressure. Attempts to purify compound 15 by column chromatography were unsuccessful and no desired product was isolated from this procedure. ESIMS: \textit{m/z} 703 [M+Na]\textsuperscript{+}.

3.9.15 Chemoenzymatic Synthesis of UDP-2F-GlcA 22

\textit{α-1-Phospho-2-deoxy-2-fluoroglucose} 20 was synthesized according to literature described procedures.\textsuperscript{172,175} A variation from these procedures was in the phosphorylation of 19, forming compound 20 (Figure 3.22). The neat mixture was heated to 65 °C to ensure full conversion to the desired product. Compound 20 (25 mg, 0.096 mmol) was dissolved in 70 mM Tris-HCl, pH 7.8, (40 mL) containing UDP-glucose pyrophosphorylase (25.5 units), UTP (200 mg) and MgSO\textsubscript{4} (72 mg, 292 \mu mol). The incubation was carried out at 30 °C for 12 h. The
progress of the reaction was monitored by $^{31}$P NMR spectroscopy and ESI-mass spectrometry. Once the reaction was complete, the enzymes were removed by centrifugal ultrafiltration and the remaining solution was loaded onto a DE-52 column (220 mL) and eluted with a linear gradient of 0 - 400 mM triethylammonium bicarbonate buffer (800 mL total). The triethylammonium bicarbonate buffer was prepared by bubbling CO$_2$ into 400 mM triethylamine solution over 10 h. The separation was monitored by an UV detector (Spectra/Chrom Flow Thru UV monitor controller, Spectrum) with a 254 nm filter. The fractions containing UDP-2-deoxy-2-fluoro-D-glucose 21 were lyophilized to dryness. Compound 21 was then dissolved in 50 mM phosphate buffer (10 mL), pH 8.0, containing 5 mM NAD$^+$, 2 mM DTT, and freshly prepared UDP-glucose dehydrogenase (700 mg of crude protein extract/mL of lysis buffer), and incubated at 37 °C for 3 h. The progress of the reaction was followed by ESI-mass spectrometry. The sugar was purified by a DE-52 column as described above, and fractions containing compound 22 were lyophilized to dryness. $^1$H-NMR (D$_2$O): δ 7.87 (d, 1H, $J_{5,6}$ = 8.1 Hz, H-6); 5.92-5.87 (m, 2H, H1', H5); 5.70 (dd, 1H, $J_1''$,$p_b$ = 7.98 Hz, $J_{1''}$$,2''$ = 3.51 Hz, H1''); 4.45, 4.32 (m, 1H, H-2''); 4.28-4.07 (m, 5H, H-2',3',4',5'); 4.06. (d, 1H, $J_{4',5'}$ = 10.2 Hz, H-5''); 3.94 (m, 1H, H-3''); 3.45 (t, 1H, $J_{3''}$$,4''$ = 9.7, $J_{3''}$$,5''$ = 10.2 Hz, H-4''). $^{31}$P NMR (D$_2$O): δ ~9.17 (d, $J_{Pa,pb}$ = 19.7 Hz), ~10.97 (d, $J_{Pa,pb}$ = 9.7 Hz). $^{19}$F NMR (D$_2$O): δ ~201.18 (s). ESIMS: $m/z$ 581.1 [M–H]$^-$.  

3.9.16 AXS1 Incubation with UDP-2F-GlcA

A sample of AXS1 (75 mg) was added to a solution of UDP-2F-GlcA (3.0 mg, 5.0 μmol) dissolved in solution of sodium phosphate buffer pH 8.0 and 1 mM NAD$^+$ (1.0 mL final volume). The reaction was incubated at 25 °C. $^{31}$P and $^{19}$F NMR and negative ESI-mass spectra were acquired at timed intervals. After 24 h, all the starting material was converted to UDP-2F-Xyl ($m/z$ 537 [M-H]$^-$) and the enzyme was removed by centrifugal ultrafiltration. The resultant
filtrate was loaded onto a DE-52 column (220 mL) and eluted with a linear gradient of a 0 to 0.5 M triethylammonium bicarbonate buffer. The $A_{254}$ of the eluant was monitored, and UV-active fractions were analyzed by negative ESI-mass spectrometry. Those containing UDP-4-ketoxylose were lyophilized to dryness. The lyophilized sugar was dissolved in 10 mL H$_2$O and lyophilized again. This procedure was repeated twice more to yield the UDP-2F-Xyl as its triethylammonium salt. $^1$H-NMR (D$_2$O): $\delta$ 7.86 (d, 1H, $J_{5,6} = 8.1$ Hz, H-6); 5.89-5.82 (m, 2H, H1', H5), 5.61 (dd, 1H, $J_{1',2''} = 7.38$ Hz, $J_{1'',2''} = 3.51$ Hz, H1''); 4.38, 4.21 (m, 1H, H-2''); 4.28-4.07 (m, 5H, H-2',3',4',5'); 3.86 (m, 1H, H-3''); 3.63-3.51 (m, 2H, H-4'', H-5''); 3.43 (m, 1H, H-5''). $^{31}$P NMR (D$_2$O): $\delta$ -9.27 (d, $J_{PA,PP} = 19.6$ Hz), -10.87 (d, $J_{PA,PP} = 19.6$ Hz). $^{19}$F NMR (D$_2$O): $\delta$ -201.36 (s) ESIMS: m/z 537.1 [M–H]$^-$. 

3.9.17 Chemoenzymatic Synthesis of UDP-3F-GlcA 28

3-Deoxy-3-fluoro-D-glucose was prepared according to a previously published synthetic route. A variation in this procedure was in installation of the fluorine group to form compound 26 (Figure 3.26). Compound 25 (400 mg, 1.53 mmol) was dissolved in 6 mL dry CH$_2$Cl$_2$ in a flame-dried flask and the solution was chilled under argon to $-40$ °C in a dry ice/acetonitrile bath. DAST (0.19 mL, 1.4 mmol) was added slowly via syringe and the solution was stirred at $-40$ °C for 1 h, and then allowed to warm to rt over 5 h. The solution was then cooled to $-10$ °C and 2.5 mL of methanol was added. The solvent was then removed under reduced pressure and the crude mixture was separated by silica gel chromatography. Silica gel column chromatography (9:1 petroleum ether/ethyl acetate) yielded 195 mg (48% yield) of 1,2:5,6-di-O-isopropylidene-3-deoxy-3-fluoro-D-glucose 26. Compound 28 was prepared from compound 27 by the conditions described in Section 3.9.15. $^1$H-NMR (D$_2$O): $\delta$ 7.87 (d, 1H, $J_{5,6} = 8.1$ Hz, H-6); 5.92-5.87 (m, 2H, H-1', H-5), 5.70 (m, 1H, H-1''); 4.52 (dt, 1H, $J_{2'',3''} = J_{3'',4''} = 175$
9.22 Hz, $J_{3', F} = 54.1$ Hz, H-3'); 4.28-4.07 (m, 5H, H-2',3',4',5'); 4.06 (d, 1H, $J_{4', 5'} = 10.34$ Hz, H-5'); 3.80-3.64 (m, 2H, H-2'', H-4''). $^{31}$P NMR (D$_2$O): $\delta$ -10.75 (d, $J_{Pa, P\beta} = 20.4$ Hz), -12.67 (d, $J_{Pa, P\beta} = 20.4$ Hz). $^{19}$F NMR (D$_2$O): $\delta$ -200.993 (s) ESIMS: $m/z$ 581.1 [M-H].

### 3.9.18 Chemoenzymatic Synthesis of UDP-[U-13C]-3F-GlcA

UDP-[U-13C]-3F-GlcA was prepared from [U-13C]-D-glucose using the same procedure that was used to prepare UDP-3F-GlcA 28. $^{13}$C NMR (D$_2$O): $\delta$ 175.1 (d, $J_{C-5', C-6'} = 60.3$ Hz, C-6'); 94.8 (d, $J_{C-1', C-2'} = 51.39$ Hz, C-1'); 93.07 (dt, $J_{C-2', C-3'} = J_{C-3', C-4'} = 43.2$ Hz, $J_{C-3', F} = 83.2$ Hz); 71.81 (m, C-2'); 69.54 (m, C-4'', C-5''). ESIMS: $m/z$ 587.1 [M-H].

### 3.9.19 AXS1 Incubation with UDP-3F-GlcA and UDP-[U-13C]-3F-GlcA

A sample of AXS1 (125 mg) was added to a solution of UDP-3F-GlcA (3.0 mg, 5.0 µmol) dissolved in sodium phosphate buffer pH 8.0 and 1 mM NAD$^+$ (1.0 mL final volume). The reaction was incubated at 25 °C. $^{31}$P and $^{19}$F NMR and negative ESI-mass spectra were acquired at timed intervals. Once all the starting material was converted to UDP and free fluoride, the enzyme was removed by centrifugal ultrafiltration and $^{31}$P/$^{19}$F NMR spectra were acquired again. A similar procedure was followed with UDP-[U-13C]-3F-GlcA and the reaction was monitored by $^{13}$C NMR spectroscopy.
Bibliography


177


Appendix
Figure A.1: H NMR spectrum of 2,4-diacetamido-2,4,6-trideoxymannose (300 MHz, D$_2$O, 25°C).
Figure A.2. $^1$H NMR spectrum of UDP-4-ketoxylose (400 MHz, D$_2$O, 25°C).
Figure A.3: $^1$H NMR spectrum of 2,3,3'-tribenzyl-D-apiose 14 (400 MHz, CDCl$_3$, 25°C).
Figure A.4. $^1$H NMR spectrum of UDP-2-deoxy-2-fluoro-D-glucuronic acid as a triethylammonium salt 22 (400 MHz, D$_2$O, 25°C).
Figure A.5 Partial $^1$H COSY spectrum of UDP-2F-Xyl (400 MHz, D$_2$O, 25 °C).
Figure A.6 $^1$H NMR spectrum of UDP-3-deoxy-3-fluoro-D-glucuronic acid as a triethylammonium salt 28 (400 MHz, D$_2$O, 25 °C).
Figure A.7 $^{13}$C NMR spectrum of UDP-[U-$^{13}$C]-3-deoxy-3-fluoro-D-glucuronic acid as a triethylammonium salt (150 MHz, D$_2$O, 25 °C).